VOICE ONSET TIME IN CHILDREN WITH AND WITHOUT VOCAL FOLD NODULES

A Thesis
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
MASTER OF ARTS
OF SPEECH, LANGUAGE, AND HEARING SCIENCE

---

by
Lauren A. Colletti
May 2022

Thesis Approvals:

Elizabeth Heller Murray, Thesis Advisor, Communication Sciences and Disorders
Edwin Maas, Communication Sciences and Disorders
Susan Caspari, Communication Sciences and Disorders

ABSTRACT

Purpose: This study examined voice onset time (VOT) in children with and without vocal fold nodules (VFN). The purpose of this study was to provide further evidence regarding the need for individualized research and treatment dedicated to the pediatric population. The pediatric population has a distinctly different laryngeal mechanism than adults, as they are still developing. Although the pediatric system is anatomically different from that of a fully mature adult system, treatment for children with VFN is largely based on adult research. This study examined the VOTs of voiceless consonants, as the transition from the voiceless consonant to the subsequent vowel requires significant vocal and articulatory control and coordination. Measures of VOT change throughout the maturation as VOT follows a significant developmental pattern. Children with and without VFN were enlisted in order to examine the effects VFN have on VOT.

Hypotheses: We hypothesize that children with VFN would have differences in 1) average VOT values compared to the control group, with no prediction for direction of difference (shorter and longer), and 2) between-word variability of VOT values compared to the control group, with no prediction for direction of difference (more variable and less variable).

Methods: Participant data were retrospectively collected and included children between 6 and 12 years old with VFN and age- and sex-matched controls. Participants were recorded producing the six CAPE-V sentences. Four voiceless consonants were selected for VOT analysis. Praat was utilized to manually mark the vocal onset of the stop consonant by the current researcher. A previous researcher identified the vocal offset, and each placement was confirmed by the current researcher. VOT was calculated as the time between the stop consonant burst and the vocal onset of the vowel.

Results: There was no significant difference between the VFN and the control groups in average VOT or VOT variability. Within the VFN group, participants who were more dysphonic (lower cepstral peak prominence (CPP) values) had more variable VOT values. Participants in the VFN group had lower CPP values than the control group, suggesting that CPP measures are a reliable indicator of dysphonia. Additionally, within the VFN group, male children had lower CPP values than female children.

Conclusion: Although no group difference was found, the within-group analyses indicated that VFNs impacted productions. Children with VFN who were more dysphonic had increased VOT variability. This may suggest that VFN impact a child's ability to phonate therefore causing more variability within productions. Future research is needed to study the impact dysphonia treatment for children with VFN may have on VOT values. Additionally, a longitudinal study of the impact of VFNs on VOT values during developmental stages may be warranted.

ACKNOWLEDGEMENTS

I’d like to acknowledge my thesis advisor Dr. Elizabeth Heller Murry for guiding me through not only this Master’s Thesis but supporting me throughout my graduate school career. Your clear passion for your work is infectious and inspirational. I would also like to thank my committee members Dr. Edwin Maas and Susan Caspari for their continued dedication to their work and the education of young clinicians in the field. Thank you all for supporting me throughout this process and always believing in me. I would like to acknowledge Andie Chao for her prior work, which has allowed me to complete this thesis.

Additionally, I would like to acknowledge my fellow graduate cohort for their continued dedication to the field and their continued persistence in completing our education throughout this unprecedented time. To my friends who I have had the immense pleasure of coming to know during my time in graduate school; I would have not made it through this process without you. I found solace in knowing that I had true friends to fall to in times of need. I would also like to acknowledge my friends from college who have seen the journey I have been on since we first met as undergraduates. Thank you for supporting me through the years and continually showing up for me. Lastly, I would like to extend my most sincere gratitude to my family who has been my biggest support system throughout my life. Thank you for helping me through this journey in so many ways. I know I can’t say thank you enough because my gratitude for what you have done for me is endless.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Voice Disorders in Children

The overall prevalence of voice disorders in children ranges from 1 to 30 % (Kiliç et al., 2004; Leeper Jr et al., 1980; Tavares, et al., 2011). The reported variable rates may be partially attributed to different data collection methods, as prevalence statistics from parental reports are lower than those acquired from clinical data (Tavares, et al., 2011). Vocal fold nodules (VFN) are the most prevalent voice disorder in children, seen in 17% to 30% of school-aged children (Ongkasuwan & Friedman, 2013; Shah et al., 2005; Tavares et al., 2011; Tezcaner et al., 2009). VFN are a bilateral thickening of the outermost tissue of the vocal folds, known as the epithelium. They occur primarily at the initial point of vocal fold contact, between the anterior one-third and two-thirds of the vocal folds (Kiliç et al., 2004; Sonbay Yılma et al., 2021; Tezcaner et al., 2009). The presence of VFN can cause dysphonia, that is, an altered vocal quality noted in the auditory-perceptual domain. If the VFN are large enough, they can prevent the vocal folds from completely closing during phonation, often creating an hourglass closure pattern. This allows for the early escape of air through the anterior and posterior portions of the vocal folds, leading to a perceptually breathy vocal quality. VFN can also lead to a perceptually strained voice as children may use excessive muscle tension or hyperfunctional vocal behavior to phonate if the VFN prevents full closure. Overall, VFN are the leading cause of chronic dysphonia in the pediatric population (Nuss et al., 2010; Shah et al., 2005).

VFN can occur at any point during childhood, with cases documented in children as young as a few months of age (Shah et al., 2005). In children 3-12 years of age, boys are twice as likely to be diagnosed with VFN than girls (Angelillo et al., 2008; Shah et al., 2005). Despite this, the prevalence of VFN in adulthood is significantly higher in females than males (Angelillo et al. 2008; Shah et al. 2005; Sonbay Yılma et al. 2021), adding further support to the differences between children and adults with VFN. Treatment for dysphonia is essential as chronic dysphonia can negatively impact a child's voice use, behaviors, school performance, social participation, and other aspects of health and daily life (Connor et al., 2008).

VFN can develop from vocal abuse or misuse of the vocal mechanism (Kiliç et al., 2004). Vocal abuse, also called phonotrauma, can occur in situations conducive to yelling, such as participating in sports or speaking in noisy environments. Vocal misuse is when the individual uses the vocal mechanism in an inefficient or inappropriate way, such as speaking at a pitch that is too high/low or using increased vocal strain (Hillman et al., 1989). VFN etiology can also be exacerbated by conditions such as respiratory reflux and allergic laryngitis (Shah et al., 2005). Respiratory reflux and allergic laryngitis can cause laryngeal mucosa edema, making the vocal folds more susceptible to phonotrauma (Zeitels & Healy, 2003). These etiologies in VFN are seen in children and adults; however, the laryngeal system of children is distinctly different, potentially making children's vocal folds more susceptible.

