2020

Vocal Fry: Acoustics, Airflow, and EGG Analysis of Various Types

Kiersten McCormick
Bowling Green State University, kierstm@bgsu.edu

Dr. Ronald Scherer
Bowling Green State University, ronalds@bgsu.edu

Katherine Proctor
Bowling Green State University, klproct@bgsu.edu

VOCAL FRY: ACOUSTICS, AIRFLOW, AND EGG ANALYSES OF VARIOUS TYPES

KIERSTEN MCCORMICK

HONORS PROJECT

Submitted to the Honors College
at Bowling Green State University in partial fulfillment of the requirements for graduation with

UNIVERSITY HONORS

Dr. Ronald Scherer, Communication Sciences and Disorders,
College of Health and Human Services, Advisor

Dr. Lewis Fulcher, Physics and Astronomy,
College of Arts and Sciences, Advisor

INTRODUCTION

Vocal fry is a vocal register that is produced at a frequency below that of one's typical modal or speaking vocal register. It is relatively low in intensity and appears as distinct pulses when viewing the microphone signal (Hollien, 1974). It was suggested in early speech science literature that the presence of vocal fry in speech should be classified as a voice disorder due to the roughness of its production and sound quality (Moore and von Leden, 1958). Since then, numerous aspects of vocal fry have been studied surrounding its cultural significance and physiological production mechanisms. Instead of being classified as a voice disorder, vocal fry eventually was deemed as a vocal register by Harry Hollien in 1966. It is important to note that the use of vocal fry, when excessive and consistent, can be indicative of a voice disorder (Hollien et al., 1996). However, when used linguistically or culturally, vocal fry is an important psychosocial and normal aspect of communication. Vocal fry can indicate a number of different social cues to listeners, including boredom or assertion of authority. It can also be used as a social linguistic tool to build relationships (Quenqua, 2012).

Vocal fry has a number of defining physiological, acoustic, perceptual, and aerodynamic properties. The physiological aspect of vocal fry used extensively in the current project is revealed through the electroglottographic (EGG) waveforms during phonation while a person is speaking. The study reports different patterns of the vocal fry phenomenon and their physiological and aerodynamic correlates. Current categories of vocal fry found in the research literature include: "prototypical fry", "multiply pulsed voice", "aperiodic voice" (Keating, Garellek, & Kreiman, 2015), and "onset fry" (Perrine, 2018). The current study reports on more patterns than suggested in the past.

Average acoustic properties of vocal fry include the relatively low fundamental frequencies ("pitch of the voice") of 49.14 Hz for males and 48.1 Hz for females. There is a statistically significant difference between the fundamental frequencies of vocal fry and those of modal register for both males and females. Vocal fry is also seen to be produced over a relatively small section of the entire length of the vocal folds. This suggests that it is unlikely that the difference in length of the vocal folds between males and females is related to any differences in the resulting frequency of vocal fry produced. However, the masses of the vocal fold tissue of males and females are inherently different, being greater in the male larynx, suggesting that mass differences could contribute to the gender differences of the fundamental frequency of vocal fry (Blomgren, Chen, Ng, and Gilbert, 1998).

Regarding the perceptual aspects of vocal fry, it is shown that listeners pick up on specific acoustic qualities that differentiate vocal fry from other vocal registers. The most important factors include turbulent noise, pulse skewedness, pulse width, glottal closure abruptness, and amount of closed time. When compared to modal, falsetto, and breathy voice, vocal fry has been shown to have low turbulent noise, high pulse skewing, short pulse width, and the most abrupt glottal closure (0.7% t) (Childers & Lee, 1991).

There has been little research done on the aerodynamic aspects of vocal fry. It has been shown that the average airflow is between 2.0-79.1 ml/s (McGlone, 1976). While this measure is important to note, there are still many aspects of the aerodynamics of vocal fry to be explored, including variants among categories of vocal fry and differences between males and females. Such a study has research and clinical importance and would enhance the understanding of the production of vocal fry.

