

UVM ScholarWorks

Cross-Linguistic Pitch Differences in English and French Bilinguals: Timothée Chalamet and Lily-Rose Depp

Item Type	undergraduate thesis
Authors	Bellino, Kate Isabella
Download date	2026-05-09 13:43:41
Item License	http://creativecommons.org/licenses/by-nc-nd/3.0/
Link to Item	https://hdl.handle.net/20.500.14849/5251

CROSS-LINGUISTIC PITCH DIFFERENCES IN FRENCH AND ENGLISH BILINGUALS:
TIMOTHÉE CHALAMET AND LILY-ROSE DEPP

Kate Bellino

A thesis submitted to the faculty at The University of Vermont in fulfillment of the requirements
for honors with a degree of Bachelor of Arts in the Department of Linguistics

University of Vermont
Spring 2020

Table of Contents

Abbreviations	3
List of Figures	3
List of Tables	3
Abstract	4
Introduction	5
Overview of Pitch	5
Pitch and Gender	6
Children.....	6
Adults	8
Pitch Differences Across Languages	10
Pitch Differences in Bilinguals	12
Methodology	14
Material	14
Informants	18
Data Collection	19
Results	24
Median F0	24
Mean F0	25
F0 Variance	28
Discussion and Conclusion	29
Future Directions	33
Acknowledgments	34
Bibliography	35

Abbreviations:

F0	Fundamental Frequency
Hz	Hertz
TCE.....	Timothée Chalamet English
TCF.....	Timothée Chalamet French
LRDE	Lily-Rose Depp English
LRDF.....	Lily-Rose Depp French
IQR.....	Interquartile Range

Figures:

1. Spectrogram of Timothée Chalamet speaking English with force alignment.....	6
2. Creaky voice when Depp says 'love'	17
3. Spurious F0 value manually corrected in Depp's English speech in 'thirteen'.....	22
4. Improper automatic forced alignment in the word 'drugs' for Chalamet in English	23
5. Chalamet F0 across each word in each language	25
6. Depp F0 across each word in each language	26
7. Mean F0 for Chalamet and Depp compared in English and French	28
8. Boxplot F0 range by informant and language	29

Tables:

1. Median pitch values in semitones by informant and language	24
2. Paired sample statistics for Chalamet	26
3. Paired samples test for Chalamet.....	27
4. Paired sample statistics for Depp.....	27
5. Paired samples test for Depp	27

**Cross-Linguistic Pitch Differences in English and French Bilinguals:
Timothée Chalamet and Lily-Rose Depp**

By Kate Bellino

Abstract

Pitch differences between men and women are partially linked to differences in average larynx size. Men have been shown to have a lower pitch range than women due in part to physical factors such as longer larynges. However, even with a natural difference in average pitch in men and women, social factors in various cultures can affect this natural pitch difference, either helping to distinguish men from women or shrinking this gap. Bilinguals have the potential to be active participants in the multiple cultures their fluency allows them to; therefore, I am examining if bilinguals manipulate their average pitch depending on the language they are speaking reflective of social factors. This study looks specifically at two English/French bilingual actors, Timothée Chalamet and Lily-Rose Depp. This study measured their fundamental frequency (F0) using interviews in English and French. The F0 was found to be higher and to have a smaller range (in semitones) in French than English for both informants. These differences in F0 demonstrate that there is a pitch difference between these two bilingual informants when speaking English and French, and future research could analyze if this expands to a larger sample size of English/French bilinguals.

Introduction

Overview of pitch

Vocal cord vibration in the larynx produces speech. This vibration results in changes in air pressure that vibrate off of the eardrum of a listener. These vibrations can be viewed in the form of a sound wave where the amplitude of the wave is the intensity or volume, while the wavelength correlates with the frequency of the vibration of the vocal cords. This is also known as tone or pitch. Lengthening or shortening the vocal cords, thus causing an increase or decrease in tension of the cords, can also alter the pitch, either raising or lowering it. The vocal cords only vibrate for voiced phonemes, therefore pitch cannot be measured for any voiceless phonemes and is primarily measured for vowels. Multiple vibrations occur at once, which results in harmonics. The slowest and strongest vibration is the first harmonic, which is known as the fundamental frequency (F0). These various harmonics can be graphed in the form of spectrograms (see Figure 1), which plot time on the x-axis and place frequency on the y-axis. The higher the frequency, the higher the F0 value will be. The F0 is the harmonic that is analyzed because it is the perceived pitch of a speaker. The graph below demonstrates a sound wave adjacent to its spectrogram counterpart lined up with each phoneme (sound). The blue dashed line indicates pitch in Hertz.

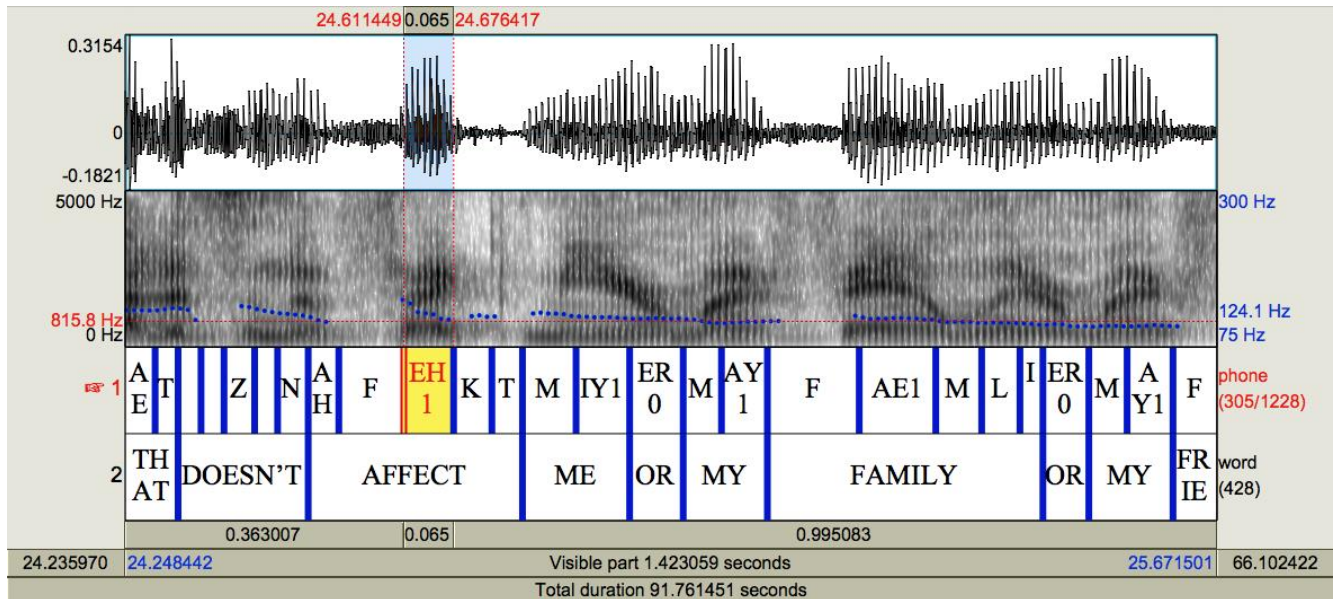


Figure 1. Spectrogram of Timothée Chalamet speaking English with forced alignment.

Pitch and gender

A. Children

Previous research on babies and children of different sexes¹ found that boys and girls have similar larynx sizes before puberty, as hormones have not yet affected the length of the larynx. Babies should then have an F0 that is independent of their sex as their larynges have not yet been affected by changes in hormones. In a study done by Lieberman (1967), a 10-month-old baby altered his pitch when babbling depending on if he was in the presence of his mother or father. When alone, his pitch was 430 Hz, while with his mother it was 390 Hz, and with his father 340 Hz. A 13-month-old girl also had a variation in her pitch, ranging from 290 and 390 Hz, depending on if she was interacting with her mother or father. As there are not yet physical differences in larynges between males and females at this age, this demonstrates that there are acquired pitch manipulation strategies based on whom a speaker is interacting with. Also,

¹ When talking about physical differences in men and women, I am referring to the sexes of male and female. Gender is then a personal identification of how individuals want to present themselves or which gender they identify with. This study is assuming that the participants in these studies are all cisgender and that there are only two sexes and two genders. This gender binary is problematic for many reasons, especially for the fact that there are more than these two sexes and genders. This study simplifies and excludes many people in society. The goal is not to do this, but to be a basis for exploring how gender identity can affect pitch, as it is another way individuals can present themselves. The ultimate goal is that future research can expand from this and be more inclusive of other identities.

prepubescent girls were found to speak with a higher F0 than the boys, and vice versa, even though the lengths of their larynges were still the same size (Eckert & McConnell-Ginet, 2003). Significant differences in pitch were found at various age levels (between 6 and 10) for various speech tasks for males, not females, which could indicate a cultural influence was present that may have caused male children to lower their F0 (Sorenson, 1989). Perry et al. (2001) found that although pre-pubertal children of both genders speak with similar F0s, adults can correctly identify a child's gender starting at the age of four. It is not until puberty when both hormonal changes and physical developments occur, that differences in pitch biologically begin to develop between the genders. An individual can manipulate their F0 but are constrained by the mass and dimension of their vocal cords (Graddol & Swann, 1983). Raising or lowering the larynx, tightening or slackening the vocal folds, along with adjusting the placement of the jaw can change the F0 (Sagisaka et al., 2012). This biomechanical ability to control pitch is on a neuromuscular level through the contraction of the cricothyroid—the muscle that aids in phonation in the larynx— and vocalis muscles—a part of the larynx where the vocal cords are attached (Hong et al., 1998; Zhang, 2016). These physical adjustments give a prepubescent child the ability to personalize and adjust their overall pitch before physical changes have occurred that would cause the biological differences in pitch between males and females. This demonstrates how the differences in pitch at this age can be attributed to the various social and gender norms that are imposed on each group, not differences in sex. This can be carried into adulthood, resulting in pitch differences in adults to be attributed to not only physical differences (differences between sexes) but also from social and gender norms learned during childhood.