There are significant structural differences in the laryngeal mechanism between children and adults. Children's vocal folds are smaller than adults, with differences in microstructure and elasticity, with the mature three-layer vocal fold structure not fully emerge until around seven years of age (Hartnick et al., 2005). Furthermore, the developing vocal system interacts with a developing articulatory system during speech production (Kent, 1976). Thus, it is imperative that treatment for VFN in children is not based solely on adult voice therapies, but is optimized for the pediatric mechanism. This thesis will examine the interaction of two developing systems critical for producing intelligible speech, the vocal and articulatory systems.

Voice Onset Time Development

Voice onset time (VOT) is the time between the release of the stop consonant and the initiation of the subsequent vowel (Lisker & Abramson, 1964). This thesis will focus on VOTs of voiceless consonants, as they require intricate control and coordination of the vocal and articulatory systems. During the production of a voiceless consonant, the vocal folds remain abducted (open). Then, they adduct (close) and begin vibrating to produce the vowel. Therefore, productions that involve a voiceless consonant followed by a vowel require a great deal of coordination from the articulatory and vocal mechanisms.

In English-speaking adults, VOT is an acoustic temporal cue used by listeners to determine if a consonant is voiced or voiceless (Whiteside et al., 2004). Adults produce voiced stops (/b/, /d/, /g/) with an average VOT of 15 milliseconds (ms) or less, while voiceless stops (/p/, /t/, /d/) have an average VOT of 30-40 ms (Hitchcock & Koenig, 2013). However, children's VOTs are different than adults, with a clear development pattern seen in early childhood. VOT development can be described using four primary features: accuracy, discreteness, range, and overshoot. *Accuracy* for VOT is compared to typical VOT in adults, 20ms for bilabials, and 30ms for alveolars, with any production over this cutoff considered inaccurate. Children typically produce accurate VOTs (i.e., comparable to adults) around three years of age (Hitchcock & Koening, 2013). *Discreteness* is the distinction between voiceless and voiced consonants. As children become more accurate in their productions, voiced and voiceless cognates become more discrete. In 9-18-month-olds, there is typically no distinction between voiced and voiceless cognates. Acoustic analyses indicate these cognates begin becoming distinct around 18- 28 months of age (Macken & Barton, 1980; Whiteside & Marshall, 2001).

Children begin to show a decreased range of VOT productions, less VOT variability across, and increased VOT accuracy during this early childhood period. Children 6-18 months old typically have VOTs that *range* from 0 to 30 ms (Whiteside & Marshall, 2001). As children further develop and begin decreasing the range, they also may demonstrate *overshoot* of voiceless stops, where children produce these phonemes for longer than 100ms (Hitchcock & Koenig, 2013; Whiteside & Marshall, 2001). Overshoot begins to decrease around the age of four, yet considerable production variability remains (Macken & Barton, 1980; Whiteside & Marshall, 2001; Yu et al., 2015). Around six years of age, average VOT values are comparable to adults (Kent, 1976; Whiteside & Marshall, 2001).

Although average VOTs values are comparable to adults around six years of age, considerable VOT variability persists until around 8-11 years of age (Gilbert & Purves, 1977; Whiteside & Marshall, 2001; Yu et al., 2015; Zlatin & Koenigsknecht, 1976). Moreover, this time period of decreasing VOT variability (e.g., 8-11 years) coincides with the prepuberty/puberty phase, where vocal sexual dimorphism begins to emerge in children. Males aged 8-11 years produce larger VOT values for plosives and overall had more variable productions than those of females (Whiteside & Marshall, 2001; Yu et al., 2015). Sex differences affecting VOT may be related to the differing anatomical changes during development (Whiteside & Marshall, 2001; Yu et al., 2015). Males undergo rapid laryngeal changes during early adolescence, causing their fundamental frequency (f_o) to decline at the onset of puberty. These laryngeal changes may also affect vocal fold function and therefore impact VOT values (Hasek et al., 1980; Seikel et al., 2000; Whiteside & Marshall, 2001). It is important to note that these studies have not yet considered gender identity or societal norms on speaking behavior based on gender. However, they are discussed to outline additional context to differences that may be noted in children.

VOT and Dysphonia

Only a few studies have examined VOT in individuals with dysphonia. McKenna and colleagues (2020) examined VOT in adults with voice disorders. They found that participants with vocal hyperfunction exhibited more variable VOTs than a cohort of age- and sex-matched vocally healthy individuals. The authors also found preliminary evidence for shorter voiceless VOTs in speakers with more dysphonic voices, defined as having a Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V; Kemester et. Al., 2009) overall severity rating ≥ 25 (McKenna et al., 2020). Heller Murray and Chao (2021) examined the relationship between the variability of VOT and dysphonia in children. Although the voice disorder status of these children was unknown, perceptual evaluation was used to categorize children with CAPE-V overall severity scores greater than 22.5 as dysphonic. The authors found a negative correlation between VOT variability and vocal variability within the non-dysphonic group, but no significant correlations were found in the dysphonic group. Upon further investigation, the authors also found that children who presented with more severe dysphonia had increased vocal variability. The authors suggested that dysphonia severity may impact the relationship between vocal variability and VOT variability (Heller Murray & Chao, 2021). However, additional work is needed to examine this relationship in children diagnosed with a voice disorder. Both the typical developmental trends of VOT development and the relationships between VOT and dysphonia suggest that understanding VOT and VOT variability measures would be valuable in children with VFN.

Cepstral Peak Prominence

In order to clinically evaluate and diagnose dysphonia, many clinicians employ an auditory-perceptual evaluation (e.g., CAPE-V). Auditory-perceptual measures utilize a clinician’s subjective judgment of a speaker’s vocal quality (e.g., breathiness, hoarseness, etc.), pitch, and loudness based on their age and gender presentation. Due to the subjective nature of auditory-perceptual evaluations, objective measures of dysphonia are used in conjunction with auditory-perceptual evaluations to provide a reliable dysphonia severity rating. The acoustic measure of cepstral peak prominence (CPP) is recommended for an objective measurement of dysphonia (Patel et al., 2018; Murton et al., 2020). CPP values provide clinicians with support for their subjective dysphonia ratings and have been shown to be reliable objective measures within both the adult and pediatric populations (Aydinli et al., 2019; Infusino et al., 2015; Murton et al., 2020). Although a majority of the work is done on adults, a recent study conducted by Aydinli and colleagues found a correlation between CPP and dysphonia in children (Aydinli et a., 2019).

