

Creaky Voice in American English: How Are American Women Who Use Creaky Voice Perceived? A Literature Review

Pernille Christina Hald Meier 

1. Introduction

“Vocal fry” has joined more traditional young-women voice mannerisms such as run-ons, breathiness and the dreaded question marks in sentences (known by linguists as uptalk) to undermine these women’s authority in newly distinctive ways’ (Wolf 2015). These views that vocal fry is an emerging vocal trait that inhibits women in the social and professional sphere are propagated in present-day media and attributed to the fame of influential figures such as the Kardashians (Rhodan 2014; Wolf 2015). However, studies have found that men also use vocal fry even though they might use it to a lesser extent than women (Yuasa 2010, 325; Dallaston and Docherty 2020, 12). Nonetheless, as can be seen in Dallaston and Docherty’s (2020, 12) literature review, most of the studies focused on ascertaining the presence of vocal fry in women, and only a few also studied the presence of vocal fry in male speakers. This literature review, in order to provide an all-encompassing view of the results, also compared the methods, the speakers and the evaluators in the different studies, as these variables could influence the results that the studies presented (Dallaston and Docherty 2020). It is important to also incorporate these variables in literature reviews pertaining to topics such as speech evaluations in order to deliberate on the different results. Especially considering that studies regarding the perception of speakers of vocal fry had contradictory findings where the use of vocal fry has been perceived as both a positive quality (Yuasa 2010, 329–30) and detrimental to a woman’s career (Anderson et al. 2014, 3). Even though these contradictory results have been presented, no literature review of the articles published since 2010 has been written. Thus, this paper will review papers published between 2010 and 2022 which investigate the perception of Americans who use creaky voice and compare their methodology, the people evaluated, the evaluator and the findings.

In the following section, it will be discussed what vocal fry is, who uses vocal fry and how vocal fry can be measured. This will be followed by the methods with which the papers were

gathered and which questions were of interest for this review. In the fourth section, the speech samples, the evaluation and the results of the studies will be compared, and in the end, a conclusion will be drawn.

2. Previous works

The topic of this paper is situated in the larger field of speech evaluation. Speech evaluation is often studied in connection to evaluations of foreign-accented speech (Kalin and Rayko 1978; Lindemann 2003; Kinzler et al. 2009) or the evaluation of differently accented varieties such as the evaluation of American, Australian and New Zealand English (Bayard et al. 2001). Similarly, the study of the perception of different voice qualities, such as vocal fry, also uses methods such as these to evaluate how the listeners perceive the people who speak with that voice quality (Yuasa 2010; Anderson et al. 2014; Parker and Borrie 2018; Gallena and Pinto 2021; Conner 2022; Pointer, van Mersbergen, and Nanjundeswaran 2022; Taylor, Wheeler-Hegland, and Logan 2022).

The speech quality of interest in this study is vocal fry, also known as ‘creaky voice’ and ‘laryngealization’ (Laver 1980, 125). In terms of the fundamental frequency, ‘the mean fundamental frequency for creak has been found to be 34.6 Hz, in an average range for male speakers’ (Laver 1980, 122). Thus, vocal fry is on the lower side of the fundamental frequency spectrum (Laver 1980, 122). According to Laver (1980, 124), vocal fry sounds like ‘continual, separate taps in rapid sequence’ and can be recognised by this quality by the listeners. However, vocal fry can also be found with the help of acoustic analysis software such as Praat by their unique waveforms (Parker and Borrie 2018, 540).

Even though popular media postulates that vocal fry is an emerging speech quality used by young American women (Rhodan 2014; Wolf 2015), research has found that vocal fry is not exclusively used by young American women; rather, men are also found to use this quality though to a lesser extent (Yuasa 2010, 325; Abdelli-Beruh, Wolk, and Slavin 2014, 187; Dallaston and Docherty 2020, 12). Moreover, it is not only young Americans who use this variety, but vocal fry has also been found in the speech of 35–50 year old American women (Oliveira et al. 2016, 685–86). Additionally, in comparison to American English, research has found evidence for the use of vocal fry in some varieties in the United Kingdom, where men use vocal fry more than women (Dallaston and Docherty 2020, 12). This was established in the last century, indicating that the use of vocal fry is not an emerging quality in the English language (Dallaston and Docherty 2020, 12). Furthermore, American English and British English are not the only varieties in which vocal fry is present, as vocal fry has been identified in speech produced by New Zealand women (Hornibrook, Ormond, and Maclagan 2018, 37–38). Therefore, though young American women seemingly use

vocal fry more than other groups in the United States of America, this does not mean that they are the only ones to use this speech quality compared to other varieties of English or even in the U.S. itself.