B. Adults

Studies have been done to look at if body size affects pitch. In the animal kingdom, typically, the larger the animal, the lower the pitch the animal will have. Studies have been done to see if physical size differences have an effect on pitch within humans. Evans et al. (2006) found that there was a significant negative relationship in males between their measurements of body size and weight with their F0. However, van Bezooijen (1995) found that body size differences among women were not statistically significant. Based on Graddol and Swann's findings in (1983), they found that body build is not related to a woman's F0, while the men's pitches reflected the physical sizes of their larynges. Body size or build did not affect a woman's F0, but pitch differences were still found amongst groups of women, indicating that there are other factors present that influence pitch. Since body size is not a substantial factor, unlike in men, cultural differences might contribute to the differences in pitch for women's speech more than men's (Graddol & Swann, 1983).

Graddol and Swann (1983) argue that one influence of a speaker's F0 is physical and anatomical disposition. A speaker can manipulate their F0 but is limited by the constraints of the dimensions of the vocal cords. This manipulation results in muscular tension; thus, specific frequencies are more difficult or uncomfortable, which can lead to laryngeal abuse; therefore there is a preferred F0 range per speaker. The size of the vocal tract can also aid in ease as particular frequencies resonate better in a particular size vocal tract. This shows that a speaker has a preferred F0 range that is based on their physical components in their vocal tract. This aids in demonstrating that differences in pitch in bilinguals are in part socially influenced because there should be a theoretical natural pitch range for an individual. Additionally, Graddol and Swann (1998) cite studies in which pitch is also influenced by factors such as menstrual cycles,

the singing voice of an individual's ancestors, the volume of the testis, and whether an individual had sexual intercourse recently. This all results in different typical pitch ranges for men and women, demonstrating that pitch differences between these two groups are attributed to various physical reasons.

In adults, women are perceived to have higher pitches than men, which correlate with higher F0 values. In a study that investigated correlations between glottal waveform parameters and the perception of particular voice types, Karlsson (1988) found that voices perceived as higher pitch had high F0 values, demonstrating how pitch plays a vital role in identifying the sex of a speaker. F0 and formant spacing reflect the vocal apparatus' anatomical dimorphisms that aid in sex identification (Cartei & Reby, 2013). In addition, Whiteside (1988) found that a speaker's sex could accurately be identified from brief vowel segments. This signifies that there are strong perceived distinctions between male and female voices that can require only milliseconds' worth of speech. In addition to determining a speaker's sex based on of their F0, personality judgments are made about a speaker based on of their F0, regardless of if the speaker possesses those personality traits. This phenomenon called is called "voice stereotypy" (Aronovitch, 1976, p. 208). Aronovitch (1976) found that particular vocal characteristics have a significant relationship to many personality traits with American college students as both speakers and listeners. For women— kindness, emotionality, and humor were found to have a negative correlation with F0 mean. Maturity has a positive correlation with F0 mean. Self-doubt and submissiveness were found to have a positive correlation with F0 variance in males. For men, significant correlations were found in regards to the degree of variability of F0 while for women significant correlations were found for average F0. Since American men and women were assigned personality traits based off of two different manners—F0 variation and mean F0—

this reflects that there are different cultural stereotypes between American men and women or else there would have significant correlations would have been found for the same manners.

Pitch differences across languages

Males and females have physical differences that affect their pitch, but people from various cultures do not necessarily have physical differences that would result in varying natural pitch ranges depending on the languages. Although this is the case, various studies have shown that there are different pitch ranges between languages, which demonstrate that there are more factors than just sex or gender. In a study done by Majewski et al. (1972), adult Polish males were found to have an average pitch of 137.6 Hz while American males were found to have a lower average pitch of 118.9 Hz. Size was determined not to be a contributing factor. Another study looking at U.S. male university students found their average pitch to be 123 Hz, still lower than Polish males (Hollien & Jackson, 1972). Comparing Tagalog, Spanish, Japanese, and English, English speakers had a lower F0 than Japanese and Spanish speakers but higher F0 than Tagalog (Hanley et al., 1966). In a study by Andrianopoulos et al. (2001), Mandarin Chinese male speakers were found to have a higher F0 than standard American English and Hindi Indian male speakers. This was also the case with female Mandarin Chinese speakers, who were found to have higher F0 values than American English and Hindi Indian female speakers. This is similar to the findings of Keating and Kuo (2012) that Mandarin Chinese speakers have higher F0 values and have more extensive F0 ranges compared to English speakers.

Some of these cross-linguistic pitch differences can be attributed to differing gender stereotypes across societies, such as a culture's idea of masculinity and femininity. In a study done by van Bezooijen (1995), Japanese women were found to use a higher range of F0 than

Dutch women. Loveday (1981) similarly found Japanese women to speak at a considerably higher pitch than (British) English females. van Bezooijen's study is the only one that focused on analyzing pitch differences in conjunction with gender stereotypes. They first had individuals rate their voices on scales that included short/tall, weak/strong, dependent/independent, and modest/arrogant. This study found that in Dutch society, the differences between men and women differed in half a point, while the Japanese had an average difference of one point and a half (on a seven-point scale). Both of these findings were significant differences between men and women. A correlation between pitch and gender stereotypes can be constructed from these results. This survey also demonstrated that there is an ideal pitch range for men and women in each culture. The study found that higher pitches are associated with powerlessness in both Dutch and Japanese cultures, although the differentiation between men and women is much stronger for the Japanese. In Japan, there is a preference for higher pitches in women, while in the Netherlands, the preference is for medium or low pitches in women. These differences reflect contrasting cultural views of how pitch contributes to ideal masculinity or femininity in their society; in Japanese society, pitch appears to be a reliable marker of gender identity, evidenced by the amplification of difference between women and men when compared to Dutch speakers. In other words, this study demonstrated that the ideal man and woman are constructed as extremes from each other in Japanese society as opposed to a more neutral difference among the Dutch. Their third hypothesis was also found to be true, in which the Dutch, both men and women, found low to medium pitches more attractive in women, whereas the Japanese found medium to high pitches to be more attractive. Based on this study and in conjunction with the findings earlier that children manipulate their pitch, pitch is clearly a paramount resource that people make use of to orient to gender norms.

In summary, although there are no physical differences between people from around the world that would affect pitch, each culture has its own set of social norms that can affect how pitch is manipulated in order to adhere to these norms. All of these studies demonstrate that there can be overall F0 differences from language to language. Also, there may be social influences affecting the overall pitch in one language over the other, but what if a person speaks two of these languages? Can a person exhibit two different fundamental frequency averages based on the language they are speaking, or is fundamental frequency user-dependent? Since bilinguals are connected to two different societies through being fluent in two languages, it is possible they could either rely more on one culture's ideologies, blend the two, or switch between the two depending on which language is being spoken.

Pitch differences in bilinguals

Some previous studies have looked at F0 averages in bilinguals. In Passoni et al. (2018), Japanese/English bilinguals were found to speak with a higher fundamental frequency when speaking English than when speaking Japanese, although this could have been attributed to the fact that a particular speaker mentioned how they “hates[sic] that female way of speaking in Japanese.” However, no further discussion on gender as a factor was discussed. Previous studies on these languages have not yielded a consistent pattern: Graham (2015) found that Japanese/English bilinguals spoke with a higher fundamental frequency when speaking Japanese, while studies done by Loveday (1981) and Ohara (1999) found that only female Japanese/English bilinguals had a higher fundamental frequency when speaking Japanese. Cheng (2020) investigated Korean/English bilinguals and concluded that Korean was spoken with a higher fundamental frequency than English. Cox (2010) similarly found that with

Spanish/English bilinguals, Spanish was spoken with a higher fundamental frequency than English. No discussion was made for gender as a factor in these studies. Lee and Sidtis (2017) found that bilinguals spoke Mandarin or Korean with a higher F0 than when speaking English. Their study mostly focused on the effects on pitch of performing various speech tasks. They found that people spoke with a higher F0 when performing reading tasks regardless of language; therefore, when analyzing F0, measurements from reading tasks are not as reflective of natural speech as other tasks may be. They also briefly discussed that in Korean and Japanese societies, women are expected to exhibit a certain level of politeness and femininity, which may be portrayed through F0. These studies demonstrate that an individual can have varying mean fundamental frequencies based on the language spoken. It is also important to note that differences in bilingual fundamental frequencies were found within two Western languages, which arguably have more similar cultural views on gender norms when compared to an Eastern language and Western language (and thus the two cultures associated with those languages). Even without a stark contrast between cultures or language families, pitch differences are significant.