Unlike time-based measures that examine cycle-to-cycle changes in pitch (jitter) or amplitude (shimmer), CPP does not rely on accurate tracking of the fundamental frequency. This is important for individuals with voice disorders, as dysphonic voices can have reduced periodicity or increased noise, making these time-based measures unreliable. CPP is derived from the cepstrum, previously describe as a “spectrum of a spectrum” (Murton et al., 2020). To examine a signal in the cepstral domain, the time domain waveform undergoes its first Fourier transform, converting it to the spectral domain. Next, the logarithm of the spectrum is taken, and an additional Fourier transform is done to convert the signal to the cepstral domain (Herman-Ackah et al., 2003; Hillenbrand et al., 1994; Hillenbrand & Houde, 1996). After this transformation, the cepstral domain allows the user to examine the degree of harmonic organization, e.g., how periodic the signal is. The measure of CPP is calculated in the cepstral domain, and provides a measure of how high the cepstral peak (associated with the fundamental period) emerges from the cepstral noise (Hillenbrand et al., 1994; Hillenbrand & Houde, 1996). A signal with a low CPP value, as seen in dysphonic voices, is one in which the amplitude of the cepstral peak is not very distinct from the rest of the vocal noise, indicating decreased periodicity. Healthy voices will have high amplitude cepstral peaks which are distinct from the noise in the signal.

Stop Consonant Identification

In order to reduce measurement variability, this thesis focused on strict criteria for identifying the boundaries of a VOT production, including identification of the stop consonant burst and identification of the vocal onset. The beginning of the VOT requires accurate measurement of the burst, which is the release of the stop consonant closure. Previous work has highlighted three primary methods for identifying the stop consonant burst: wideband spectrograms, airflow oscillograms, and waveforms from the microphone signal. The first method is using a wideband spectrogram of the production measured from a microphone signal (Figure 1). A wideband spectrogram is optimal for temporal information about the signal rather than the signal's frequency information, which would be more optimally viewed using a narrowband spectrogram. The user would then identify the portion of the spectrum with an abrupt change, depicting vertical broadband noise (Lisker & Abramsons, 1964).

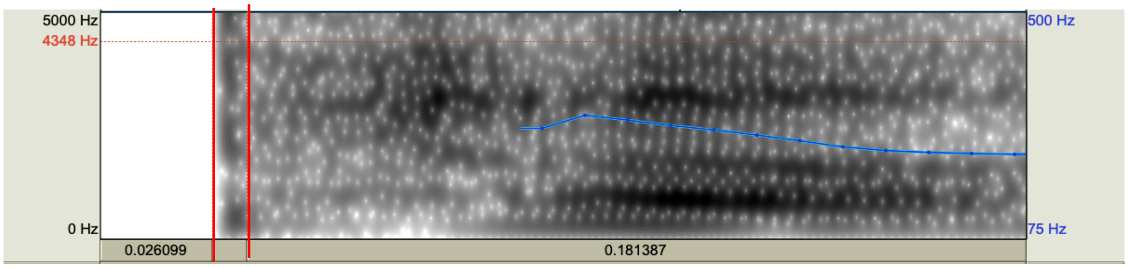

Figure 1. Wideband Spectrogram of production /pa/ with stop consonant identified between red lines. Blue line indicates the pitch of the vowel.

The second method that also uses a microphone signal is the examination of the waveform. Utilizing this method, the user would identify the first abrupt change as the stop consonant and mark the spot where the noisy signal begins (Brown et al., 1993) (Figure 2). To reduce measurement variability, some researchers noted they mark the waveform at the last spot at the zero crossing before a sudden change in the waveform (Karlsson et al., 2004).

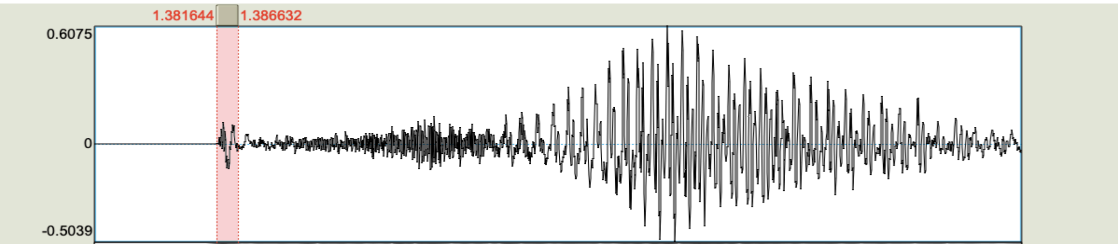

Figure 2. Waveform of production /pa/ with stop consonant identified in highlighted red area.

The third method to identify stop consonant bursts uses airflow oscillograms to examine intraoral pressure changes (Koenig, 2000). The airflow is collected via a face mask attached to a pneumotachograph. The onset of a stop consonant can be identified when viewing the oscillogram and locating the beginning of the precipitous fall in the air pressure wave. However, this method requires the purchase of expensive, specialized aerodynamic equipment. As there is not a significant benefit of this method over the other two, this additional expense does not appear necessary for accurate analysis.

Many researchers have proposed using both method 1 (wideband spectrogram) in conjunction with information from method 2 (waveform) to accurately identify the stop consonant burst (McCrea & Morris, 2005; Petrosino et al., 1993; Whiteside et al., 2004). We have chosen to do this combination of methods in the current thesis. Expanded further in the Methods section, wideband spectrograms were used to broadly identify burst location. The waveform was then examined, and the zero-crossing point before the sudden change in waveform shape was identified.

Vocal Onset Identification

The end of the VOT is the initiation of voicing of the vowel. Vocal onset identification provides insight into the coordination of the vocal folds during the coarticulation of a voiceless consonant into the start of the vowel. The primary methods for identifying vocal onset are in the wideband spectrogram, examining the waveform of a microphone, or examining the waveform of an accelerometer signal. Previous work has used wideband spectrograms in the identification of vocal onset. Vocal onset can be determined by identifying glottal pulsing represented by the first regularly spaced vertical striations on the spectrogram (Klatt, 1975; Lisker & Abramson, 1964; Watson & Alfonso, 1982).

Another method of detecting vocal onset is by using the waveform of an accelerometer or contact microphone. An accelerometer collects sound wave information by recording a signal from the neck's surface (Lien & Stepp, 2014). An accelerometer helps to reduce extraneous high-frequency noises such as vocal tract resonance and environmental noise, but there is a higher chance of movement artifacts since the microphone is placed on the neck. While accelerometers have some advantages over more traditional microphone signals, these differences are not big enough to warrant purchasing of specialized equipment (Lien & Stepp, 2014). Furthermore, these signals are not routinely recorded, and thus this method cannot be used in most retrospective analyses.

The final, and most common, method of identifying vocal onset time is the examination of the waveform from a microphone signal. Previous work has shown that this is the most accurate if the microphone signal is low-pass filtered, as the unprocessed signal contains noise from the stop consonant and the environment that can impede accurate detection of the vocal onset (Lien & Stepp, 2014). Low-pass filtering attenuates the signal above a determined frequency. This helps to reduce excessive noise caused by environmental factors and results in a more reliable identification of vocal onset than the unprocessed signal (Lien & Stepp, 2014). The current thesis used a low-pass filtering technique of the microphone signal and identified the start of the vowel in the waveform of this signal (see Methods for more details).