The purpose of this study was to record individuals reading a prepared passage, note occurrences of vocal fry, and identify recently determined vocal fry patterns (from a recent Honor's project by Kate Proctor, BGSU) to determine the aerodynamic and glottographic correlates of the vocal fry productions. The goal was to highlight the origins of vocal fry productions, essentially explaining the typically studied acoustic records of vocal fry through how those acoustics were created through laryngeal function and corresponding aerodynamics. This report is of two examples for illustration of this phenomenon.

METHODS

This project was approved by the BGSU IRB, project number 1349239-2.

Subjects

The participants in this study included six young adults (3 males and 3 females) who had self-reported lack of a history of voice, speech, language, or hearing problems or therapy and were untrained relative to voice and speech performance. Each participant was also a native English speaker and was determined to be in good vocal and overall health on the day recordings were made.

Equipment

All voice recordings were conducted in an IAC sound treated booth (room 181 HHS) using a condenser microphone (C 420 III PP MicroMic from AKG Acoustics; frequency response 20-20,000 Hz). The aerodynamic measures oral air pressure and airflow were obtained using the Glottal Enterprises aerodynamic system (MSIF-2 S/N 2049S) which includes a face mask and an intraoral tube connected to a pressure transducer. An electroglottograph (KAY Model 6103) was used to obtain the EGG waveforms. To digitize the concurrent analog signals obtained (oral air pressure, airflow, EGG, and audio), a 14 bit DATAQ A/D converter (Model DI-2108 Series) was used with Windac Pro + software. Sigplot software (MATLAB-based) was used when analyzing pressure, airflow, and EGG, and PRAAT software was used when looking for instances of vocal fry. Both the microphone signal and the corresponding spectrogram were studied to determine vocal fry during the running speech.

Protocol

Participants were recruited using BGSU's SONA website. Students can access this site to volunteer to participate in various research studies in exchange for Sona credit in class. The subjects used the Sona website to sign up for a phone interview time. During the phone interview (with the PI, Kiersten McCormick), participants were screened relative to inclusion and exclusion criteria. Those who passed the phone screening were asked to choose a time to come into the Health and Human Services Building to fill out a preliminary heath questionnaire, sign the consent form, and perform the study.

When the participants arrived in the Health and Human Services Building, they completed the necessary paperwork and were given a rundown of the protocol. They were shown a demonstration of how to properly apply the flow mask to their face and had EGG sensors placed on both sides of their laryngeal thyroid cartilage by a member of the research team. Participants were asked to repeat a string of five /pa/ repetitions at normal conversational loudness two times: first without the mask and then using the mask. This was intended to familiarize subjects with the mask and to ensure there was no air leakage around the rim of the mask during speech. Participants then read the Rainbow Passage in a normal manner at the same loudness level two times: first without the mask and then using the mask.

Analysis

Occurrences of vocal fry were counted based on auditory and visual indications. According to Hollien et al. (1965), vocal fry is commonly described as "rough" and having a distinct "popping" characteristic. Using this auditory criterion, research team members Kiersten McCormick and Ronald Scherer marked the instances of auditory vocal fry present in each recording of the Rainbow Passage. The primary characteristic was the prominent negative pressure spike of the microphone signal and its associated dark line energy excitation in the spectrogram. The members of the research team then analyzed the recordings for occurrences of vocal fry and used the new vocal fry pattern designations to label and study the productions.

RESULTS

The vocal fry categorization scheme in which data were analyzed includes various categories and sub-categories of vocal fry based on the pattern of acoustic pulses (i.e., negative acoustic pressure transients). The present paper focuses on the relationship between the microphone signals of different categories of vocal fry and their corresponding electroglottographic and airflow signals of two of the most commonly occurring types: Single Pulse Fry and Double Pulse Fry.