The current study tests F0 level and span in English and French using two bilingual celebrities, one male and one female. Building on the findings comparing English and Spanish (Cox, 2010; Hanley et al., 1966), I expect to find that French is spoken with a higher F0 mean than English for both the male and female speaker. Furthermore, I predict that the female bilingual speaker will speak with a smaller pitch range than the male bilingual, as women have been found to speak with a smaller pitch range than men (Graddol, 1986; Traunmüller & Erikson, 1995).

Methodology:
Material

The two bilingual speakers this study will focus on are the actors Timothée Chalamet and Lily-Rose Depp. These actors were chosen due to their similar age, comparable popularity (judged by the number of Instagram followers), both having been immersed in both countries where French and English are spoken, and both having one parent who is a native speaker of English and one who is a native speaker of French. The two both star in the Netflix original movie, *The King*, which was released November 1, 2019. Although Chalamet and Depp are not the most well-recognized actors, their work is being streamed on the largest subscription streaming service, and each actor has earned numerous nominations and awards.

Available speech data from these celebrities are either from interviews or films. When comparing speech from interviews with speech from films, interviews will be more authentic as the speech is not rehearsed and is therefore spontaneous and more natural. Previous research has found that spontaneous speech tasks, when compared to other tasks, do result in more natural productions of speech that make up a speaker's voice profile. A voice profile includes rate, production levels, and pitch (Kreiman & Sidtis, 2011). Specifically related to pitch, the average F0 in spontaneous speech samples from children, young adults, older men, and adult trained singers is significantly higher than in reading tasks (Hudson & Holbrook, 1982; Mysak, 1959; Ramig & Ringel, 1983; Sorenson, 1989). For this study, in order to best emulate spontaneous speech, data was collected from interviews found online. Two interviews for each actor were included (one English and one French) so that there are a total of four interviews analyzed.

The interlocutor for the French interviews is Yann Barthès, a journalist and the host of *Quotidien*, a French television talk-show program that airs Monday through Friday. At the moment this was written, their YouTube channel has 160,332,199 views. Segments on the show

include news segments, interviews, and comedy. Using the same interlocutor and source of media ensures that the informants are in similar speech environments. The interlocutor, the media outlet, the general topics of conversation, and the type of audience are the same.

For the English interviews, the same interlocutor could not be found, but the same media outlet and format of interviews were found. Both interviews are taken from *W Magazine's* YouTube channel, which is an American magazine that has stories about culture, fashion, art, celebrities, and film. At the time of writing, their YouTube channel had 293,109,653 views.

Different researchers have used a range of methods for measuring pitch. Baken and Orlikoff (2000) determined that 14-second samples are sufficient for measuring speaking fundamental frequency accurately. Jassem (1971) demonstrated that values within two standard deviations of the mean F0 of 50 seconds of spoken data represented 95 percent of all observations, thus giving an accurate portrayal of a person's "compass" (Jassem, 1971, p. 59) of the speaking voice. Other researchers adopted this method of using more than 14-second samples (Baken & Orlikoff, 2000) when analyzing female speakers' F0 range: Gilbert and Weismer (1974) required 30 seconds of continuous speech from informants, de Pinto and Hollien (1982) used around 1 minute worth of data from reading a passage, and Henton (1989) used at least 1.5 minutes worth of continuous speech data. According to Nolan (1983), 40 seconds of speech is sufficient to gain an accurate portrayal of a speaker's range of variation. Past 60 seconds, any additional variation not already present is unexpected. In a study looking at pitch differences in Black students versus Caucasian students in the US, Gelfer and Denor (2014) collected 30-60 seconds of continuous speech so that 30 seconds of that could be measured for each of their informants. An average of 36 seconds was used. Another study looking at F0 of young Black adults used 40 seconds (Hudson & Holbrook, 1982). This is based on the average time it would

take an adult to read the Rainbow Passage (Fairbanks, 1960), a reading passage that is often used in acoustic voice research. The approximate total length of conversation that data was collected from in a study by Yuasa (2008) was 10 minutes, of which approximately one minute was used for pitch range analysis. In yet another study, which investigated the perception of distinct voices with glottal waveform parameters and glottal airflow, data was collected from excerpts from novels read out loud that were 45 to 70 seconds in length (Karlsson, 1988).

In studies that have focused on bilingual speakers, there is a similar variation in the methodology applied. Passoni et al. (2018), rather than using a particular length of speech, extracted F0 values from 20 sentences per language per person and measured the F0 min and F0 max across each sentence. Cox (2010) took F0 measurements from spontaneous speech of informants describing a narrative and used three segments from English and three segments from Spanish for each of the four participants, with each segment averaging 3 seconds in length. Cheng (2020) collected 10-20 minutes worth of spontaneous speech through interviews for Korean and 20-30 minutes worth of spontaneous speech for English. F0 measurements were taken every 5 ms and averaged across the entire duration of each vowel. The average F0 value was determined from using the average F0 across the stressed vowels in English while the F0 was averaged from measurements of every vowel for Korean.

In all these aforementioned studies that analyzed average F0 values between various populations, between 30 seconds to 90 seconds worth of continuous speech was used for the analysis (de Pinto and Hollien, 1982; Gelfer and Denor, 2014; Gilbert & Weismer, 1974; Henton, 1989; Hudson & Holbrook, 1982; Karlsson, 1988; Yuasa, 2008). Cheng (2020) used a much longer speech sample than other studies because he also analyzed F0 over the length of the interview to see if time affected F0. Time was found not affect F0. The total length of speech

used is congruous with previous research that claimed at least 14 seconds of speech was necessary (Baken and Orlikoff, 2000) and that 50 seconds of spoken data represents 95 percent within two standard deviations of all speech observations of an individual (Jassem, 1971). Based on these previous findings, 90 seconds of usable continuous speech, was collected per interview for this study. This study did not investigate the effect of time in conjunction with F0 like in Cheng (2020); therefore, the timeframe chosen corresponded with the majority of other studies that analyzed F0 (Gilbert & Weismer, 1974; de Pinto and Hollien, 1982; Henton, 1989; Gelfer and Denor, 2014; Hudson and Holbrook, 1982; Yuasa, 2008; Karlsson, 1988). 90 seconds, as opposed to the minimum of 30 seconds, was chosen to allow for more data points.

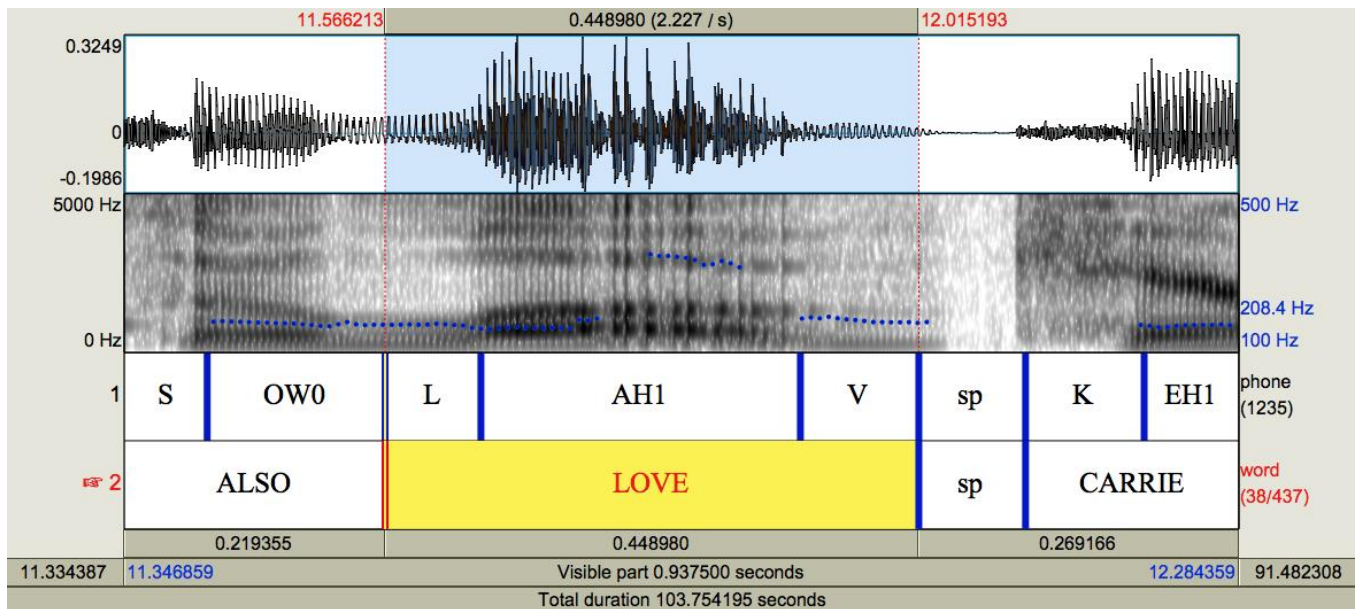


Figure 2. Creaky voice when Depp says ‘love’.