Current Work

The current study will utilize data collected from age- and sex-matched children with and without VFN between 6 and 12 years of age. We hypothesize that children with and without VFN will have a difference in average VOT. Children with VFN may have a shorter average VOT (due to anticipated difficulty during phonation) or a longer average VOT (due to difficulty initiating phonation in the presence of VFN). We also hypothesize that children with VFN will have differences in VOT variability, either reduced VOT variability (due to reduced vocal flexibility in the presence of VFN) or increased VOT variability (potentially due to decreased vocal fold control and increased vocal fold instability).

2. METHODS

Participants

Data from children with voice disorders were obtained from a clinical database at Boston Children's Hospital. Participants included in this thesis were 20 children with VFN (11 boys, 9 girls, average 8.8 years). Initial participant selection was done for a previous study conducted with researchers at Boston University (Heller Murray et al., 2020). Participants with VFN were retrospectively selected from the clinical database with the following inclusion criteria; (1) between 6.0 and 12.5 years of age; (2) had a primary diagnosis of vocal fold nodules from a board-certified laryngologist; (3) no prior voice therapy history; (4) received an overall voice severity score greater than or equal to 25 rated on CAPE-V (Kempster et al., 2009) determined by a certified speech-language pathologist during the initial clinical evaluation; (5) no history of other speech, language, or hearing concerns noted during evaluation; (6) usable, high-quality voice recordings were obtained during the initial clinical evaluation; and (7) accents were representative of a fluent English speaker from the northeast region. The inclusion criteria for the current study were consistent with this previous study.

A control group of 20 children without VFN (11 boys, 9 girls, average 8.9 years) were selected for this thesis. Control participants were originally recruited from Boston and its surrounding communities for the original study. Participants spoke English as their primary language, had no history of a voice disorder, and had not received speech or language therapy within the last year. Children aged 7;0 and older provided verbal assent and dissent from children under 7;0 was respected, while guardians provided written consent. Children without VFN underwent and passed an audiometric hearing screening at 25 dB or greater at the following frequencies: 0.25, 0.5, 1, 2, 4, 6, & 8 kHz. Participants in the VFN and control group were age and sex-matched.

Recording Procedures

All recordings were completed in a sound-treated room and were performed by a certified speech-language pathologist. Recordings done for children with VFN were completed during their clinical evaluation with the Computerized Speech Lab (Pentax Medical), with a 32.0 kHz sampling rate and a 16-bit resolution. Information about the microphone used during recordings was not available. Recordings done for the control group were conducted with a dynamic headset microphone (model WH20XLR) and acquired with the MOTU Ultralite mk3 hybrid soundcard (MOTU, Cambridge, MA, USA), sampled at 44.1 kHz with a 16-bit resolution.

Children repeated each of the six CAPE-V sentences one to three times. Four vocal onset speech segments were parsed from the recordings to analyze VOT (Table 1). Consistent with previous work (e.g., Heller Murray & Stepp, 2020; Lien & Stepp, 2014), the acoustic samples were low-pass filtered using a fifth-order Butterworth filter. A cutoff value of 680 Hz was selected for the filter, as it was 100 Hz higher than the highest f_o measured in the sample. This filtering aimed to reduce extraneous noise from the vocal tract and environment, thus making the vocal cycles easier to identify.

Table 1.

Stimuli recorded, CAPE-V sentences

CAPE-V Sentences	Analysis section	Vocal onsets
The blue spot is on the key again.	key	/ki/
Peter will keep at the peak.	Peter	/pi/
	keep	/ki/
	peak	/pi/

Data Processing and Acoustic Analysis

The start (stop consonant burst) and end (start of the vowel) were manually identified in Praat (Boersma & Weenick, 2022). Custom Praat scripts opened the recorded stimuli and the TextGrid area underneath the spectrogram for the manual analysis (Figure 3).

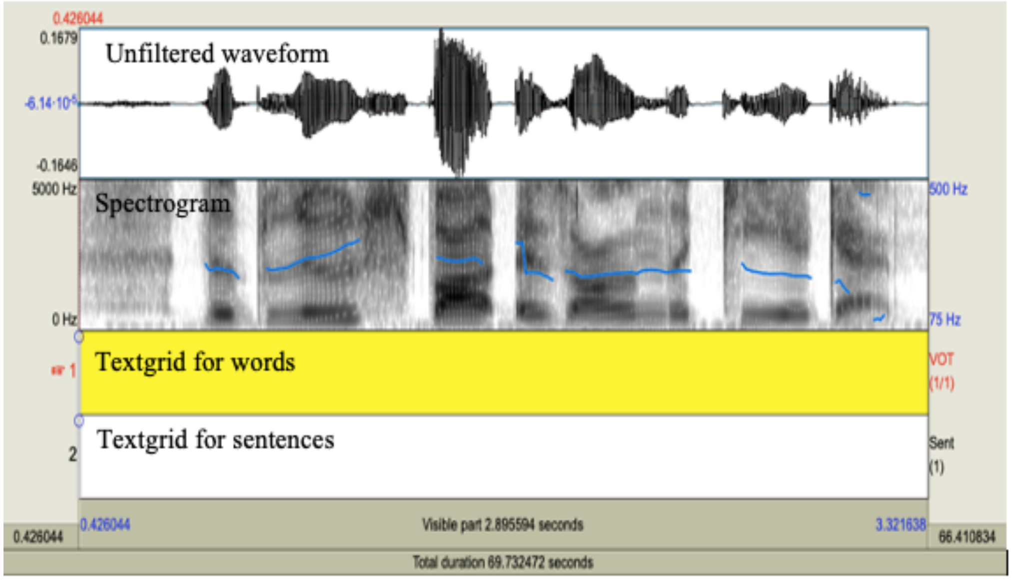

Figure 3. Example of the unfiltered waveform, spectrogram, and textgrids in Praat.

The onset of the stop consonant burst was identified on the unfiltered (raw) signal. By using the unfiltered signal, both perceptual and visual examination of the burst could occur, as low-pass filtering would distort the sound and remove some of the desired noise needed for identification. First, the stop consonant was identified perceptually by listening to the speech sample. Second, the spike in energy in the unfiltered waveform (characteristic of the burst) was identified. Third, the spectrogram was examined to confirm the waveform location is consistent with the location of the dark vertical band in the spectrogram. The identified area was then enlarged on the waveform, and the cursor was placed at the 0 crossing on the y-axis before the onset of the burst (Figure 4).

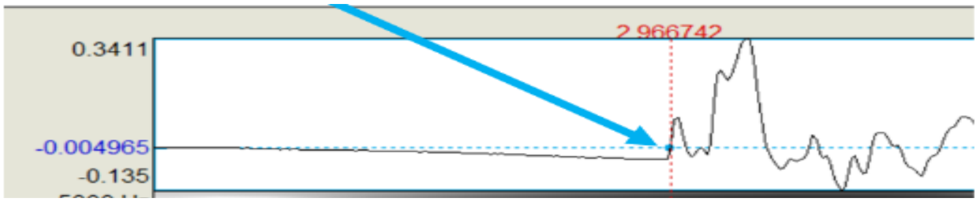

Figure 4. Blue arrow representing 0-crossing at y-axis.

