



Perceptions of Vocal Fry by English L2 Undergraduate Students in Korea

Seoyoung Kim

Hankuk University of Foreign Studies, South Korea

Sueah Lee

Hankuk University of Foreign Studies, South Korea

Max Watson*

Assistant Professor, Hankuk University of Foreign Studies, South Korea

Abstract

Vocal fry, also known as glottalization, is prevalent in the U.S. Many studies have been conducted regarding listener perceptions, but there is a lack of research targeting the perceptions of Asians, specifically Koreans, who speak English as a second or other language. Setting the goal of finding out the perceptions of Korean university students toward vocal fry as the purpose of this study, student listeners were queried on their perceptions of vocal fry acceptability between speakers demonstrating vocal fry and speaking conventionally. The basis of this study is a survey presenting paired audio samples in a random order of young U.S. native English speakers speaking twice: once conventionally and once with vocal fry. The questions seek to uncover listener perceptions of the acceptableness and professionalism of vocal fry. The result of the survey, conforming to our hypothesis, is that vocal fry is perceived negatively, especially for young female speakers. Also, listeners report a higher-than-average negative correlation for vocal fry when the speaker is imagined as a professor. These results suggest that college-aged Koreans, like the general U.S. populace, express negative perceptions of speakers demonstrating vocal fry.

Keywords: vocal fry, glottalization, English linguistics, Korea

INTRODUCTION

This study examines the attitudes and perceptions towards vocal fry, specifically students attending courses in the English Linguistics & Language Technology major at Hankuk University of Foreign Studies (HUFSS) in Seoul, South Korea whose mother tongues are not English. Vocal fry is a manner of speech that many people have been using for decades, but that has only been significantly recognized by the masses after being adopted by current media personalities that demonstrate vocal fry in their voices, such as Kim Kardashian, Keshha, and Katy Perry. Some researchers have organized studies of

how vocal fry affects women in the workplace and academic settings, such as that conducted by Anderson et al. (2014). Some cases have even shown people's negative responses towards vocal fry, going as far as to say a person with vocal fry sounds manipulative and untrustworthy (Borrie & Delfino, 2017). However, there is a lack of research querying users of English as a second or other language on their perceptions of speakers demonstrating vocal fry. Therefore, this study seeks to discover Hufs student perceptions of vocal fry in imagined settings, with the hypothesis that they perceive speaker vocal fry negatively.

LITERATURE REVIEW

Vocal fry, also known as glottalization, pulse phonation, or "creaky voice," is a form of phonation characterized by distinct laryngeal vibratory patterns, sounds in the lower register of the vocal cords, as well as distinct acoustic features, and is a distinct vocal quality (Wolk et al., 2012). According to Fessenden (2011), creaky voice consists of intermittent irregular vibrations of the vocal cords in the larynx. The vibration happening in the larynx, the glottalization, is the lowest staccato vibration that can be produced during the vocalization of speech. Since Hollien et al.'s 1966 study, vocal fry has been recognized as the lowest of the three vocal registers, which also include falsetto and modal—the usual speaking register. Additionally, they claimed that the recognition and systematic investigation of a shift from modal register to vocal fry might be triggered by linguistic factors as well as the field of speech language pathology, instead of being limited to it.

Williamson (2015) reports that during the production of voiceless sounds the glottis is open: the vocal folds are held apart as the air passes through then the folds do not vibrate (abducted) and are relaxed. This is also the state of the glottis for restful breathing. During the production of voiced sounds, however, the glottis is opened and closed in rapid succession as the vocal folds vibrate. Two further states of the glottis can be identified. One is when the vocal folds are held relatively close together under tension but without vibration. This is the configuration adopted for whispering. The final state is when the vocal folds are held firmly together, closing off the glottis all together. This configuration is adopted when lifting weights; where it serves to prevent air escaping from filled lungs, thereby holding the chest in an expanded, rigid position against which the arms can gain better leverage. This state can also be momentarily adopted during speech to produce what is known as a glottal stop. It is produced by the sudden shutting and opening of the glottis and is transcribed in IPA as /ʔ/.