Other speech variations can occur, such as falsetto and creak, that could substantially raise or lower a speaker’s pitch and may be absent from a minute’s worth of speech. Creaky voice or vocal fry are low-frequency glottal pulses with a particular level of aperiodicity that can be seen in Figure 2 above. Creaky voice occurs predominantly at pitches below the modal

register (Hollien et al., 1966). In addition, voice quality does not determine the meaning of words in English, therefore there are no internal constraints on the use of creaky voice. This allows for an English speaker to utilize this pitch manipulation strategy as a stylistic effect. Some studies have shown that there is an increase in use among young adults, particularly among young American women (Dallaston and Docherty, 2020). This increase in creaky voice could lower a speaker's overall pitch. In addition, it could be used as a tool to narrow the gap between men and women's pitch. Despite popular stereotypes regarding gendered speech, in English, contradictory of when expressed in Hertz, in semitones women have a smaller pitch range than men (Graddol, 1986; Traunmüller and Erikson, 1995). It is only when factoring in the perceived "liveliness of speech" (p.1) or "prosodic explicitness" (p.4) that women then have a larger pitch span than men. The contrastive stress, placement of focus, and emotions all affect the perceived liveliness of speech. Surprise, interest, joy, contempt, and anger all are reflected in an increase of F0 variation, while little F0 variation is reflected by emotionally depressed, sad, or ashamed speakers. Traunmüller and Erikson (1995) used four classifications for lively speech. An increase in pitch span for women can be attributed to the use of falsetto, which in English is more socially acceptable for women to use and could be considered more lively speech. Once falsetto is omitted from speech or data, women have a smaller pitch range than men ('t Hart, et al. 1990). Pitch range is interesting to examine because it is a variable that allows a speaker to further customize and manipulate their voice in order to further express themselves.

Informants

Timothée Chalamet, at the time of the study, is a 24-year-old actor who was born in New York City to a French father and an American mother. He has 5.5 million followers on

Instagram. He is fluent in Standard English and French and spent his summers growing up in a small French village while attending school in New York. As a result, Chalamet claims that his time completely immersed in French culture resulted in cross-cultural identity issues and “a little bit of ambiguity in the self-identity sense, which helps a lot creatively because I don’t feel as constricted by who I am. I almost don’t really know the answer to that” (Marotta, 2017). He has been nominated for an Academy Award – making him the youngest best-actor nominee since 1939, three BAFTAs, two Golden Globes, and four Screen Actors Guild Awards. Most of Chalamet’s work is in English, but there are instances in which he speaks French in some of his films such as in *Call Me by Your Name* and *The King*.

Lily-Rose Depp, at the time of the study, is a 20-year-old actress who was born in France to an American father (Johnny Depp) and a French mother (Vanessa Paradis). Depp has 3.5 million followers on Instagram, which is the only form of social media she uses. She is fluent in Standard English and French, and up until 2012, Depp split her time between Los Angeles and France. Since then, she has spent most of her time stateside. Depp has been nominated for two César awards (France’s highest film honor). Depp is in both Anglo and Franco productions– *The Planetarium* (2016) was the first time she read a script entirely in French (Sagansky, 2016), and her future films are both in English and French.

Data Collection

Three of the interviews were downloaded using the website ytmp3.cc/en13/, which converts YouTube videos into MP3 files. The fourth interview had to be downloaded using savethevideo.com since it was not available on YouTube. The MP3 files were converted from stereo track to mono track and converted into wav files using the program Audacity. Following

the manner of many sociolinguistic interviews, the pure continuous speech by the informant was taken from the middle 80% of the total interview allowing for the informant's heightened awareness of being recorded to be minimized (Cheng, 2020). This allows for the informant to be as comfortable as possible so that their speech most closely emulated natural speech when not being recorded (Peterson, 1996). A heightened awareness of being recorded could result in more careful speech use or possibly higher F0 values.

Setting pitch ranges is a technical requirement in order to analyze pitch. A floor set too low would cause rapid F0 changes to be missed while a ceiling too high would result in missing very low F0 values (Yuasa, 2008). The pitch range for Chalamet was set to 75 to 300 Hz while the pitch range for Depp was set to 100 to 500 Hz, based on Pratt manual recommendations (Boersma and Weenink, 2018) for lower-pitched speakers (males) and higher-pitched speakers (females). Passoni et al. (2018) used these ranges, while Mennen et al. (2012) used similar values, only expanding them slightly for manual correction of problematic cases. This also helps eliminate certain instances of creaky voice, which can go as low as 40 Hz (Boersma and Weenink, 2018). The pitch floor determines the length of the analysis window ($3/\text{floor value in hertz} = \text{analysis window}$). For example, if the pitch floor is set to 75 Hz, the analysis window is 0.04 seconds. The more periods of the sound wave in the analysis window, the higher the precision will be for measurements of pitch based on the spectrogram (Boersma & Weenink, 2018). The time step, which indicates how often Praat computes F0, varies from study to study, with durations of 5 ms (Cheng 2020), 10 ms (Passoni et al., 2018), and 20 ms (Gelfer & Denor, 2013) used. The standard time step in Praat is set to 10 ms; thus, that is the value that was used for this study.

The middle audio of all four interviews was trimmed in Praat so that extraneous noises, including coughs, laughs, long pauses, music, and interruptions, were excluded. Instances of creaky voice were also excluded from the selected data and checked to make sure no creak was present. Although creaky voice is often a feature of fluent speech, it causes problems with pitch tracking (Keating and Kuo, 2012). Once 90 seconds of continuous speech was trimmed from the original full interview so that no extraneous noises or creak was included, phonemes were aligned for each interview either by using a forced aligner or by hand (detailed below).

The English interviews were transcribed into Standard English in a Text file, and The University of Wisconsin at Milwaukee's free Online Forced Aligner was used only after it was made sure no extraneous audio was present in the selected audio clip. The alignment of phonemes and the spectrogram were manually checked in Praat and adjusted so that the audio and text for each phoneme matched. The Forced Aligner then output the average F0 across each phoneme. Only the values for vowels with primary or secondary stress were kept as in Cheng (2020). Stress was determined by using the IPA transcription entry for each word in an English dictionary in unclear cases only (Dictionary.com, n.d.).

For the French interviews, the phonemes were aligned by hand in Praat (following Cheng, 2020; Passoni et al., 2018). As French is a syllable-timed language (English is stress-timed language), the final syllable is generally considered the stressed syllable. The average F0 across the vowel in the final syllable of each word was collected using the command *Get Mean* (Pépiot, 2012).

The pitch tracking was performed using Praat's standard autocorrelation method as the standard algorithm for F0 tracking (following Cox, 2010; Mennen et al., 2012; Possoni et al., 2018). Each F0 contour was inspected for spurious F0 values and manually corrected (Mennen et

al., 2012; Gelfer & Denor, 2013). Figure 3 exhibits an example of spurious F0 values in the stressed vowel in the word ‘thirteen’. In order to correct this, the phoneme alignment bar was manually moved so the spurious values were excluded from the area being measured (in this example, IY1). Alignment only needed to be adjusted so that the vowel phonemes were properly aligned because only the F0 was measured from vowels, not from any consonants. There were only a few instances where the force aligner was not lined up correctly (see Figure 4).

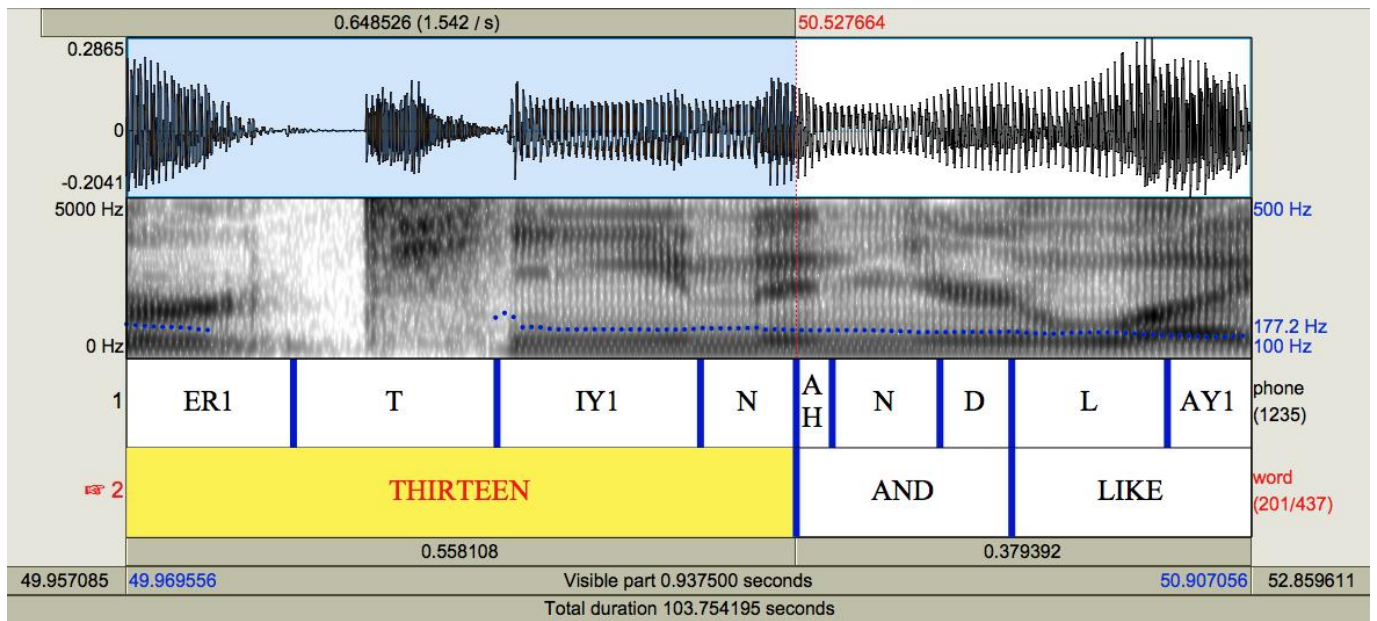


Figure 3. Spurious F0 value manually corrected in Depp’s English speech in ‘thirteen’.