Vocal onset identification was completed by a previous researcher. The current researcher confirmed vocal onset. Identification for vocal onset was performed on the filtered signal, with the unfiltered signal remaining present for auditory-perceptual and visual examination. Identification of vocal onset occurred after the stop consonant was identified, and this previously marked location was visible during vocal onset analysis. First, the portion of the waveform that contained the vowel onset was enlarged. Next, the researcher zoomed in on the area of coarticulation between the stop consonant and the vowel, that is, where stop consonant noise ended and periodicity in the waveform began (Figure 5).

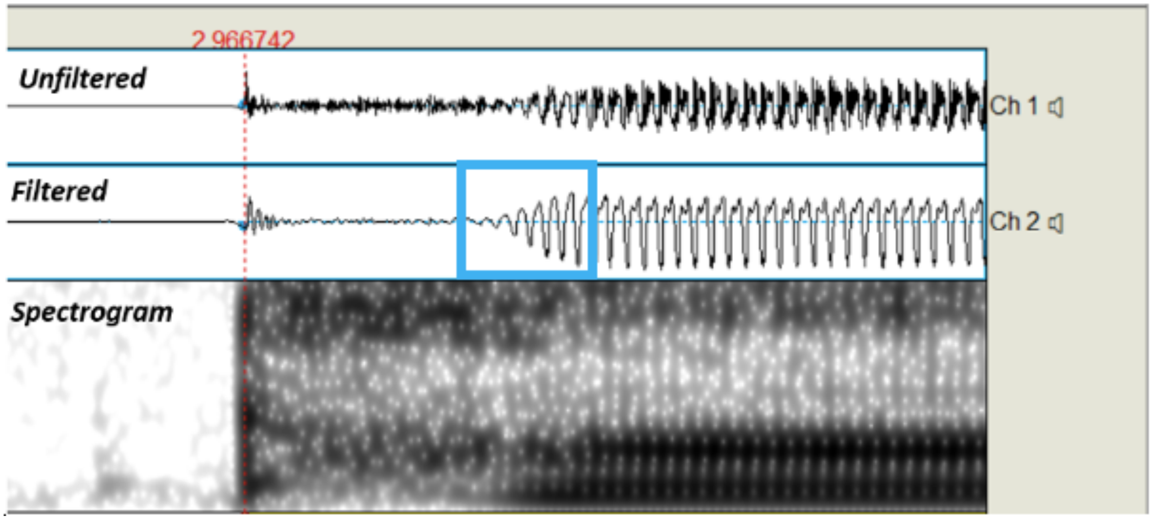

Figure 5. Broad vocal onset identification seen in blue box on the filtered waveform. Unfiltered waveform used for visual and auditory cue and spectrogram used for visual cue.

Then, the waveform was examined from right to left, moving from the midline of the vowel to the coarticulation boundary until it was visually unclear what to define as a true vocalic cycle (Figure 6). The waveform cycles were examined for periodicity and consistency in shape to the previous cycle. Once an undefinable cycle was encountered, a cursor was placed at that location on the filtered signal and on the unfiltered signal. The filtered waveform can appear more periodic due to the frequencies it filters out, so the comparison to the unfiltered waveform helps to provide further evidence of where a true

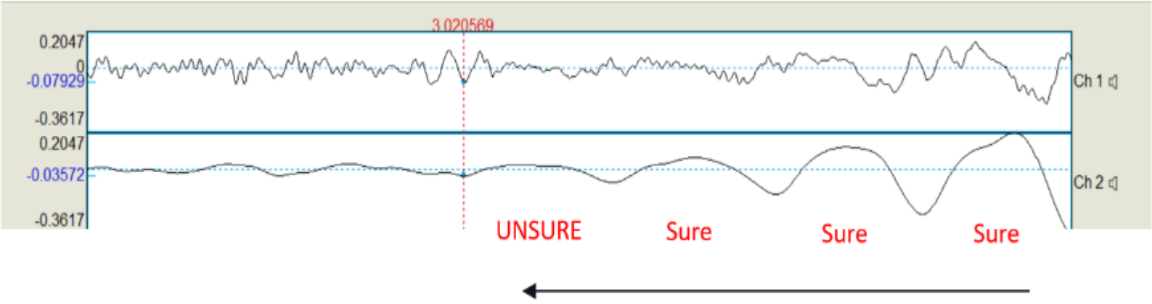

Figure 6. Undefinable cycle at coarticulation boundary. Unfiltered signal above and filtered signal below.

cycle is. When looking at the unfiltered signal, the cycles the researcher was confident with from the filtered signal were confirmed, and changes in the selected cursor location were done as needed.

Lastly, the spectrogram was examined for a voicing bar, a dark band of energy located at the individual's corresponding f_o (Figure 7). Shown in Figure 7, the voicing bar appears during voiced segments of speech with a faint transition during vocal onset. This faint transition does not appear for everyone, and the windowing on the spectrogram is different than on the waveform, so this was mainly used as a failsafe after identification in the filtered and unfiltered waveform. After checking all three displays as needed, the vowel cycle determined to be the onset of voicing was marked in either the peak or trough. After the vocal onset was marked in the TextGrid, the researcher wrote the word in between the two markers. Finally, the onset and offset of the entire sentence were marked to determine the total sentence length, denoting any incongruencies in the speech sample (e.g., the participant mispronounced the target word). All marked boundaries and notes were exported for analysis to excel and JMP (SAS Institute Inc., 2021).

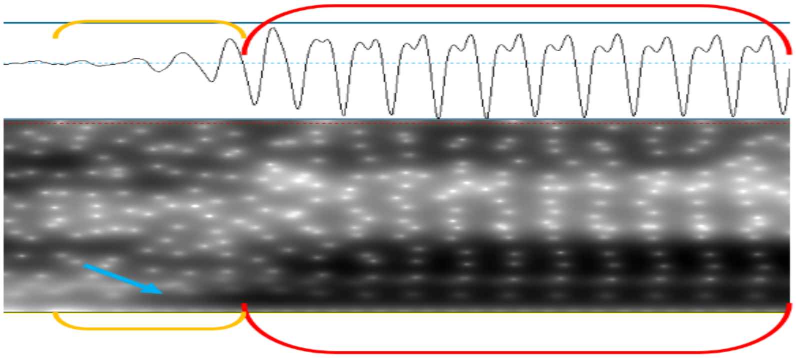

Figure 7. Voicing bar seen in red bracket and transition in yellow bracket with the onset of the transition indicated via the clue arrow. Filtered signal above, spectrogram below.