Yet, why are these women under critique by popular media? This question cannot be answered with any certainty. However, Chao and Bursten (2021) have presented a possible reason for the critique. They base their explanation on the assumption that listeners interpret women as assuming male traits by consciously or unconsciously using a voice quality that is associated with men (Chao and Bursten 2021, 46). Yuasa (2010, 318) proposes a similar reason for women to use vocal fry as she reasons that women use vocal fry to ‘project a male-like authoritative image’. According to Chao and Bursten (2021, 50-53), what follows an utterance spoken with vocal fry by a woman, is an interpretation of the speaker sounding ironic, which leads to a negative response as the listener interprets the speaker to use a voice quality that they are not entitled to. This hypothesis offers a possible explanation for the negative focus on especially female speakers of vocal fry.

3. Material and methods

In order to find the papers suitable for this literature review, the three search terms ‘creaky voice’, ‘vocal fry’ and ‘glottal fry’ were chosen to search for papers on ScienceDirect and in the different databases available on Ebscohost. The search terms were restricted to these three due to the scope of this paper. The criteria for choosing the papers were that the papers had to discuss the perception of Americans, both male and female, using vocal fry and that they were published from 2010 onwards. This method has the disadvantage that some relevant literature for this review might not have been included; however, based on the theory sections in the chosen papers, the most relevant papers were chosen for this study. This led to choosing the papers depicted in Table 1.

Authors	Year	Titel
Yuasa	2010	Creaky Voice: A New Feminine Voice Quality for Young Urban-oriented, Upwardly mobile American Women?
Anderson et al.	2014	Vocal Fry May Undermine the Success of Young Women in the Labor Market
Parker and Borrie	2018	Judgments of Intelligence and Likability of Young Adult Female Speakers of American English. The Influence of Vocal Fry and the Surrounding Acoustic-Prosodic Context
Ligon et al.	2019	Perceived Desirability of Vocal Fry Among Female Speech Communication Disorders Graduate Students
Gallena and Pinto	2021	How Graduate Students With Vocal Fry Are Perceived by Speech-Language Pathologists
Conner	2022	Creaky, She Spoke: Examining f ₀ , Vocal Creak, and Perceptions of Young Women's Professionalism
Pointer, van Mersbergen and Nanjundeswaran	2022	Listeners' Attitudes Towards Young Women With Glottal Fry
Taylor, Wheeler-Hegland and Logan	2022	Impact of Vocal Fry and Speaker Gender on Listener Perceptions of Speaker Personal Attributes

Table 1. Table of all the papers reviewed in this literary review

Of these papers, the paper written by Connor (2022) is a working paper and the papers written by Pointer, Van Mersbergen, and Nanjundeswaran (2022) and Taylor, Wheeler-Hegland, and Logan (2022) are articles in press.

This paper focuses on the differences in methodology used in the eight papers, which includes from whom the speech samples were collected and how, and the evaluators. It was also of interest how these methodologies might influence the differences in the results. Especially as the papers also discussed some of these possibilities in their own paper but also the papers written by others. In order to analyse the papers for the review, the following questions were asked:

- Who was evaluated?
- How were the speech samples collected and chosen?
- Who was evaluating?
- How was the data collected?

- What were the results of the study?

These questions were compiled and answered in a table to facilitate direct comparison between the papers.

4. Analysis and discussion

4.1 The Evaluators and the Speech Samples

In this chapter, the questions of interest were who was evaluated and how were the speech samples collected and chosen. These two variables are the foundation on which most of the studies rest as they describe the speech stimulus presented to the evaluators. As can be seen in Table 2, two of the studies used speech samples produced by both male and female speakers; thus, a comparison between the perceptions of male and female speakers could be made (Anderson et al. 2014, 2; Taylor, Wheeler-Hegland and Logan, 2022, 3). Only one study did not use any kind of speech stimulus, instead asking the participants to remember the voice quality when using adjectives provided to describe the voice quality (Ligon et al. 2019, 805.e23). The speakers were all college-aged or in their thirties, as indicated in Table 2.