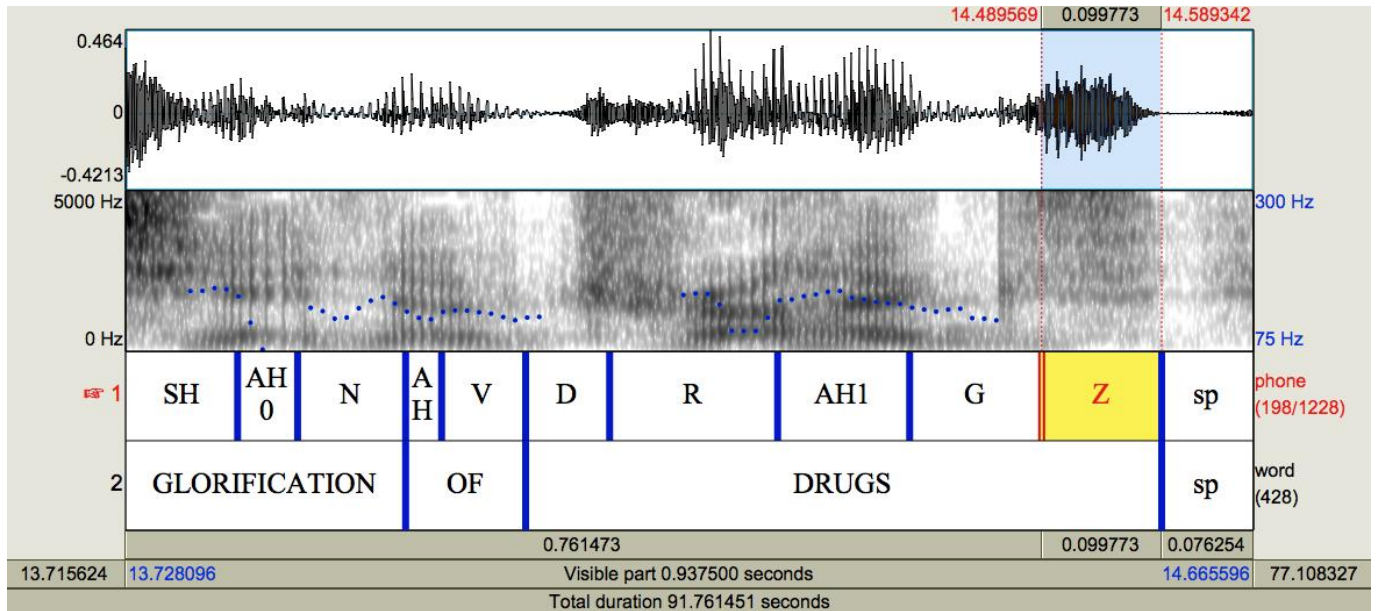


Figure 4. Improper automatic forced alignment in the word ‘drugs’ for Chalameet in English.

The human auditory system responds to sounds logarithmically—semitones— as opposed to linearly—Hertz (Graddol, 1986; ‘t Hart et al., 1990). Because pitch values are being compared, a log scale with semitones is the best way to express perceptual changes in frequencies (Cheng, 2020; Gelfer and Denor, 2013; Graddol, 1986; Traunmüller and Erikson, 1993). In the current study, the values in Hertz were converted to semitones with a base of 100 Hz for further analysis (Cheng, 2020; Gelfer & Denor, 2013; Passoni et al., 2018). The formula used to convert Hertz to semitones was (from Cheng, 2020): $\text{semitone} = [\log(\text{Hertz}) - \log(100)] / \log(2^{1/12})$. Graddol (1986) also found that when expressed in Hertz, women had a larger pitch range than men, but when expressed in semitones, the pitch range used by women was smaller than that used by men. This has to do with how the pitch ranges of women are higher thus the change in Hertz is more drastic. Because the human auditory system perceives sound in semitones, a log scale, these pitch ranges are actually more stable than men's. In addition, Graddol and Swann (1983) and Traunmüller & Eriksson (1995) found that pitch distribution is not a normal curve but is

positively skewed; there are more instances of lower frequencies than higher thus the median is a more accurate representation of average pitch rather than the mean. Because of this, the median pitch was also calculated.

Results

Median f0

Table 1 demonstrates the median pitch values for both informants in both English and French. With both informants, the median pitch values were higher when speaking French as compared to English. Depp had a 3.11 semitone increase from English to French, while Chalamet only had an increase of 1.42 semitones. In Hertz, Chalamet's median F0 for English was 127.16 Hz, while his median F0 value for French was 138.06 Hz. Depp had a median F0 in Hertz for English of 182.92 Hz and for French 218.97 Hz.

Table 1. Median pitch values in semitones by informant and language.

	Timothée Chalamet	Lily-Rose Depp
English median pitch (semitone)	4.159399891	10.45445816
French median pitch (semitone)	5.583474913	13.56843172

Mean f_0

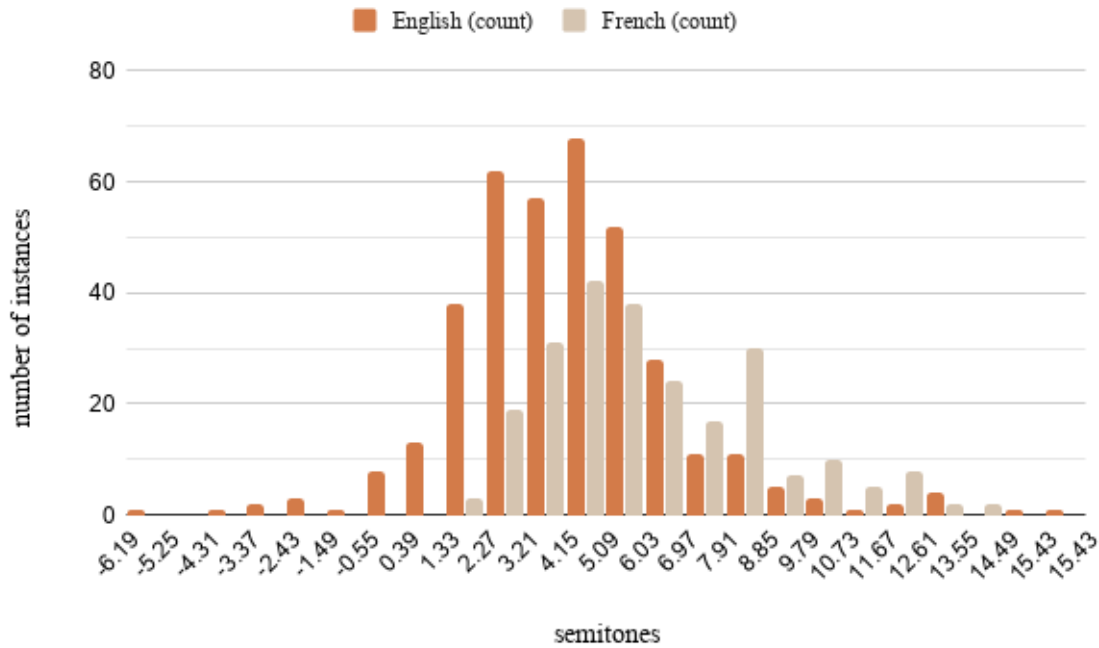


Figure 5. Chalamet F0 across each word in each language.

The mean F0 for when Chalamet spoke English was 4.27 semitones and 6.22 semitones for when he spoke French. The mean is larger than the median (The median for English was 4.16 semitones and for French, 5.58 semitones) in both instances resulting in a skew to the right, as demonstrated in Figure 5.