3. RESULTS

Prior to data analysis, the following data cleaning tasks were completed: 1) VOTs greater than 200 ms were removed (n = 2 instances), 2) an instance perceived as vocal play with subsequent audible elongation was removed (n = 1 instance), and 3) participants who had less than 4 usable data points were removed (n = 4 participants). Additional analyses were conducted to determine whether factors of noninterest (pitch, sentence length, and number of instances) were significantly different in the two groups (control, VFN). There were no significant group differences in pitch ($t(30.5) = 0.37$, $p = 0.71$), sentence length ($t(33.7) = 0.84$, $p = 0.41$), or number of instances ($t(27.1) = -1.76$, $p = 0.09$). Therefore, the remaining analyses are unlikely to be influenced by factors of pitch, sentence length, or the number of instances. Furthermore, high measures of intrarater ($r = .99$) and interrater ($r = .96$) suggest that the manual VOT analysis was a reliable methodology. Outlining specific methods for conducting VOT analysis proved to be successful in providing reliable data across the same and different raters.

Cepstral Peak Prominence

Independent t-tests were calculated to compare CPP values between the VFN and the control groups. The VFN group (M=10.57, SD=2.02) had significantly lower CPP values ($t(28.8)= 4.58$, $p < 0.001$) as compared to the control group (M=13.44, SD=1.65). Within the control group, there was no significant difference in CPP between male and female participants ($t(14.2) = 0.78$, $p = 0.45$). There was a significant difference in CPP between male and female participants in the VFN group ($t(11.25)= 2.18$, $p = 0.05$), with lower CPP values found in male participants (M=9.53, SD=0.96) compared to female participants (M=11.38, SD=2.30, Figure 8).

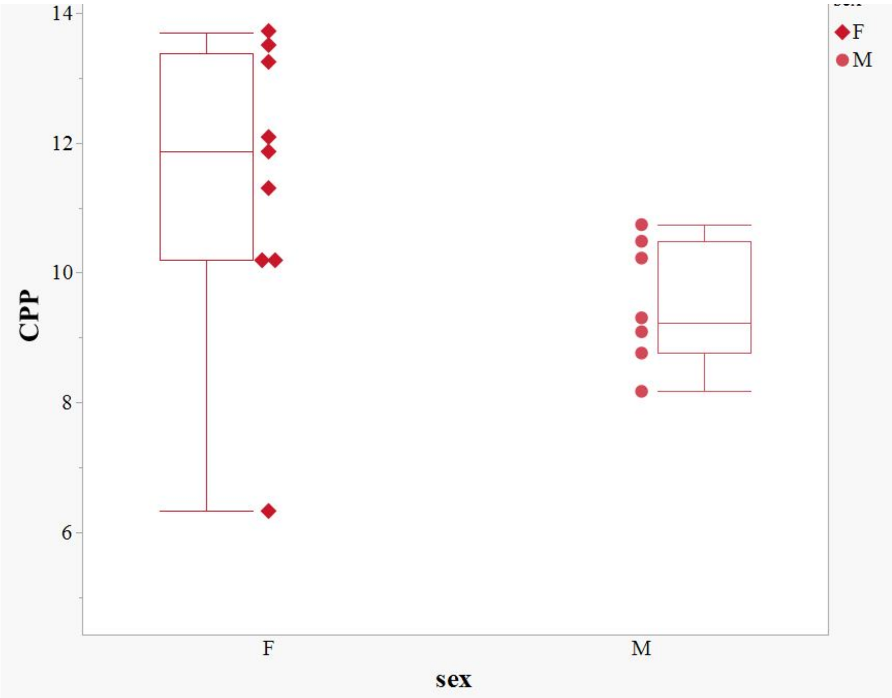

Figure 8. Distribution of CPP values within the vocal fold nodule group for female (red diamonds) and male (red circles).

Voice Onset Time Average and Variability

VOT variability, measured as the standard deviation of VOT productions, was not normally distributed and therefore was logarithmically transformed before subsequent data analysis. Two linear regressions examined whether average VOT or VOT variability were predicted by CPP, group, age, group x CPP, group x age, and age x CPP. Neither regression equation was significant (See Table 2); however, examination of the individual predictors for VOT variability suggested that additional examination of the potential relationship between CPP and VOT variability may be warranted. A significant negative correlation was found (*r* =0.35, *p* = 0.04) between CPP and VOT variability across all subjects. Additional correlations were conducted within each group. There was a significant negative correlation (*r* =-0.60, *p* = 0.01) between CPP and VOT variability within the VFN group, with a decrease in CPP (indicating a more dysphonic voice) associated with an increase in variability. There was no significant correlation found (*r* =0.11, *p* = 0.63) between CPP and VOT variability for the control group (Figure 9).

Table 2.

Linear Regression of Average VOT and VOT variability with their predictors

Predictors	Average VOT				VOT Variability			
	$R^2 = 0.25$, $p = 0.19$				$R^2 = 0.27$, $p = 0.13$			
	Estimate	Std error	t	*p*	Estimate	Std error	t	*p*
CPP	0.003	0.01	0.28	0.78	0.04	0.02	2.41	0.02*
Group	0.02	0.03	0.64	0.53	0.04	0.04	1.01	0.32
Age	0.002	0.01	0.21	0.84	0.02	0.02	1.00	0.33
Group x CPP	0.02	0.01	1.72	0.1	0.02	0.02	1.37	0.18
Group x Age	0.006	0.01	0.40	0.7	0.02	0.02	0.9	0.38