Single Pulse Fry

Single pulse fry is defined as the production of only one pulse per vocal fry period. The fundamental frequency of each complete vocal fry cycle was found to be at least one octave below the speaker's fundamental frequency in the reading immediately before or after the production of the pulses in question for them to be categorized as vocal fry. Single pulse fry can either occur in isolation, termed *isolated single pulse fry*, or in a series of single pulses, which can be further categorized into two distinct sub-categories: *even single pulse sequences* and *uneven single pulse sequences*. Criteria for an *even single pulse sequence* include each pulse having the same or similar periods of damping in between them (i.e., the period of each vocal cycle must be nearly unchanging between each pulse) and the amplitudes of each pulse must be consistent with one another. In contrast, criteria for an *uneven pulse fry sequence* requires variation amongst the period or amplitude between pulses.

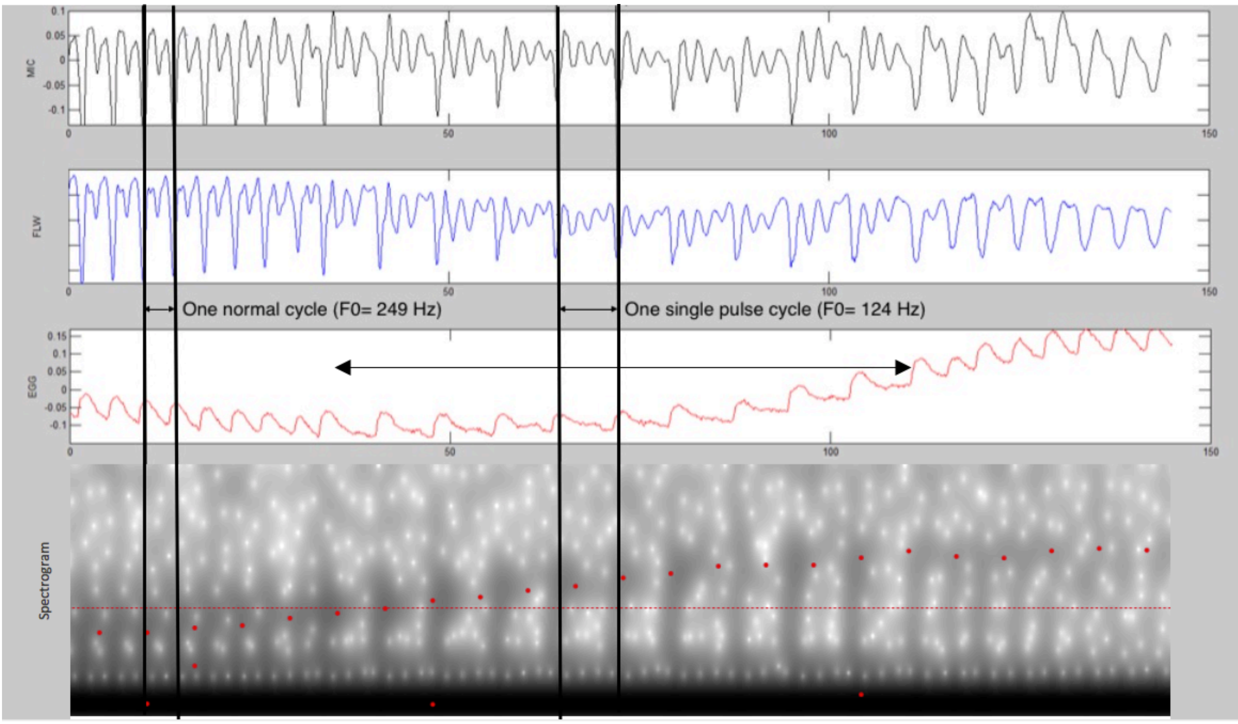

Figure 1: Microphone, Airflow, Electroglottographic, and Spectrogram signals of an instance of an *even single pulses fry sequence* from participant F3. The figure shows a sequence of 10 single pulses of vocal fry (between the arrow tips). The vertical bars on the left show the period for the normal fundamental frequency in modal vocal register produced by the subject prior to producing vocal fry, and the vertical bars in the middle of the figure indicate the period of a single instance of a single pulse fry. The subject is saying "**is**" at the time of the vocal fry.