Article	Speakers
Yuasa (2010, 328)	<ul style="list-style-type: none"> • 1 female Californian in their 20s or 30s
Anderson et al. (2014, 2)	<ul style="list-style-type: none"> • 7 women (19-27 years) • 7 men (20-30 years)
Parker and Borrie (2018, 540)	<ul style="list-style-type: none"> • 8 female college students (18-28 years)
Ligon et al. (2019)	<ul style="list-style-type: none"> • No speech stimulus
Gallena and Pinto (2021, 1557)	<ul style="list-style-type: none"> • 32 female graduates students from the Speech-Language-Hearing Sciences Department (21-29 years)
Conner (2022, 23-24)	<ul style="list-style-type: none"> • 5 women from Midland region (USA) (18-25 years)
Pointer, van Mersbergen and Nanjundeswaran (2022, 3)	<ul style="list-style-type: none"> • 3 female graduate students • Neutral dialects
Taylor, Wheeler-Hegland and Logan (2022, 2)	<ul style="list-style-type: none"> • 1 cisgender male • 1 cisgender female

Table 2. Data about the speakers used in the different studies

Another speech sample related aspect that can influence the results is how many voices were used for each variety studied. This is acknowledged by Anderson et al. (2014, 6) when they were discussing the difference between their own paper and Yuasa (2010) as Anderson et al. (2014,6) ‘[used] multiple voice exemplars, thereby avoiding idiosyncratic confounds that may arise when using only one speaker as a stimulus’. Nonetheless, even though this is a possible drawback, half of the studies used only one voice sample for each variety. Yuasa (2010, 328), for example, used two samples from the same speaker from all of their recorded samples for the evaluation of creaky voice. Meanwhile, Parker and Borrie (2018, 540) stated that ‘the speakers were explicitly selected based on their habitual vocal features, with different combinations of voice pitch (low or high), speech rate (slow or fast), and vocal fry (presence or absence)’. Thus, they also provided only one speaker for each variety studied. Taylor, Wheeler-Hegland, and Logan (2022, 2) had a similar approach with the difference that they used one male and one female speaker to voice all these different combinations, resulting in also only providing one speaker for each variety. While also providing only one speaker for each variety, Pointer, van Mersbergen, and Nanjundeswaran (2022, 3) had only three varieties that were studied. Subsequently, all four of these studies might not depict accurate perceptions as the possibility exists that other factors have influenced the evaluators. The three studies that worked with multiple speakers might have eradicated this possibility which would lead to more accurate results (Anderson et al. 2014, 2; Conner 2022, 23–24). Therefore, the choice of how many voices are chosen is an important factor that can influence the accuracy of the results.

The studies chose different speech modes to record and present to the evaluators. The speech stimuli chosen for 5 of the studies were spoken with the natural voice quality of the speaker (Yuasa 2010, 328; Parker and Borrie 2018, 540; Gallena and Pinto 2021, 1557; Conner 2022, 23–24; Pointer, van Mersbergen, and Nanjundeswaran 2022, 3). In most of these cases, this meant that one person was recorded for either vocal fry or for an absence of vocal fry and any other condition that was included in their study. However, Conner (2022, 23) examined the voice samples from the Nationwide Speech Project Corpus and divided the speech samples so that they fit her ‘six experimental conditions; two different levels for f_0 , and three different levels for position of creak’. In order to also evaluate one speaker for both vocal fry and the absence of vocal fry, the other two studies had the speakers record one sample in their natural voice and the other sample with a coached vocal fry (Anderson et al. 2014, 2; Taylor, Wheeler-Hegland, and Logan 2022, 2). However, this could result in the voice sample sounding unnatural as the person who has been coached might not be able to imitate the use of vocal fry completely (Anderson et al. 2014, 6). Furthermore, this could negatively affect the results by obscuring the actual perceptions.