According to Borrie and Delfino (2017), it is said that sporadic glottalization should be able to be differentiated from ongoing glottalization, which has been found that it is often associated with vocal disorders. However, in response to claims that vocal fry can damage vocal cords, Johns Hopkins Medicine (n.d.) otolaryngologist Dr. Akst, M.D., stated that vocal fry does not cause physiological harm, although there is a concern of it becoming a habit. However, if it is used for long periods of time, it may end up causing negative effects by making the throat feel sore and tired, owing to the small puffs of air that are sent through the vocal folds, causing them to slap together instead of rubbing smoothly. If

vocal folds slap against each other for extended periods of time when they are supposed to smoothly rub against one another, fatigue can grow.

Fessenden (2011) reports that people previously believed in the false myth that vocal fry only appears in the speeches of people with a voice disorder that was believed to lead to vocal cord damage because it has frequently been associated with abnormal vocal laryngeal outputs. However, as seen rather widely in the U.S. media, the use of creak also occurs among speakers with conventional voice quality. However, vocal fry, with its low tones and exaggerated features of voice vibrations, has been attributed to less competent, undereducated, and untrustworthy individuals by many professionals in business settings (Anderson et al., 2014).

Many in the public have the erroneous belief that only women tend to fry their voices, but males engage in the practice as well. However, for females, understanding the negative impact of vocal fry is especially important. The results of Anderson et al.'s (2014) study suggest that women are penalized for vocal fry more than men. While there's no clear reason for this, it may be because vocal fry causes the speaker's voice to drop to an abnormally low pitch. Since men's voices are typically lower than women's voices, the drop may be more noticeable in women, making people think it's sometimes unattractive or uncomfortable to hear women talking with fry. However, it has been hypothesized that if someone has just the right amount of fry in their voices, it is said that the person speaking with fry may give off a relaxed impression (Johns Hopkins Medicine, n.d.).

METHODOLOGY

Vocal Prompts

The prompts used in this study were procured from Anderson et al.'s (2014) survey. They were selected because of their previously demonstrated validity, as well as accessibility. These prompts were produced as per Fraccaro et al.'s (2013) direction in which speakers imitate the vocal fry affection. Specifically, "seven young adult females (average age 24 years, range 19–27 years) and seven young adult males (average age 26 years, range 20–30 years) were recorded speaking the phrase, 'Thank you for considering me for this opportunity' in both their normal tone of voice and in vocal fry" (Anderson et al., 2014).

Participants and Procedures

The experiment was offered online and in the classroom to 86 undergraduate students enrolled in Prof. Watson's Fall 2018 courses in the English Linguistics department at Hankuk University of Foreign Studies in Seoul, South Korea. Of the invitees, 79 students attempted the survey, with 69 completing it. The survey was administered through LimeSurvey, an open-source online statistical survey web app that was installed by the researchers on the Google Cloud Platform. The listeners received the survey through the Canvas learning management system used in their course, with the task presented as a URL through the "quizzes" feature. Participants were informed of the task directions and the expected five-minute duration and offered 3 extra credit attendance points in the course in which they received the quiz. Participation was not required.

Following the research model of Anderson et al. (2014), the experiment was a between-subjects design. Participants were assigned randomly to listen and judge either the male or female voice sample group. Because Anderson et al. (2014) found in pretesting that listeners would be fatigued by having to analyze female and male voices in the same sitting, each group was randomly tasked to analyze either male or female voices. Of the 69 completed surveys, 41 listened to female speakers, but only 28 were randomly assigned to listen to male speakers (32% of total listeners). The counterbalance of the design may be attributed to an overexpression of chance due to limited sample size.

All audio files were presented on a single web page, divided into matched pairs of a single voice actor performing conventional vocalization or fry. The HTML5 audio player presented separate recordings with their own controls for each of the audio pairs, one labeled "Recording A" and the other "Recording B". This player was chosen for simplicity and compatibility across most current mobile and computer browsers. The order of each audio pair was randomized for each participant. As well, the presentation of the normal or fry sample was randomized for each participant between the A and B recordings.