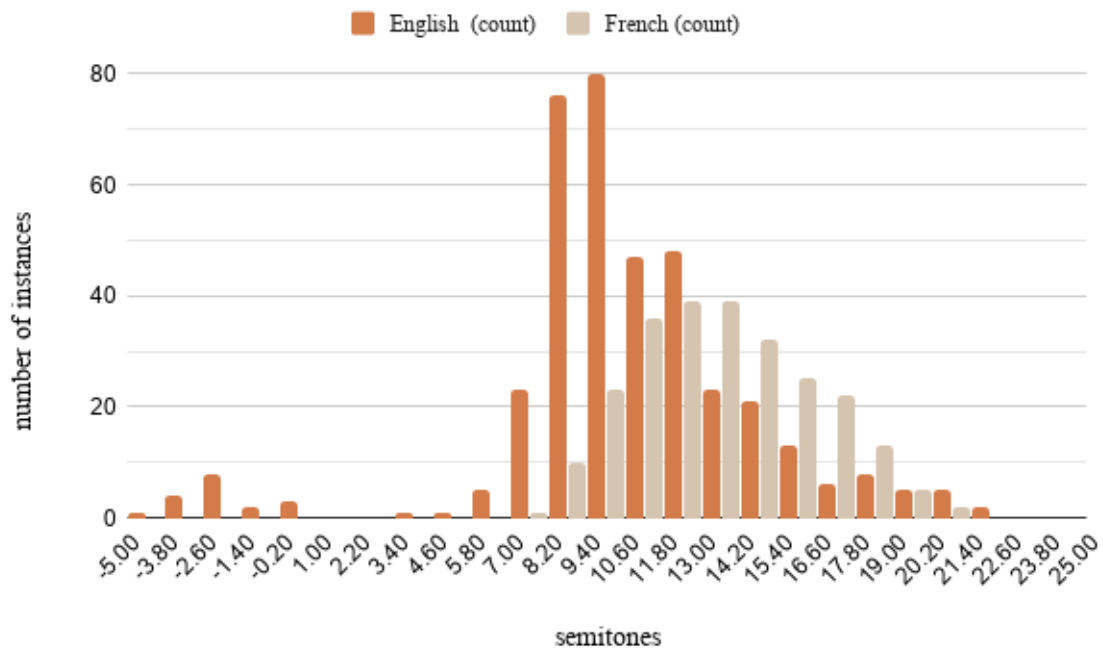


Figure 6. Depp F0 across each word in each language.

The mean F0 when Depp spoke English was 10.78 semitones and 13.66 semitones for French. The mean F0 is larger than the median where the median for English was 10.45 semitones and 13.57 semitones for French, resulting in a right-skewed distribution, as demonstrated in Figure 6.

IBM SPSS Statistics for Mac (2019) was used to run a paired-samples t-test to compare mean F0 values by each informant when speaking English and when speaking (see Tables 2-3 for Chalamet and Table 4-5 for Depp).

Table 2. Paired sample statistics for Chalamet.

	Mean	N	Std. Deviation	Std. Error Mean
Chalamet English	4.12218	238	2.44562	0.15853
Chalamet French	6.22366	238	2.62415	0.17010

Table 3. Paired samples test for Chalamet.

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Chalamet English - Chalamet French	-2.10149	3.65473	0.23690	-2.56819	-1.634795	-8.871	237	0.000

There was a significant difference in mean F0 for English (M=4.12, SD= 2.45) and French (M=6.22, SD=2.62) conditions; $t(237) = -8.87$, $p < .001$ for Chalamet.

Table 4. Paired sample statistics for Depp.

	Mean	N	Std. Deviation	Std. Error Mean
Depp English	11.03674	247	3.65305	0.23244
Depp French	13.66485	247	2.74708	0.17479

Table 5. Paired samples test for Depp.

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 2	Depp English - Depp French	-2.62811	4.53404	0.28849	-3.19634	-2.05988	-9.110	246	0.000

There was also a significant difference in mean F0 for English (M=11.04, SD= 3.65) and French (M=13.66, SD=2.75) conditions; $t(247) = -9.110$, $p < .001$ for Depp.

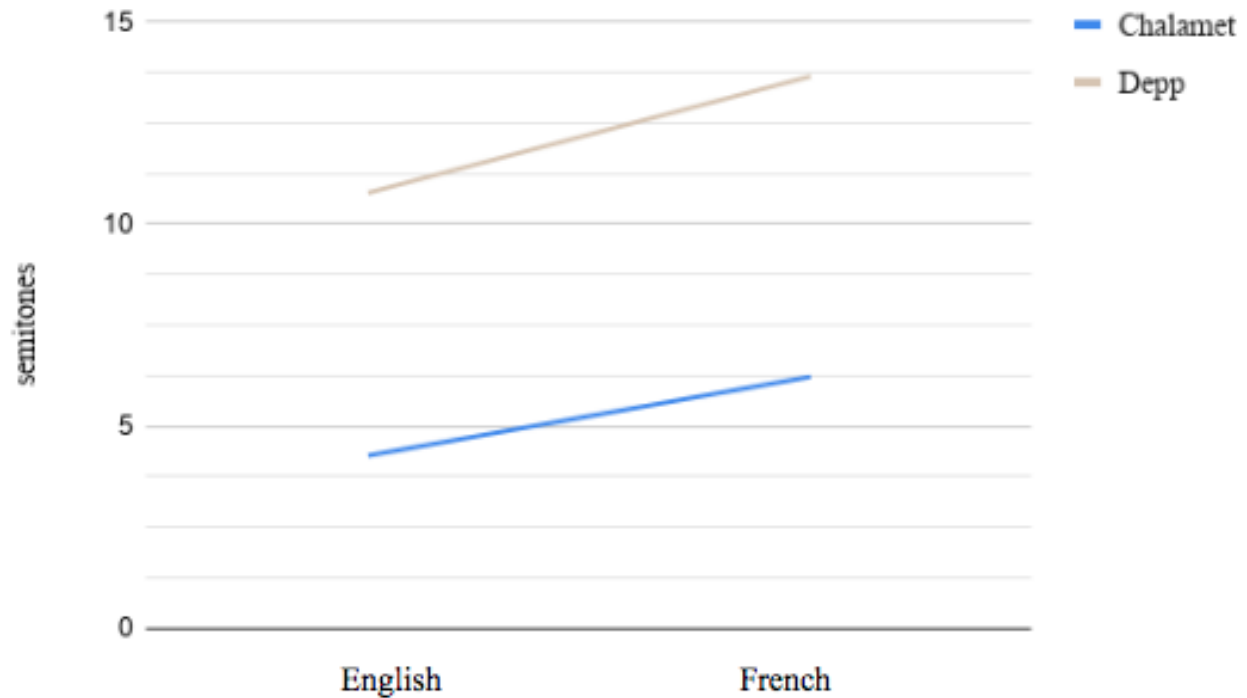


Figure 7. Mean F0 for Chalamet and Depp compared in English and French.

In Figure 7, the slope for Depp from English to French was 2.88 while for Chalamet it was 1.95, therefore Depp had a steeper rise from English to French than Chalamet.

F0 Variance:

F0 values varied more for both informants when speaking English versus when speaking French. Chalamet’s F0 range for English was 21.62 semitone, while for French it was 12.43 semitones. When expressed in Hertz, his F0 range for English was 173.91 Hz and for French 117.29 Hz. Depp’s F0 range for English was 26.91 semitones (287.54 Hz) and for French 12.91 semitones (177.34 Hz). The interquartile range (IQR) for each informant and language is shown in Figure 8. The IQR represents the middle 50% dispersion of a particular dataset so that the between the 75th and 25th percentiles are represented, thus aiding in excluding any outlier values. Figure 8 shows Chalamet and Depp both had more extensive ranges of F0 values for English

when compared to French. When comparing the IQR values, French in both cases had wider IQR values but narrower maximum and minimum values than English. In addition, Depp's lower IQR value did not overlap with Chalamet's, but her minimum F0 values did overlap with Chalamet's. Her English minimum almost extended to Chalamet's English F0 minimum.

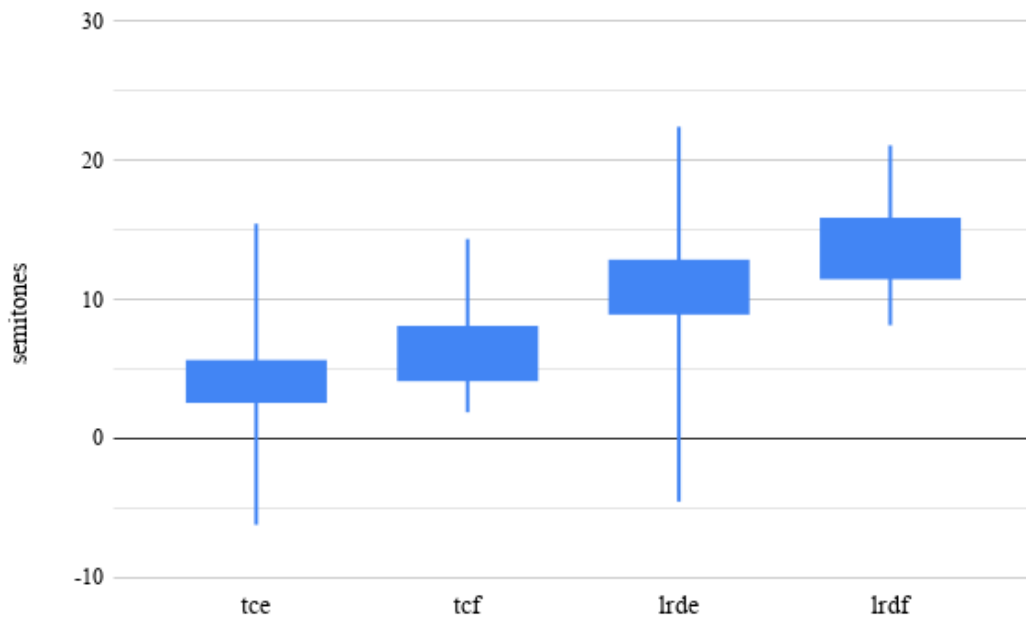


Figure 8. Box plot of F0 range by informant and language.

Discussion & Conclusion

The hypothesis that in the two English and French bilinguals, French is spoken at a higher F0 was demonstrated to be correct. The hypothesis that the female bilingual speakers speak with smaller pitch ranges (Graddol, 1986; Traunmüller & Erikson, 1995) was not supported as Depp had wider F0 ranges when compared with Chalamet for both English and French.