Microphone

Analysis of the acoustic microphone signal was the initial indicator to which this instance of vocal fry belongs. As indicated in **Figure 1**, the period of each vocal fry pulse is approximately one octave below the participant's fundamental frequency immediately prior to the production of vocal fry, confirming that this series of pulses qualify for vocal fry categorization. There is only one strong acoustic transient present per vocal fry cycle as seen as the vertical dark bands in the spectrogram below the EGG waveform, and the amplitude and period of each pulse is fairly uniform across each pulse. This indicates that **Figure 1** is an example of *even single pulse fry*.

Electroglottograph

Analysis of the electroglottograph (EGG) signal confirms that this instance of vocal fry belongs to the *even single pulse* category. The EGG waveform indicates that the vocal folds made contact one time per each vocal fry cycle, where the rise of the EGG waveform matches the closure of the glottis and the corresponding airflow peak and acoustic excitation peak. The EGG waveform essentially explains the creation of the vocal fry, the cause of the duration of the vocal fry cycle, and the corresponding acoustic pressures and airflows per vocal fry cycle. The EGG waveform suggests that the fry production is produced by a hyperadducted glottis that has a long-closed portion before opening for a short segment of time and then closing to again repeat the relatively long glottal closure time.

Airflow

The airflow signal shown in **Figure 1** gives insight as to the mechanisms explaining the glottal and acoustic phenomena responsible for *even single pulse fry*. The airflow signal is a wideband signal (frequency range 0 Hz to about 1500 Hz) and mimics the waveshape of the microphone signal, since the airflow is actually an acoustic airflow signal. The airflow signal is related to the acoustic pressure signal through acoustic impedance at the surface of the flow mask.

Double Pulse Fry

Double pulse fry is defined as the production of two pulses per vocal fry period, where the duration of damping between each pulse in the vocal fry period must be independent of the speaker's normal fundamental frequency in the immediately surrounding "normal" pulses. Each of the two pulses in the cycle can be labeled as either a *primary pulse* (i.e., the pulse with the greater amplitude relative to the other pulse in the period) or a *secondary pulse* (i.e., the pulse with the smaller amplitude relative to the other pulse in the period). The patterns in which the primary and secondary pulses appear in each pulse determine which subcategory the double pulse cycle belongs. There are three sub-categories of double pulse fry, including *primary-secondary double pulse* in which the primary pulse precedes the secondary pulse, *secondary-primary double pulse* in which the secondary pulse precedes the primary pulse, and *primary-primary double pulse* in which the double pulse consists of two consecutive primary pulses that have approximately the same amplitude.