The last factor that could influence the perception of the evaluators is the source material and if it is free speech or recited. Anderson et al. (2014, 2) and Pointer, van Mersbergen, and Nanjundeswaran (2022, 3) both had the speakers recite a text. For example, for Anderson et al. (2014, 2), the phrase ‘Thank you for considering me for this opportunity’ was recited. Additionally, in the studies by Parker and Borrie (2018, 540), Gallena and Pinto (2021, 1558) and Taylor, Wheeler-Hegland, and Logan (2022, 3), passages from the Rainbow Passage were read. In general, it can be argued that ‘the simulated interview script utilized in [the] study is a monologue reading rather than a dialogue that would occur in a face-to-face interview. The use of a monologue-reading task could have altered other vocal aspects, such as intonation and rate of speech, which would naturally occur in an interview’ (Pointer, van Mersbergen, and Nanjundeswaran 2022, 7). Thus, all of the studies reviewed that utilised a script or read a passage could be affected by this. Additionally, Ligon et al. (2019, 32) argue that the phrase ‘Thank you for considering me for this opportunity’ used by Anderson et al. (2014, 2) is not a natural environment for vocal fry. Hence, it would help explain the negative results Anderson et al. (2019) depicted in their study. Especially considering the contradictory result presented by Yuasa (2010). Yuasa (2010, 328) was the only study where spontaneous speech in the form of conversations about food was played for the evaluators. One of the most problematic studies in regard to their speech stimuli is Gallena and Pinto (2021). The reason for this is that they had the 32 speakers record two samples, one of them reading the Rainbow Passage and the other of them describing how to make a peanut butter and jelly sandwich. In the end, only the read sample was presented to the evaluator, as the other recording was used to validate the speakers’ use of vocal fry (Gallena and Pinto 2021, 1558). Additionally, from the samples recorded, they chose the five samples with the least vocal fry as the vocal fry free samples and the seven stimuli with the most recorded vocal fry were used as the speech stimuli for the evaluators (Gallena and Pinto 2021, 1558). However, the vocal fry free samples exhibited in 3-7 syllables vocal fry and the other seven stimuli exhibited in 15-51 syllables vocal fry. Thus, not only is the gap between the two versions not particularly big, the range shown in the vocal fry samples is quite large. Hence, I find that the samples chosen for the study of perceptions of vocal fry can be quite problematic.

4.2 Evaluators

The evaluators who have participated in the study are of importance as the sample size might influence the results. The two studies with the most participants are Anderson et al. (2014, 2) and Parker and Borrie (2018, 540), with 800 and 463 participants, respectively. The three studies with the least amount of evaluators had 60 (Pointer, van Mersbergen, and Nanjundeswaran 2022, 3), 29

(Taylor, Wheeler-Hegland, and Logan 2022, 5) and 23 participants (Ligon et al. 2019, 805.e23). Subsequently, it has to be considered if these sizes are large enough to make statements about the larger population. The theorem that is used to argue sample size is ‘the central limit theorem, which states the means of a random set of samples from a population (the sampling distribution) approach normality as the size of the sample increases’ (Uttley 2019, 153). This theorem applies in most cases if the sample size is over 30; however, the bigger the size of the study, the closer the distribution approaches normality (Uttley 2019, 153). Thus, if a study does not have at least 30 participants, then a statement regarding the larger population should always be considered with caution. In the three aforementioned studies, only the study by Pointer, van Mersbergen, and Nanjundeswaran (2022, 3) has more than 30 evaluators; nonetheless, they only have 30 male and 30 female participants. The other two studies do not comply with the minimum of 30 participants, which would allow them, statistically speaking, to make statements regarding a larger population (Ligon et al. 2019, 805.e23; Taylor, Wheeler-Hegland, and Logan 2022, 5). This is even more poignant for Taylor, Wheeler-Hegland, and Logan’s (2022, 4) study, as the 29 participants were further divided into two groups with 15 and 14 evaluators, respectively. Each of these groups evaluated different aspects (Taylor, Wheeler-Hegland, and Logan 2022, 4). Thus, the sample size can influence the validity of the results presented.

However, not only sample size influences how one views the result, but the composition of the sample is also important. Anderson et al. (2014, 2) not only have the highest number of participants, but their composition is also the most balanced as they used a nearly equal amount for each gender and each of the three age groups. Taylor, Wheeler-Hegland, and Logan (2022, 4), on the other hand, do not, sample size aside, have a balanced group composition on which they could base a comparison as they have done. The reason for this is that in group 1, there were seven men and eight women, and in group 2, four men and ten women (Taylor, Wheeler-Hegland, and Logan 2022, 4). There are many more women in comparison to men in group 2, especially considering the sizes of the groups. Hence, a direct comparison between the perceptions of the two groups might not depict an accurate picture of the perceptions that people have of people using vocal fry.