The present study analyzed the F0 mean, median, and span of spontaneous speech produced by two English/French bilinguals. Although statistically significant differences were found for each informant in F0 when speaking English and French, the data cannot explain the

cause of difference. Only speculations can be made. It was not able to be determined if the differences are attributed to varying gender stereotypes or because of a natural speaking pitch that occurs for each language.

First, Chalamet had a mean English pitch of 4.27 semitones or 129.48 Hz. This coincides with previous research that found U.S. male university students to have an average pitch of 123 Hz (Hollien & Jackson, 1972). Chalamet's mean pitch for French was 6.22 semitones or 144.98 Hz. The difference between Chalamet's mean English and French F0 values were found to be statistically significant. This rise in F0 from English to French also echoes previous research that found English speakers to have lower F0 values than Spanish speakers (Cox, 2010; Hanley et al., 1966). Considering the strong linguistic ties between Spanish and French, it would be reasonable to extrapolate this behavior and speculate that French pitch values would follow a similar pattern to English. However, there could be confounding cultural factors that could potentially affect pitch values. For example, gender socialization patterns or cultural norms. However, these are beyond the scope of this paper and will not be explored further.

Depp had a mean pitch of 10.78 semitones or 191.55 Hz for English and 13.66 semitones or 223.00 Hz. This also echoes the previous research that English is spoken at a lower pitch range than Spanish (Cox, 2010; Hanley et al., 1966). It is beyond the scope of this paper to explore cultural stereotypes or gender socialization patterns, although they could be factors that could help explain the results. In addition, it is also important to note that previous studies have found the median to be more representative of an informant's F0 values (Graddol & Swann, 1983; Traunmüller & Eriksson, 1995; Cheng, 2020). Not every study used median though, so this discussion used mean in order to be able to compare data to the previous studies that utilized the mean.

When analyzing median, for both Chalamet and Depp, the data was skewed to the right. This is concurrent with Graddol and Swann's (1983) findings where the median is a more accurate portrayal of a speaker's F0 because pitch is not a perfect normal curve. There is a skew to the right because it has been found that there are more instances of lower F0 values in speech, thus the median takes into account the skewness that naturally occurs in the data.

Depp had a steeper increase in mean F0 from English to French when compared to Chalamet. This demonstrates that Depp has a bigger change in F0 from English to French than Chalamet. A difference in rate between speakers demonstrates that there are different factors influencing pitch differences from English to French or else the rate would have been the same between Chalamet and Depp. This also demonstrates that there might be a wider difference between male and female pitch values for French than in English. This raises the question if this is a result of physiological speaking differences from one language to the other or if there are social factors causing a greater divide in pitch values between men and women in French. Again, this goes beyond the scope of this study because an analysis on social differences between Americans and the French need to be studied further, and more informants are needed in order to determine if this difference is individual based or if it's gender based.

An interesting find that does not echo previous studies is the pitch range between Depp and Chalamet. When expressed in either semitones or Hertz, Depp had a larger pitch range than Chalamet. However, the findings of Graddol (1986) and Traunmüller and Erikson (1995) indicated that women have a smaller pitch range than men, expressed in semitones. A possible source for this discrepancy is the nature of the interview topic. The interviews with Chalamet were slightly more serious than Depp's because, in both his English and French interviews, Chalamet discussed his film *Beautiful Boy*, which deals with drug addiction. Recalling from

Traunmüller and Erikson (1995), higher F0 variation is a characteristic of interest and joy when speaking. A higher degree of liveliness in Depp's speech than Chalamet's could be attributed to her having a larger pitch span than Chalamet ('t Hart et al., 1990).

Also, for each speaker, the IQR boundaries were wider for French while the minimum and maximum values were wider for English. This shows that most English F0 values are more concentrated than French, but in English there is a wider range that less frequent F0 values are spoken at. This wider max/min range that English is spoken at results in the spoken F0 range of English to extend past each speaker's own French F0 range. In addition, Depp had a minimum F0 range that went lower than Chalamet's French range. This is important to note because it demonstrates that there is overlap in men and women's pitch ranges. Although Depp's English IQR range does not extend as low as Chalamet's French IQR range, they both have minimum/maximum F0 measurements that extend into each other's ranges. Depp's minimum for English also extends almost as low as Chalamet's in English as well. This shows that females might take advantage of more F0 lower manipulation techniques than techniques that would raise their F0.

As an aside, creak, a phenomena utilized largely in American speech, was excluded from the data due to complicating pitch tracking. Including it would cause the following difference in the data: both Depp and Chalamet's mean/median F0 values would be lower and the variation would increase as there would be an increase of lower F0 values. Because creaky voice is a feature more utilized in American speech, this could cause the overall mean/median F0 values for English to lower a different amount than French. In addition, Depp utilized creak more than Chalamet, therefore her mean/median F0 values would also cause her F0 values to lower at a different rate from Chalamet.

Future Directions

This study gathered data across two participants, essentially creating a dataset with cardinality two. Future research should be much more comprehensive across multiple participants for each gender. Additional studies that look at monolingual English and French speakers could aid in establishing a basis median F0 value range. Also, a survey similar to van Bezooijen (1995) that looks at English and French speakers could aid in understanding the sociocultural aspects of pitch differences as well. The combination of these studies captures multiple facets of spoken language that could all contribute to the F0 values of a speaker, and clarify how large of a role cultural and gender norms affect the F0 value versus it being due to purely the language itself.

In addition, although there are conflicting results about the frequency of creak in males and females in American English, lower F0 values in English can be attributed to higher uses of creaky voice amongst Americans in English regardless of gender. Previous studies looked at the use amongst monolinguals, not bilinguals (Dallaston & Docherty, 2020; Melvin & Clopper, 2015; Yuasa, 2010). Both males and females employ creaky voice, including Depp and Chalamet. With both informants being fluent in American English, a possibility for future research could be to hone in on creaky voice affecting native American English speaker's engagement with other languages. This would be due to its nature as a sociophonetic marker for fluent American English speakers. Do bilinguals have a higher creaky voice usage when speaking languages besides American English?

Acknowledgments:

I am grateful for Professor Maeve Eberhardt of the Department of Linguistics for helping me the past year and a half on this project. Professor Toni Cook of the Department of Linguistics for always having my back the past four years. Professor Alexander Wurthmann of the Department of Chemistry for serving as my committee chair. Andrew Cheng for meeting with me to further discuss his studies, helping me find a free force aligner online, and making sure I could get it to work. Zach Acosta for letting me use his updated Excel software.

Bibliography:

- Andrianopoulos, M.V., Darrown, K.N., and Chen, J. (2001). Multimodal standardization of voice among four multicultural populations: Fundamental frequency and spectral characteristics. *Journal of Voice*, 15(2), 194-219. [https://doi.org/10.1016/S0892-1997\(01\)00021-2](https://doi.org/10.1016/S0892-1997(01)00021-2)
- Aronovitch, C.D. (1976). The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *The Journal of Social Psychology*, 99(2), 207-220. <https://doi.org/10.1080/00224545.1976.9924774>
- Baken, R.J. & Orlikoff, R.F. (2000). *Clinical measurements of speech and voice* (2nd ed.). Singular Thomson Learning.
- Boersma, P. & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.42, retrieved __ August 2018 from <http://www.praat.org/>
- Cartei, V. & Reby, D. (2013). Effect of formant frequency spacing on perceived gender in pre-pubertal children's voices. *PLoS ONE*, 8(12): e801022. <https://doi.org/10.1371/journal.pone.0081022>
- Chan, M. (2017, October 20). *Comparing gender equality in France and the U.S.* France-Amérique. <https://france-amerique.com/en/comparing-gender-equality-in-france-and-the-u-s/>
- Cheng, A. (2020). Cross-linguistic f0 differences in bilingual speakers of English and Korean. *The Journal of the Acoustical Society of America*, 147(2), EL67-EL73. <https://doi.org/10.1121/10.0000498>
- Cox, A. (2010). Pitch differences in female bilingual Spanish/English speakers. Boise State University ScholarWorks.
- Dallaston, K. & Docherty, G. (2020). The quantitative prevalence of creaky voice (vocal fry) in varieties of English: A systematic review of the literature. *PLoS ONE*, 15(3), 1-18. <https://doi.org/10.1371/journal.pone.0229960>
- Dictionary.com*. (n.d.) Retrieved from <https://www.dictionary.com/>
- Eckert, P. & McConnell-Ginet, S. (2003). *Language and Gender*. Cambridge University Press.
- Evans, S., Neave, N., & Wakelin, D. (2006). Relationships between vocal characteristics and body size and shape in human males: An evolutionary explanation for a deep male voice. *Biological Psychology*, 72(2), 160-163. <https://doi.org/10.1016/j.biopsycho.2005.09.003>
- Fairbanks, G. (1960). *Voice and articulation drillbook* (2nd ed.). Harper & Row.