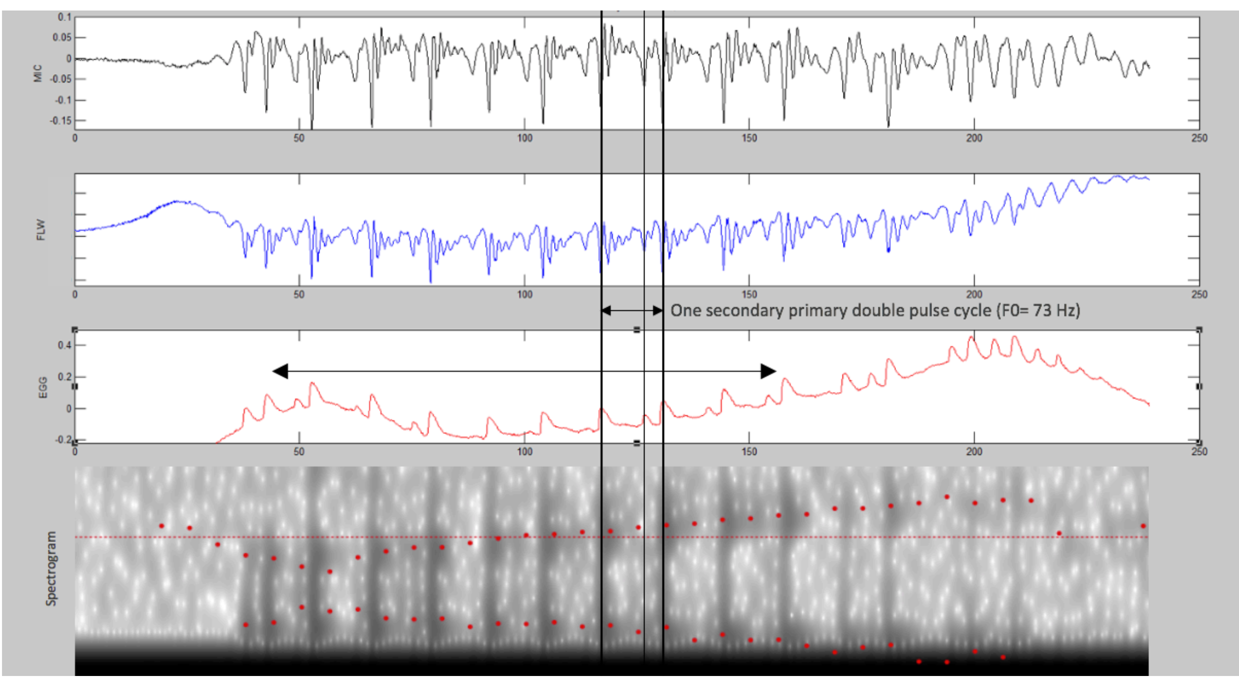

Figure 2: Microphone, Airflow, and Electroglottographic signals of an instance of *secondary-primary double pulse fry* from participant F1. The arrow indicates an eight double pulse fry sequence. The subject is saying "**sorry**" at the time of the vocal fry.

Microphone

Analysis of the acoustic microphone signal allowed for the initial categorization of this instance of vocal fry as *secondary-primary double pulse fry*. As indicated in **Figure 2**, the period of each vocal fry cycle is lower than one octave below the participant's average fundamental frequency immediately following the fry production, confirming that this series of pulses qualify for vocal fry categorization. Furthermore, the period in between each pulse in the overarching vocal fry cycle is independent of the participant's average fundamental frequency immediately following the fry production, indicating that the two pulses in the fry cycle qualify as *double pulse fry*. Lastly, the appearance of two pulses in the secondary-primary pattern is indicative that this is an instance of *secondary-primary double pulse fry*. The arrow in the figure shows the location of the 8-cycle sequence of the *secondary-primary double pulse fry* example.

Electroglottograph

Analysis of the electroglottograph (EGG) signal further confirmed that this instance of vocal fry belongs to the *secondary-primary double pulse* category. It can be seen in **Figure 2** that both the secondary pulse and the primary pulse in each vocal fry cycle as seen in the acoustic microphone signal directly corresponds to a peak in the EGG signal cycle. This confirms that the source of the secondary and primary acoustic excitation is due to a double-motion (lateral-medial) of the vocal folds that creates the double-peak airflow signal to produce double pulse fry. Furthermore, it can be seen in **Figure 2** that each secondary pulse corresponds to a peak in amplitude on the EGG that is notably small relative to the primary pulses of both signals. This suggests that the secondary pulses in each double pulse fry requires a weaker point of contact in terms of surface area made by the vocal folds in their production relative to the production of primary pulses. While there is not a concrete relationship between amplitude of acoustic microphone signals and amplitude of EGG signals, there is a notable correlation between the two.

Airflow

The airflow signal again confirms the similarity between the airflow waveshape and the microphone waveshape as discussed above. For the double pulse example, as noted above, the double pulsing of the airflow is consistent with the double pulsing of both the EGG signal and the acoustic signal.

DISCUSSION AND CONCLUSIONS

Comparison of Vocal Fry Patterns

The patterns of vocal fry in this study were categorized based on a set of criteria described in a previous study with Proctor et al. (2019). While the present paper discusses only two of the six categories from that study, similar confirmations have been made with the other categories as well (and will be discussed in a manuscript for publication).