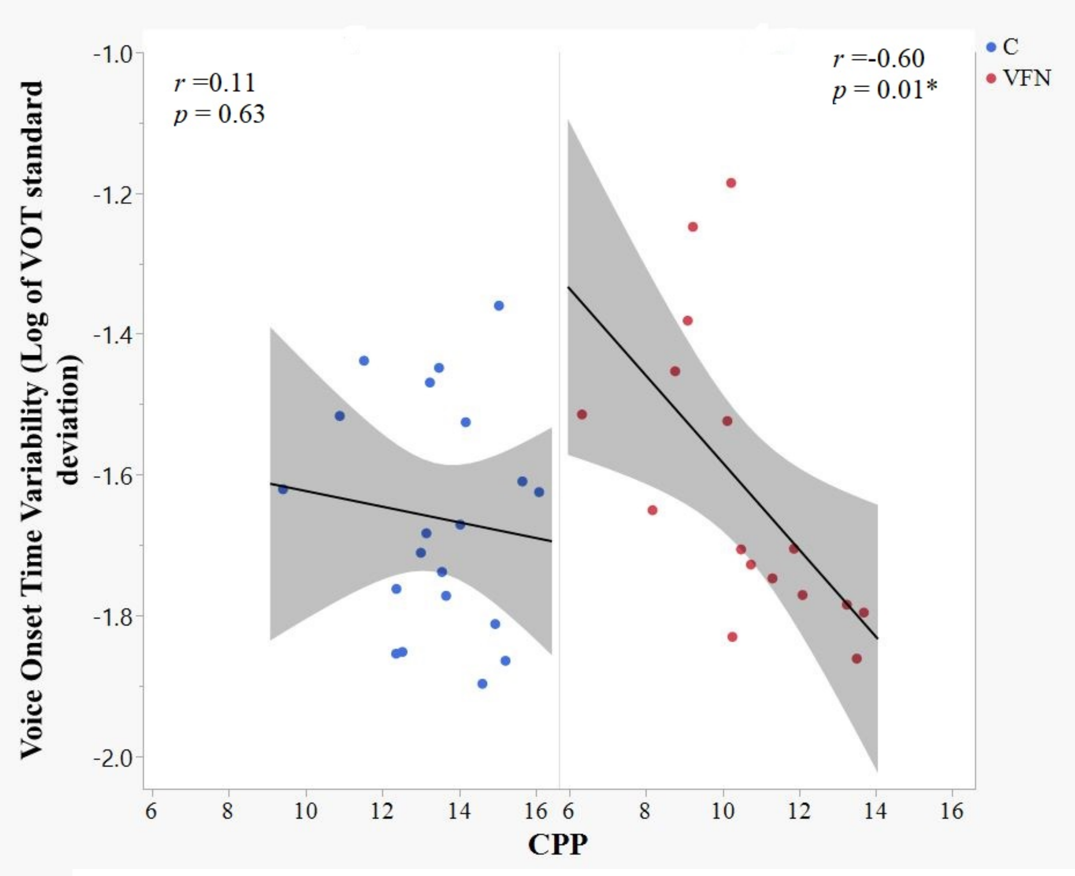

Figure 9. Group difference of VOT variability as a function of CPP for the control group (blue circles) and VFN group (red circles) with the slope and the *p*-value marked with the asterisk in the VFN group showing significance.

4. DISCUSSION

This study aimed to examine the effects of VFN on VOT average and variability measures derived from voiceless stop consonants in the pediatric population. Voiceless consonants were selected as they are a speech segment that requires control of vocal movement in order to move from a voiceless consonant (vocal folds abducted) to a vowel production (vocal folds adducted). Findings from this study indicated that increased dysphonia (measured by the acoustic measure CPP) was related to increased VOT variability across all subjects. A stronger relationship between VOT variability and CPP was found in a subsequent examination of only participants with VFN. Further discussion of the results and their implications can be found in the discussion below.

Cepstral Peak Prominence

CPP was significantly different between the control and VFN groups, with lower CPP values (corresponding to a more dysphonic voice) found in children with VFN. These results support prior research examining the reliability of CPP values in indicating the presence of dysphonia within the pediatric population (Aydinli et al., 2019). Aydinli and colleagues (2019) found that CPP was a reliable method for differentiating between healthy and dysphonic voices in children, with lower CPP values correlating to more dysphonic voices (Aydinli et al., 2019). CPP is now a recommended tool for clinical use in examining dysphonic voices for patients of all ages (Patel et al., 2018). Additionally, some articles outline the normative trajectory of CPP as a function of vocal development in children (Infusino et al., 2015). Clinicians can utilize CPP to measure baseline dysphonia during an evaluation or to track patient progress and treatment efficacy throughout therapy. Infusino and colleagues (2015) created a normative database of pediatric voices, which provides clinicians with distinct values to compare CPP values calculated during clinical care. However, it is important to note that this normative database used the Analysis of Dysphonia in Speech and Voice (ADSV) program to calculate CPP (Infusino et al., 2015). In contrast, the current work used the Praat program (Boersma & Weenink, 2022) to calculate CPP. CPP values in adults can be reliably calculated using the ADSV program or Praat software; however, these individual programs use slightly different algorithms to calculate CPP, resulting in different CPP values (Murton et al., 2020). There has only been one publication regarding normative CPP values computed with Praat in the pediatric population, but this study only analyzed CPP on a vowel level (Spazzapan et al., 2022). Therefore, future research is needed to determine a normative database for pediatric CPP values during CAPE-V sentence productions using Praat, a free and easily accessible computer program.

Statistical analysis within the VFN group revealed a significant difference in CPP values by sex, with male children having lower CPP values than female children. Male children are two times more likely to be diagnosed with VFN than female children (Angelillo et al., 2008; Shah et al., 2005). During initial data collection (Heller Murray et al., 2020), inclusionary criteria for VFN included an overall dysphonia severity score of greater than or equal to 25 on the CAPE-V (Kempster et al., 2009), which is consistent with a moderately dysphonic voice. Taking into account that these participants were already screened as having at least a moderately dysphonic voice, below we propose a few possible explanations for the sex differences seen in our current study.

First, male and female children have structural differences that occur in the prepuberty and puberty phases (Whiteside & Marshall, 2001; Yu et al., 2015). During the onset of puberty, males' f_o decline due to rapid laryngeal changes (Markova et al., 2016). We postulate that males may respond to this rapid laryngeal change and accompanying vocal instability by displaying vocal misuse patterns. They may attempt to use a pitch that is too high or low for their current anatomy while simultaneously implementing strenuous vocal movements to attempt to phonate in the presence of VFN. Another possible explanation for why males in this study had lower CPP values than females is societal expectations for male voices during puberty. It may be more socially acceptable for males to have a strained voice than for females, as people may anticipate significant perceptual voice changes during puberty in male children. This perceptive difference in societal norms may lead to males not being brought to the clinic until their voice is more severely dysphonic. Therefore, early detection of dysphonia in males may go unnoticed until their dysphonia becomes more severe. This may be a cause of sex differences we see in this study, with male children not being brought into the clinic until their dysphonia is severe. A caveat to this discussion is the consideration of gender identity as an influence. This study utilized prior data, where gender identity was not an inclusion consideration in participant selection. Future research should consider gender identity and the influence of hormones and/or hormone blockers, which may have had an effect on the variables discussed in this study.

Average Voice Onset Time

VOT measures of voiceless consonants provide a snapshot of the vocal and articulatory mechanisms during coordination. During the production of a voiceless consonant preceding a vowel, the vocal folds must move from a position of abduction (open) to adduction (closed) in order to build up subglottic pressure to allow for phonation. This production is more complex than a voiced consonant, where the vocal folds remain adducted throughout the transition from the consonant to the vowel. In individuals with VFNs, this coordination would be further complicated by changes in anatomical structure. VFN thicken the vocal folds, increasing their mass. We hypothesized this increased vocal fold mass could impact vocal control during the voiceless consonant production. However, this study did not find any group differences in VOT average. One possible explanation for this is that the average participant age was 8.8 years old, ranging from 6.0 to 12.5 years. As VOT averages become adult-like around six years old (Kent, 1976; Whiteside & Marshall, 2001), it is possible that children in this study already had mature VOT productions that are not impacted by structural vocal changes.