Another aspect that can influence the results and the way one views the results is the choice of participants. Especially considering ‘that human perception is subjective’ (Gallena and Pinto 2021, 1554). While some studies, such as Anderson et al. (2014, 2), used internet survey portals to find participants and, therefore, had a blend of different people featured in their study, other studies used specific groups of people. One such example is Yuasa (2010, 328,) where the professors recruited students from the University of Iowa and the University of California to participate in

their study. In another study, Speech Communication Disorder graduate students were asked to participate (Ligon et al. 2019, 805.e23). By choosing these students, the study faces the limitation that they are not fully educated and that they are a very specific group (Ligon et al. 2019, 805.e34). Another study that used a very specific group as the evaluators is Gallena and Pinto (2021). They had 150 practising Speech-Language Pathologists evaluate the speech samples; however, the study was also geared towards exploring the perceptions in this particular field and not the perception in the United States of America. Another study that cannot necessarily make statements regarding Americans is Conner (2022, 24), as the participants were people living in the U.S. and spoke English as one of their primary languages. Hence, it does not follow that they are American citizens; a person from Australia who, at that given moment, lives in the U.S would also fit this description, but they might give different answers due to their cultural background. Thus, it is also important to indicate which variety of English the evaluators speak. All of these examples illustrate that the evaluators and their cultural and educational background are important in determining how the results should be viewed.

4.3 Methods

The manner in which the questionnaires are constructed, such as which questions are asked and how they are to be answered, also influences the results. This factor is particularly interesting considering that none of the evaluation methods were the same for any of the studies. While three studies used a Likert scale for their evaluation, they all chose different intervals. Parker and Borrie (2018, 541) used a seven-point Likert scale, and Conner (2022, 24) had a six-point Likert scale. The study with the smallest interval was Taylor, Wheeler-Hegland, and Logan (2022, 4) with a five-point Likert scale which might have been too small to gather conclusive evidence. This is something that they themselves acknowledged as a limitation of their study (Ligon et al. 2019, 8). Two other studies also used scales; however, they used scales running from 0-100, which offer a much larger range for the participants to rate the attributes of interest (Gallena and Pinto 2021, 1557; Pointer, van Mersbergen, and Nanjundeswaran 2022, 4). Meanwhile, Yuasa (2010, 328) used two question types. They used a seven-point scale from negative 3 to positive 3, with which the participants rated the voices for the different attributes and an open question regarding their impression of the speaker with vocal fry (Yuasa 2010, 328-9). Contrarily, Anderson et al. (2014, 2) did not ask the evaluators to rate voice samples on a scale; they rather asked them to evaluate two voice samples and decide for each person if they preferred their sample with vocal fry or without. Thus, they presented a direct comparison of one person with vocal fry and one without. However, the participants evaluated either the male speakers or the female speakers, not both. Therefore, the

question arises if a comparison between the results of the two groups can be compared as the pools are not the same (Anderson et al. 2014, 2). As demonstrated, all of the studies had different rating methods, which could contribute to different results.

Another variable that was different for the studies was the amount of time one could listen to the recording. Of the seven studies that had voice recordings, three of the studies did not specify how many times one could listen to the voice sample (Gallena and Pinto 2021; Pointer, van Mersbergen, and Nanjundeswaran 2022; Conner 2022). Meanwhile, only one study presented the samples only once for the participants (Parker and Borrie 2018, 541), and another study had the voice sample played four times (Taylor, Wheeler-Hegland, and Logan 2022, 4). The last two studies allowed the participants to listen to the recordings as many times as they wanted (Yuasa 2010, 328; Anderson et al. 2014, 2). The times a participant could listen to a voice sample is another variable that differentiates the studies.

Even though all the studies investigated the perceptions people had of Americans using vocal fry in their speech, they did not all investigate the same parameters. The parameter that was most often rated was 'intelligence'/'educated', which was rated in six of the studies among other parameters (Yuasa 2010, 328; Anderson et al. 2014, 3; Parker and Borrie 2018, 541; Gallena and Pinto 2021, 1565; Pointer, van Mersbergen, and Nanjundeswaran 2022, 3; Taylor, Wheeler-Hegland, and Logan 2022, 4). Another parameter that was used in more than two studies is 'competence', which was rated in three studies (Anderson et al. 2014, 3; Gallena and Pinto 2021, 1565; Conner 2022, 24). The last such parameter was 'hireable', which in the case of Gallena and Pinto (2021, 1565) was both a yes/no question and a rateable parameter, Anderson et al. (2014, 3) had the participants compare the two voice samples, and Pointer, van Mersbergen, and Nanjundeswaran (2022, 3) constructed the study around 14 hiring contrast pairs; thus, this parameter was in the centre of all 14 contrast pairs. Nonetheless, these were not the only parameters that were of interest in the different studies. Yuasa (2010, 328), for example, had five other contrast pairs that the evaluators needed to rate. Meanwhile, Taylor, Wheeler-Hegland, and Logan (2022, 4) had group 1 rate the criterion 'likeability' and 'salary' and group 2 'attractiveness' and 'intelligence'. As illustrated, there are some parameters that the different studies share, but also many attributes that are different, which can lead to different results, as either the question is different, or the attributes might be interpreted differently.