After playing a pair of recordings, listeners were asked 1. which speaker had a more desirable accent, 2. who they would want to be friends with, 3. who made more money, and 4. which they would want as a professor. These questions were presented in the same order in a list below each pair of recordings. It was hypothesized that randomizing these questions would place an unnecessary burden on the participants, being English language learners, and could negatively affect response accuracy. Our set of questions differed from Anderson et al. (2014) as their survey sought to determine the employability of those voices demonstrating vocal fry. As our participant sample would likely not have any experience in either acting as an employer or engaging in hiring, so our questions focus on desirability and professionalism. These were chosen to determine overall perceived acceptability of vocal fry. Listeners' responses were coded as 1 if they selected the vocal fry sample from the pair, and 0 if the normal voice was selected. To reduce response bias, a third option of "No Answer" was introduced to the survey, allowing for responses wherein the listener could not make a differentiation between a recorded pair. Responses involving this third option were not tabulated into the question's results.

RESULTS

Analysis looks to the listener as the unit of measure. Analyses were conducted in R (version 3.5.1). Two-tailed one-sample t-tests were used to compare listeners' average preference ratios to 0.50, a figure representing indifference between a normal voice and vocal fry. Shown in Figure 1, preference ratios for all four assessments by speaker gender are lower than 0.43, indicating that listeners preferred a normal voice over vocal fry at a rate greater than chance ($p < 0.05$).

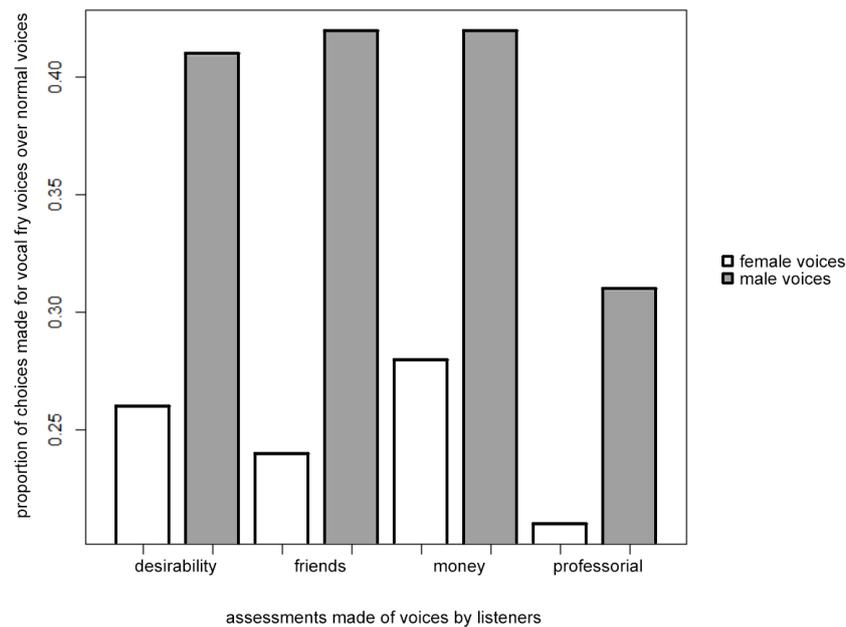


Figure 1. Average preference ratios for vocal fry over normal voice by speaker gender.

Preference ratios show the proportion of responses supporting vocal fry compared to a normal speaking voice. A preference ratio of 1 (0) reflects a strict preference for vocal fry, while 0.50 indicates no discernable preference for either vocal fry or normal voice.