Single Pulse Fry:

Proctor et al. (2019) describes single pulse fry as, "when there is only one pulse (i.e., one negative acoustic pressure transient) per vocal fry period". Using information gathered from the EGG and airflow signals in the present study, it can be confirmed that single pulse fry is created by a single glottal pulse, and the description provided in Proctor et al. (2019) is accurate.

Double Pulse Fry:

Proctor et al. (2019) describes double pulse fry as, "when there are two pulses per period of vocal fry production rather than just one". The EGG and airflow signals gathered in the present study support this description, as two distinct glottal pulses can be seen in the signals. Therefore, it can be confirmed that double pulse fry is caused by two distinct glottal pulses (provided by the EGG signal), and the description provided in Proctor et al. (2019) is accurate.

Relevance of EGG and Airflow

The EGG and airflow signals give researchers more insight as to the laryngeal behavior responsible for the production of each type of vocal fry. The EGG waveform provides information detailing the laryngeal control, specifically how different vocal fold adductory patterns (revealed by the EGG signal) correspond to the acoustic and airflow signals. In addition, information about the adduction level of the vocal folds between the speaker's normal fundamental frequency and their fundamental frequency immediately before the production of vocal fry helps researchers determine when adduction may be the highest. This further illustrates the laryngeal behavior behind each type of vocal fry.

Clinical Relevance

Research in this area in particularly relevant in the area of voice pathology. Understanding both the typical laryngeal behavior responsible for each type of vocal fry and the typical vocal fry patterns present in various vocal pathologies could contribute to a more accurate and well-informed clinical vocal diagnosis. Further research can be done in the area of vocal fry relative to various voice pathologies to gain a deeper understanding of these potential relationships.

Understanding the laryngeal behavior behind different types of vocal fry can also aid in the creation of interventions for the treatment of various pathologies related to vocal fry when necessary. If clinicians know the patterns of vocal fry a particular client is using and reliable treatment strategies for each type, they will be able to tailor treatment sessions to fit the client's individual needs. This has the potential to increase treatment effectiveness, however more research is needed to construct applicable intervention strategies.

Limitations

A limitation of this study is the rather small sample size of three males and three females of the same general age group and region of the United States. It would be beneficial to study more individuals who vary in age and region in order to get a better understanding of vocal fry patterns across an inclusive sample of individuals. However, it is noted that large samples appear unnecessary when illustrating the various types of vocal fry due to the consistent nature of the glottal cyclic action revealed by the EGG signal and consistent correspondence to the wideband airflow and acoustic (microphone) signal.

REFERENCES

Blomgren, M., Chen, Y., Ng, M. L., & Gilbert, H. R. (1998). Acoustic, aerodynamic, physiologic, and perceptual properties of modal and vocal fry registers. *The Journal of the Acoustical Society of America*, *103*(5), 2649-2658.

Childers, D. G., & Lee, C. K. (1991). Vocal quality factors: Analysis, synthesis, and perception. *the Journal of the Acoustical Society of America*, *90*(5), 2394-2410.

Hollien, H. (1972). "On vocal registers". *Journal of Phonetics*. 2,125-143.

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.

Keating, P., Garellek, M., & Kreiman, J. (2015, August). Acoustic properties of different kinds of creaky voice. In *Proceedings of the 18th International Congress of Phonetic Sciences* (pp. 0821-1).

McGlone, R. (1967). "Air flow during vocal fry phonation". *Journal of Speech and Hearing Research*. 10, 299–304.

Moore, P., and von Leden, H., Dynamic variation of the vibratory pattern in the normal larynx. *Folia Phontrica*, 10, 205-238 (1958).

Perrine, Brittany. (2018, May). *The Influence of Stress on the Voice* (Doctoral Dissertation).

Quenqua, Douglas. (2012). "They're, like, way ahead of the linguistic currrrve". The New York Times. Retrieved from https://www.nytimes.com/2012/02/28/science/young-women-often-trendsetters-in-vocal-patterns.html