A unique study in terms of their methodology was Ligon et al. (2019). In this study, the Speech Communication Disorder students were asked to rate 25 adjectives in terms of their desirability on a scale of 1 (desirable) to 3 (undesirable) (Ligon et al. 2019, 805.e23). In the next question, they were asked to rate eight different voice qualities, including vocal fry, with these same

adjectives (Ligon et al. 2019, 805.e23). These adjectives with their scores attached were then used to determine how desirable each of the voice qualities were to the individual evaluator (Ligon et al. 2019, 805.e23). Ligon et al. (2019, 805.e24) stated that the reason for employing this method was to minimise prejudices that the participants might have. One pitfall of this method was that the rating of the adjectives was done directly before utilising the adjectives in describing the voice qualities, which might lead to awareness of how they themselves rate these adjectives. However, this is the only study that employed an indirect rating system. The other possible drawback is that the students evaluated the voice qualities from memory and did not have a speech sample (Ligon et al. 2019, 805.e23). On the one hand, with no speech sample, they might have rated the voice qualities based on their knowledge and biases; on the other hand, they are also not influenced by any input from the speech sample not relevant to the study.

4.4 Results

As varied as the methods have been as varied the results have also been. When the results of the studies are compared, then in most cases, people, especially women, speaking with vocal fry are seen in a more negative light to the people speaking without vocal fry (Yuasa 2010, 330–31; Anderson et al. 2014, 3; Parker and Borrie 2018, 541–42; Gallena and Pinto 2021, 1561; Conner 2022, 27; Pointer, van Mersbergen, and Nanjundeswaran 2022, 4; Taylor, Wheeler-Hegland, and Logan 2022, 6). However, there were some studies that had the opposite result (Yuasa 2010, 330–31; Parker and Borrie 2018, 541–43). In regard to the aspect ‘intelligent’ or ‘educated’, as the aspect most of the studies had in common, Yuasa (2010, 330–331) illustrate that people speaking with vocal fry are perceived as more educated than speakers without vocal fry. These results match the rest of the parameters studied in this article (Yuasa 2010, 330–31). The other article that demonstrated positive perceptions regarding the use of vocal fry did so only partially (Parker and Borrie 2018, 541–42). ‘The results revealed that ratings of intelligence were highest for the speaker who presented with a low pitch, a fast rate, and no vocal fry’; however, the speaker who spoke with a low pitch, slow rate and vocal fry was perceived as more intelligent than the person who spoke with the same combination but without vocal fry (Parker and Borrie 2018, 541–42). Thus, it was determined that vocal fry is not a speech quality that can, on its own, dictate how people tend to perceive speakers of vocal fry (Parker and Borrie 2018, 541). Similar results have been found for the other parameter, ‘likeability’, in the study (Parker and Borrie 2018, 542–43). Meanwhile, the other studies showed a more negative perception of people using vocal fry not only for the attribute ‘intelligent’ or ‘educated’ but across the whole board (Anderson et al. 2014, 3; Gallena and Pinto 2021, 1560–1561; Pointer, van Mersbergen, and Nanjundeswaran 2022, 4–5; Taylor, Wheeler-

Hegland, and Logan 2022, 5-6). However, Taylor, Wheeler-Hegland, and Logan (2022, 6) did not come to a conclusive result in the case of the attribute 'intelligent' in male speakers, unlike the female speakers, as there was nearly no difference, rather the male speakers with vocal fry were perceived as slightly more intelligent. Additionally, Pointer, van Mersbergen, and Nanjundeswaran (2022, 4) not only demonstrate that voices without vocal fry were perceived as more intelligent but that there is also a difference in perception between the use of vocal fry where vocal fry at the end of sentences was perceived in a more positive light than continuous vocal fry. Additionally, Conner (2022, 27) determined that there was no effect on the perception of professionalism due to vocal fry only; an effect was only found in connection with other factors. All of these results create a field of perception of the use of vocal fry that needs more research.