DISCUSSION

This paper examines perceptions of vocal fry spoken by native U.S. English speakers in an acceptability context. A between-subjects experiment conducted with a sample of college-age Korean university attendees shows that vocal fry is perceived negatively compared to normal speech. Young adult voices demonstrating vocal fry are interpreted as less desirable and less professional, and subsequently less acceptable. The negative perception of vocal fry was significantly greater when the speaker was female. Yuasa (2010) reported that college-age students residing in northern California and Iowa in the U.S. regarded female vocal fry as educated, urban-oriented, and upwardly mobile. This negatively contrasts to our survey population which strongly rejected female vocal fry in all queried categories, especially in terms of being an appropriate vocalization for a female professor, which could be imagined as an educated, urban-oriented, upwardly mobile profession. Anderson et al. (2014) posited that their study may have found a negative correlation with vocal fry due to their adult-aged sample. However, our research supports their findings with a college-aged sample.

The especially negative correlation found between vocal fry and professor speech is worthy of further investigation. It could be hypothesized that our English learner listeners, regardless of their other perceptions of vocal fry, value clarity of speech which can be distorted with the fry affectation. This distortion might otherwise be considered acceptable in other situations imagined by our survey questions.

Our study also offers the suggestion that perception of vocal fry, particularly of young female speakers, could vary on a global scale. Variables may be introduced through background interference of non-English languages spoken by the listener. Wolk et al. (2012), and Belotel-Grenié and Grenié (2004) identified that some languages, such as Chinese, widely utilize vocal fry as part of their normal pronunciation. This could influence listener preferences in English. More research on non-native English speaker perceptions of vocal fry-affected English is encouraged.

In conclusion, the hypothesis for this study that HUFs students would perceive vocal fry negatively is supported. With this finding it is suggested that English as a second or other language learners avoid mimicking this speech pattern as may not be in a speakers' best interest to employ vocal fry.

REFERENCES

- Anderson, R. C., Klofstad, C. A., Mayew, W. J., & Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PloS one*, 9(5), e97506. Retrieved from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0097506>
- Belotel-Grenié, A., & Grenié, M. (2004). The creaky voice phonation and the organisation of Chinese discourse. In *International symposium on tonal aspects of languages: With emphasis on tone languages*. Retrieved from https://www.researchgate.net/profile/Agnes_Belotel-Grenie/publication/239922940_The_Creaky_Voice_Phonation_And_The_Organisation_Of_Chinese_Discourse/links/53fc6590cf2dca8ffff190c.pdf
- Borrie, S. A., & Delfino, C. R. (2017). Conversational entrainment of vocal fry in young adult female American English speakers. *Journal of Voice*, 31(4), 513-e25. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0892199716303800>
- Fessenden, M. (2011). Vocal fry creeping into US speech. *Science Mag*, 9. Retrieved from <https://www.sciencemag.org/news/2011/12/vocal-fry-creeping-us-speech>
- Fraccaro, P. J., O'Connor, J. J., Re, D. E., Jones, B. C., DeBruine, L. M., & Feinberg, D. R. (2013). Faking it: deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour*, 85(1), 127-136. Retrieved from http://www.academia.edu/download/45357586/Faking_it_deliberately_altered_voice_pit20160504-15240-nu49bo.pdf
- Hollien, H., Moore, P., Wendahl, R. W., & Michel, J. F. (1966). On the nature of vocal fry. *Journal of Speech, Language, and Hearing Research*, 9(2), 245-247. Retrieved from <https://jslhr.pubs.asha.org/article.aspx?articleid=1749843>
- Johns Hopkins Medicine (n.d.). Is vocal fry ruining my voice?. Retrieved from <https://www.hopkinsmedicine.org/health/articles-and-answers/wellbeing/is-vocal-fry-ruining-my-voice>
- Williamson, G. (2015, November 9). Glottalization. Retrieved from SLTinfo website <https://www.sltinfo.com/allo101-glottalization/>
- Wolk, L., Abdelli-Beruh, N. B., & Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice*, 26(3), e111-e116. Retrieved from [https://www.jvoice.org/article/S0892-1997\(11\)00070-1/fulltext](https://www.jvoice.org/article/S0892-1997(11)00070-1/fulltext)
- 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. Retrieved from <https://read.dukeupress.edu/american-speech/article-abstract/85/3/315/5885/Creaky-Voice-A-New-Feminine-Voice-Quality-for>