- Ferrand, C.T. & Bloom, R.L. (1996). Gender differences in children's intonational patterns. *Journal of Voice*, 10(3), 284-291. [https://doi.org/10.1016/S0892-1997\(96\)80009-9](https://doi.org/10.1016/S0892-1997(96)80009-9)
- Gelfer, M.P. & Denor, S.L. (2014). Speaking fundamental frequency and individual variability in Caucasian and African American school-age children. *American Journal of Speech-Language Pathology*, 23(3), 395-406. https://doi.org/10.1044/2014_AJSLP-13-0016
- Graddol, D. (1986). Discourse specific pitch behavior. In C. Johns-Lewis (Ed.), *Intonation in Discourse* (pp. 221-237). Croom Helm.
- Graddol, D. & Swann, J. (1983). Speaking fundamental frequency: some physical and social correlates. *Language and Speech*, 26(4), 351-366. <https://doi.org/10.1177/002383098302600403>
- Graddol, D. & Swann, J. (1998). The voice of authority. In, *Gender Voices* (pp. 12-40). Blackwell Publishers Inc.
- Hanley, T.D., Snidecor, J.C., & Ringel, R.L. (1966). Some acoustic differences among languages. *Phonetica*, 14(2), 97-107. <https://doi.org/10.1159/000258520>
- Hollien, H. & Jackson, B. (1972). Normative data on the speaking fundamental frequency characteristics of young adult males. *Journal of Phonetics*, 1, 117-120. [https://doi.org/10.1016/S0095-4470\(19\)31416-0](https://doi.org/10.1016/S0095-4470(19)31416-0)
- Hollien, H., Moore, P., Wendahl, R.W., & Michel, J.F. (1966). On the nature of vocal fry. *Journal of Speech and Hearing Research*, 9(2), 245-247. <https://doi.org/10.1044/jshr.0902.245>
- Hong, K.H., Ye, M., Kim, Y.M., Kevorkian, K.F., Kreiman, J., & Berke, G.S. (1998). Functional differences between the two bellies of the cricothyroid muscle. *Otolaryngol Head Neck Surg*, 118(5), 714-722. [https://doi.org/10.1016/S0194-5998\(98\)70251-7](https://doi.org/10.1016/S0194-5998(98)70251-7)
- Hudson, A.I. & Holbrook, A. (1982). Fundamental frequency characteristics of young Black adults: Spontaneous speaking and oral reading. *Journal of Speech and Hearing Research*, 25(1), 25-28. <https://doi.org/10.1044/jshr.2501.25>
- IBM SPSS Statistics for Mac, version 26. (2019). (IBM Corp., Armonk, N.Y., USA)
- Jassem, W. (1971). Pitch and compass of the speaking voice. *Journal of International Phonetic Association*, 1(2), 59-68. <https://doi.org/10.1017/S0025100300000256>
- Karlsson, I. (1988). Glottal waveform parameters for different speaker types. *Department for Speech, Music and Hearing Quarterly Progress and Status Report*, 29(2-3), 61-67.

- Keating, P. & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America*, 132(2), 1050-1060. <http://dx.doi.org/10.1121/1.4730893>
- Kreiman, J. & Sidtis, D. (2011). *Foundations of voice studies: an interdisciplinary approach to voice production and perception*. Wiley-Blackwell.
- Lee, B. & Van Lancker Sidtis, D. (2017). The bilingual voice: Vocal characteristics when speaking two languages across speech tasks. *Speech, Language and Hearing*, 20(3), 174-185. doi: 10.1080/2050571X.2016.1273572
- Lieberman, P. (1967). *Intonation, Perception, and Language*. M.I.T. Press.
- Loveday, L. (1981). Pitch, politeness and sexual role: An exploratory investigation into the pitch correlates of English and Japanese politeness formulae. *Language and Speech*, 24(1), 71-89. <https://doi.org/10.1177/002383098102400105>
- Majewski, W., Hollien, H., & Zalewski, J. (1972). Speaking fundamental frequency of Polish males. *Phonetica*, 25(2), 119-125. <https://doi.org/10.1159/000259375>
- Marotta, J. (2017, November 17). 'Call Me by Your Name': Timothée Chalamet is learning how to be a man, onscreen and off. IndieWire. <https://www.indiewire.com/2017/11/timothee-chalamet-call-me-by-your-name-1201898370/>
- Melvin, S. & Clopper, C.G. (2015). Gender variation in creaky voice and fundamental frequency. *The Ohio State University*.
- Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *The Journal of the Acoustical Society of America*. 131(3), 2249-2260. <https://doi.org/10.1121/1.3681950>
- Mysak, E.D. (1959). Pitch and duration characteristics of older males. *Journal of Speech and Hearing Research*, 2(1), 46-54. <https://doi.org/10.1044/jshr.0201.46>
- Nolan, F. (1983). *The Phonetic Bases of Speaker Recognition*. Cambridge University Press.
- Ohara, Y. (1999). Performing gender through voice pitch: A cross-cultural analysis of Japanese and American English. In U. Pasero & F. Braun (Eds.), *Wahrnehmung und Herstellung von Geschlecht: Perceiving and performing gender* (pp. 105-116).
- Olson, D.J. (2016). The impact of code-switching, language context, and language dominance on suprasegmental phonetics: Evidence for the role of predictability. *International Journal of Bilingualism*, 20(4), 453-472. <https://doi.org/10.1177/1367006914566204>

- Passoni, E., Mehrabi, A., Levon, E., & de Leeuw, E. (2018). Bilingualism, pitch range and social factors: preliminary results from sequential Japanese-English bilinguals. 9th International Conference on Speech Prosody. doi: 10.21437/SpeechProsody.2018-78
- Pépiot, E. (n.d.). Voice, speech and gender: male-female acoustic differences and cross-language variation in English and French speakers. XVèmes Rencontres Jeunes Chercheurs de l'ED 268, Jun 2012, Paris, France. (à paraître). ffhalshs-00764811
- Perry, T.L., Ohde, R.N., & Ashmead, D.H. (2001). The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America*, 109(6), 2988-2998. <https://doi.org/10.1121/1.1370525>
- Peterson, J. (1996). Sociolinguistic interviewer style variation: Hyperconvergence in the other informant. *(N)WAVES and MEANS: A selection of papers from NWAVE 24*, 3(1), 159-170.
- [Quotidien avec Yann Barthès]. (2016, September 19). *L'interview Pop et Wizz de Lily Rose Depp* [video]. My TF1. <https://www.tf1.fr/tmc/quotidien-avec-yann-barthes/videos/l-interview-pop-wizz-de-lily-rose-depp.html>
- Ramig, L.A. & Ringel, R.L. (1983). Effects of physiological aging on selected acoustic characteristics of voice. *Journal of Speech and Hearing Research*, 26(1), 22-30. <https://doi.org/10.1044/jshr.2601.22>
- Rosenthal, A.S. (1999). The gender-coded stereotype: An American perception of France and the French. *American Association of Teachers of France*, 72(5), 897-908.
- Sagansky, G. (2016, May 17). *Live from Cannes: Lily Rose Depp on "The Dancer" and Chanel*. W Magazine. <https://www.wmagazine.com/story/lily-rose-depp-cannes-film-festival-2016/>
- Sagisaka, Y., Campbell, N., & Higuchi, N. (1977). *Computing Prosody: Computational Models for Processing Spontaneous Speech*. Springer.
- Sorenson, D.N. (1989). A fundamental frequency investigation of children ages 6-10 years old. *Journal of Communication Disorders*, 22(2), 115-123. [https://doi.org/10.1016/0021-9924\(89\)90028-2](https://doi.org/10.1016/0021-9924(89)90028-2)
- 't Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation: An experimental-phonetic approach to speech melody*. Cambridge University Press.
- [Timothée Chalamet Fan]. (2019, January 19). *(English subtitles) Timothée Chalamet speaking French!! – Quotidien 1/18/2019* [video]. YouTube. https://www.youtube.com/watch?v=XtM_PASnXZA

- Thussu, D.K. (2006). *International communication: Continuity and change* (2nd edition). Hodder Education.
- Traunmüller, H. & Eriksson, A. (1995). The perceptual evaluation of F0 excursions in speech as evidenced in liveliness estimations. *The Journal of the Acoustical Society of America*, 97(3), 1905-1915. <https://doi.org/10.1121/1.412942>
- van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3), 235-236. <https://doi.org/10.1177/002383099503800303>
- Whiteside, S.P. (1998). Identification of a speaker's sex: A study of vowels. *Perceptual and Motor Skills*, 86(2), 579-584. <https://doi.org/10.2466/pms.1998.87.2.595>
- [W Magazine]. (2019, January 17). *Timothée Chalamet talks Hollywood rejection and auditioning for 'Beautiful Boy' | W Magazine* [video]. YouTube. <https://www.youtube.com/watch?v=UvpPE4QkrkM>
- [W Magazine]. (2019, September 3). *Lily-Rose Depp on 'The King' and her French revolution birthday party | Screen tests | W Magazine* [video]. YouTube. <https://www.youtube.com/watch?v=RVM9QG1kB-E>
- Yuasa, I.P. (2008). *Culture and gender of voice pitch: A sociophonetic comparison of the Japanese and Americans*. Equinox Publishing.
- Yuasa, I.P. (2010). Creaky voice: a new feminine voice quality for young urban-oriented upwardly mobile American women?. *American Speech*, 85(3), 315-337. <https://doi.org/10.1215/00031283-2010-018>
- Zhang, Z. (2016). Mechanics of human voice production and control. *The Journal of Acoustical Society of America*, 140(4), 2614-2635. <https://doi.org/10.1121/1.4964509>