Some of the implications of the different results garnered are the different methodologies used in the papers. One example that illustrates this is the different results for the parameter 'trustworthy' in three papers (Yuasa 2010, 330–31; Anderson et al. 2014, 3; Pointer, van Mersbergen, and Nanjundeswaran 2022, 5). Of the three studies, only Yuasa (2022, 330-331) depicted a more positive view of people using vocal fry. There are two explanations for this based on the methodology. Parker and Borrie (2018, 543) state that the difference might have occurred because the speakers had 'a low pitch and a fast rate, rendering negative listener impressions when vocal fry was present' in the study of Anderson et al. (2014). Additionally, the speech sample was not made by a person who speaks naturally with vocal fry. 'Consequently, the listeners could have been responding to both vocal fry and to differences between the pairs of voices that are not associated with vocal fry' (Anderson et al. 2014, 6). Thus, the listener might have felt that due to the difference, the speaker was tricking them; hence, found the speaker untrustworthy. However, this is only a hypothesis. Moreover, both Anderson et al. (2014, 2) and Pointer, van Mersbergen, and Nanjundeswaran (2022, 3) recited something. Incidentally, Yuasa (2010, 328) is the only study that used spontaneous speech as their speech sample. This example illustrates how the methodology can affect not only the results but also how these results are viewed and questioned.

The results of Ligon et al. (2019) need to be considered separately as they had a different methodological set-up than the other papers. Although their focus was on the perception of vocal fry, they also included other voice qualities that the participants needed to rate (Ligon et al. 2019, 805.e23). When the relative desirability of the eight voice qualities was compared, vocal fry was not the most undesirable voice quality; rather, 'breathy', 'soft or weak', 'rough' and 'strained' were all rated as more undesirable according to the adjectives chosen to describe the particular voice quality (Ligon et al. 2019, 805.e25). Moreover, 'a female vocal quality profile for each adjective was created based on the frequency the adjective was used to describe each vocal quality' (Ligon et al. 2019,

805.e25). Of the 25 adjectives used, 18 adjectives had a vocal quality profile that contained vocal fry (Ligon et al. 2019, 805.e26). The three adjectives with the highest amount of vocal fry were 'vain', 'apathetic' and 'sleepy'; however, the adjectives containing vocal fry were not only negative but also positive, for example, 'sophisticated', 'cool' and 'mature' (Ligon et al. 2019, 805.e26). Even though these results are interesting, the sample size for this study was not over 30, which needs to be taken into consideration when using the data of this study for further research (Ligon et al. 2019, 805.e23).

5. Conclusion

In conclusion, over the last few sections, the speakers and their speech samples, the evaluators, the methodology and the results were discussed. These were not only compared to each other but also evaluated for their implications on their papers. Some of the conclusions drawn from those chapters were that some of the papers had too small sample sizes (Ligon et al. 2019, 805.e23; Taylor, Wheeler-Hegland, and Logan 2022, 5) and that one needs to consider many different factors such as the participants or the speakers when conducting these studies.

Other research that could be very interesting based on the discussion in this paper is further research into the relationship of pitch, speech rate and the presence of vocal fry and how these qualities influence the perception of the speakers. Another interesting paper would be to expand on Ligon et al.'s (2019) paper to either ask a larger amount of the same group or to expand the group to include people from different social backgrounds. Additionally, an expansion of the parameters considered in most of these papers would also open the discussion of the perception of vocal fry instead of focusing on the professional sphere.

Reference List

- Abdelli-Beruh, Nassima B., Lesley Wolk, and Dianne Slavin. 2014. 'Prevalence of Vocal Fry in Young Adult Male American English Speakers'. *Journal of Voice* 28 (2): 185–90. doi:10.1016/j.jvoice.2013.08.011.
- Anderson, Rindy C., Casey A. Klofstad, William J. Mayew, and Mohan Venkatachalam. 2014. 'Vocal Fry May Undermine the Success of Young Women in the Labor Market'. *PLOS ONE* 9 (5): e97506. doi:10.1371/journal.pone.0097506.
- Bayard, Donn, Ann Weatherall, Cynthia Gallois, and Jeffery Pittam. 2001. 'Pax Americana? Accent Attitudinal Evaluations in New Zealand, Australia and America'. *Journal of Sociolinguistics* 5 (1): 22–49. doi:10.1111/1467-9481.00136.
- Chao, Monika, and Julia R. S. Bursten. 2021. 'Girl Talk: Understanding Negative Reactions to Female Vocal Fry'. *Hypatia* 36 (1): 42–59. doi:10.1017/hyp.2020.55.
- Conner, Katie. 2022. 'Creaky, She Spoke: Examining F0, Vocal Creak, and Perceptions of Young Women's Professionalism'. *University of Pennsylvania Working Paper in Linguistics*. The Ohio State University, September 19. <https://repository.upenn.edu/pwpl/vol28/iss2/4>.
- Dallaston, Katherine, and Gerard Docherty. 2020. 'The Quantitative Prevalence of Creaky Voice (Vocal Fry) in Varieties of English: A Systematic Review of the Literature'. *PLOS ONE* 15 (3): e0229960. doi:10.1371/journal.pone.0229960.
- Gallena, Sally K., and James A. Pinto. 2021. 'How Graduate Students With Vocal Fry Are Perceived by Speech-Language Pathologists'. *Perspectives of the ASHA Special Interest Groups* 6 (6): 1554–65. doi:10.1044/2021_PERSP-21-00083.
- Hornibrook, Jeremy, Tika Ormond, and Margaret Maclagan. 2018. 'Creaky Voice or Extreme Vocal Fry in Young Women'. *New Zealand Medical Journal* 131 (1486): 36–40.
- Kalin, Rudolf, and Donald S. Rayko. 1978. 'Discrimination in Evaluative Judgments against Foreign-Accented Job Candidates'. *Psychological Reports* 43 (3): 1203–9. doi:10.2466/pr0.1978.43.3f.1203.
- Kinzler, Katherine D., Kristin Shutts, Jasmine DeJesus, and Elizabeth S. Spelke. 2009. 'Accent Trumps Race in Guiding Children's Social Preferences'. *Social Cognition* 27 (4): 623–34. doi:10.1521/soco.2009.27.4.623.
- Laver, John. 1980. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.
- Ligon, Claire, Carrie Rountrey, Noopur Vaidya Rank, Michael Hull, and Aliaa Khidr. 2019. 'Perceived Desirability of Vocal Fry Among Female Speech Communication Disorders Graduate Students'. *Journal of Voice* 33 (5): 805.e21-805.e35. doi:10.1016/j.jvoice.2018.03.010.

- Lindemann, Stephanie. 2003. 'Koreans, Chinese or Indians? Attitudes and Ideologies about Non-Native English Speakers in the United States'. *Journal of Sociolinguistics* 7 (3): 348–64. doi:10.1111/1467-9481.00228.
- Oliveira, Gisele, Ashira Davidson, Rachele Holczer, Sara Kaplan, and Adina Paretzky. 2016. 'A Comparison of the Use of Glottal Fry in the Spontaneous Speech of Young and Middle-Aged American Women'. *Journal of Voice* 30 (6): 684–87. doi:10.1016/j.jvoice.2015.08.015.
- Parker, Michelle A., and Stephanie A. Borrie. 2018. 'Judgments of Intelligence and Likability of Young Adult Female Speakers of American English: The Influence of Vocal Fry and the Surrounding Acoustic-Prosodic Context'. *Journal of Voice* 32 (5): 538–45. doi:10.1016/j.jvoice.2017.08.002.
- Pointer, Natalie Foulks, Miriam van Mersbergen, and Chaya D. Nanjundeswaran. 2022. 'Listeners' Attitudes Towards Young Women With Glottal Fry'. *Journal of Voice*, October 10. doi:10.1016/j.jvoice.2022.09.007.
- Rhodan, Maya. 2014. '3 Speech Habits That Are Worse Than Vocal Fry in Job Interviews'. *Time*, April 6, 2014. <https://time.com/2820087/3-speech-habits-that-are-worse-than-vocal-fry-in-job-interviews/>.
- Taylor, Bryn, Karen Wheeler-Hegland, and Kenneth J. Logan. 2022. 'Impact of Vocal Fry and Speaker Gender on Listener Perceptions of Speaker Personal Attributes'. *Journal of Voice*, November 16. doi:10.1016/j.jvoice.2022.09.018.
- Uttley, J. 2019. 'Power Analysis, Sample Size, and Assessment of Statistical Assumptions—Improving the Evidential Value of Lighting Research'. *LEUKOS* 15 (2–3): 143–62. doi:10.1080/15502724.2018.1533851.
- Wolf, Naomi. 2015. 'Young Women, Give up the Vocal Fry and Reclaim Your Strong Female Voice'. *The Guardian*, July 24, 2015, sec. Opinion. <https://www.theguardian.com/commentisfree/2015/jul/24/vocal-fry-strong-female-voice>.
- Yuasa, Ikuko Patricia. 2010. 'Creaky Voice: A New Feminine Voice Quality for Young Urban-Oriented Upwardly Mobile American Women?' *American Speech* 85 (3): 315–37. doi:10.1215/00031283-2010-018.