Another potential explanation is that the structural changes the VFN nodules made were not large enough to affect VOT averages. Although this study utilized a board-certified laryngologist's diagnosis of vocal fold nodules, the actual characteristics of the nodules were not taken into account. VFN are defined as a bilateral thickening of the epithelium (Kiliç et al., 2004; Sonbay Yılma et al., 2021; Tezcaner et al., 2009). Thus, this general diagnosis does not include additional information on the size or characteristics of the VFN that could have affected the child's ability to phonate. As the sizes of the VFN are unknown, it is possible that the majority of the participants had small VFN that did not impact vocal control as much as larger VFN might have.

Voice Onset Time Variability

VOT variability is increased in children, decreasing until around 8-11 years of age when it reaches adult-like levels (Gilbert & Purves, 1977; Whiteside & Marshall, 2001; Yu et al., 2015; Zlatin & Koenigsknecht, 1976). As participants in this study averaged 8.8 years old, we anticipated that these participants would display a range of VOT variabilities. VOT variability was not significantly predicted by CPP, group, age, group x CPP, group x age, age x CPP. Although the overall model was not significant, results suggested that further examination into the relationship between VOT variability and CPP was warranted (See Table 2). A significant negative correlation was found between CPP and VOT variability, with increased dysphonia associated with more variability in VOT productions. As the control and VFN groups had significant differences in CPP, further analysis was done to examine the relationship between CPP and VOT variability within each group. There was no relationship between CPP and VOT variability within the control group. This was an expected finding to the small range of CPP values, likely due to the participant's healthy vocal qualities. A significant negative correlation was found within the VFN group, with more dysphonic voices (lower CPP) associated with increased VOT variability. Therefore, the presence of VFN may impact not only the periodicity of the vocal folds (measured via CPP) but also the ability to control the vocal folds during speech (resulting in increased VOT variability).

The relationship between dysphonia and VOT variability has been previously noted in adults with voice disorders. McKenna and colleagues (2020) found that adult women with voice disorders showed an increase in VOT variability as dysphonia severity increased, during voiceless production on the CAPE-V(as rated auditorily-perceptually on the CAPE-V). However, unlike the current study, McKenna and colleagues (2020) found a group difference in VOT variability, with individuals with voice disorders showing increased VOT variability compared to sex- and age-matched peers (McKenna et al., 2020). McKenna and colleagues (2020) suggest their findings support the hypothesis previously proposed by others (e.g., Stepp et al., 2017) that adults with voice disorders have disordered vocal motor control. Previous researchers have proposed that the auditory-motor deficits in adults with voice disorders may be primarily related to auditory discrimination deficits. Abur and colleagues found that individuals with vocal hyperfunction had decreased auditory discrimination abilities across both external (pure tones) and internal stimuli (their voices) (Abur et al., 2021). Accurate auditory feedback is necessary for the detection of vocal errors, a key element of vocal motor control. Control of the vocal motor system includes feedback and feedforward controls (Behroozmand & Sangtian, 2018; Guenther, 2016). The feedback mechanism conducts online monitoring of vocal output, with error correction for vocal errors initiated as needed. The feedforward mechanism stores vocal motor patterns, allowing automaticity during productions. These mechanisms can be assessed by vocal perturbation paradigms in which participants phonate and hear their own voice shifted in near real-time (e.g., adjust pitch or loudness) (Behroozmand & Sangtian, 2018; Burnett et al., 1998; Keough

& Jones, 2009). In response to this perturbation, individuals who are vocally healthy will implement a compensatory strategy opposing the perturbation (e.g., produce a decrease in pitch when a shift upward is heard) to continue phonation within the perceived auditory target range. The mechanism by which this works consists of detecting the error (feedback controls) and subsequently adjusting the production. However, if consistent errors are detected, the automaticity of the system must adapt. Therefore, the feedforward mechanism must adjust to ensure future productions are within the desired auditory range (Behroozmand & Sangtian, 2018; Burnett et al., 1998; Keough & Jones, 2009). These prior studies propose a breakdown in adults with voice disorders in the auditory discrimination mechanism, causing an inability to effectively adapt the feedforward system and increase production variability.

McKenna and colleagues note that these auditory discrimination deficits may result in larger auditory target ranges, possibly contributing to the increased VOT variability (McKenna et al., 2020). However, prior research suggests that children with VFN have comparable vocal pitch discrimination abilities to age and sex-matched peers with typical voices (Heller Murray et al., 2019). As auditory-discrimination deficits do not appear to be present in children with VFN, this may explain why the current study did not find a group difference in VOT variability. Heller Murray and colleagues (2019) did note that younger children had poorer pitch discrimination abilities than older children. Moreover, older children continued to have poorer pitch discrimination abilities as compared to adults. Further exploration into auditory discrimination deficits and vocal variability within the dysphonic pediatric population across different ages may be warranted.

Furthermore, additional work is also needed to examine if children with VFN perceive their vocal differences as ‘errors’ that require correction. Most adults who are their own primary caretakers will seek a professional evaluation if a change in their voice occurs. However, children are not their main caretakers and rely heavily on an external source (e.g., adults) to monitor changes such as their behavior, health, and safety. Children referred to a professional for dysphonia are typically brought in because someone external, like a caregiver, has noticed a change in their vocal quality. Although research has shown that children are aware of their voice in general (Connor et al., 2008), further work is needed to determine children’s abilities to detect smaller changes in vocal quality. Additional work is also needed to examine how dysphonic children perceive and detect differences in their own voices.

An alternative theory to the disordered vocal control deficit suggested in adults is that increased VOT variability in children may be part of a compensatory response to the presence of VFN. The presence of VFN can increase breathiness, as the VFN becomes the initial point of contact during phonation, leading to anterior and posterior escape of air. Children may also find it difficult to build up adequate subglottic pressure and may need to implement phonotraumatic behaviors (e.g., strain, increased muscle tension) in order to phonate. This vocal behavior leads to further vocal misuse, further exacerbating the already present VFN. These maladaptive compensatory strategies children might employ may be more severe in children with increased dysphonia. Thus, this increased use of these phonotraumatic strategies (e.g., strain, increased muscle tension) may also result in increased variability of vocal control for speech, such as what is required in VOT productions.

Another possibility is that children with increased dysphonia are using phonotraumatic behaviors to phonate, resulting in increased aphonic breaks. The presence of aphonic breaks may decrease CPP, as CPP measures the periodicity of the cepstrum. VOT variability may also be affected by aphonic breaks, as these patterns result in general instability and unreliability of vocal productions. Thus, some VOT productions may be affected by a change in periodicity while others are not. If a child experiences an aphonic break, they may try to compensate on their next production, possibly overcompensating and, in turn producing a less strained, more breathy voice until this compensatory strategy proves unsuccessful. If a child is constantly changing how they are phonating, this can increase their variability and dysphonia if they utilize phonotraumatic behaviors. Dysphonic adults may also be employing this habitual use of maladaptive vocal behaviors, but this can prove to be much more harmful in the pediatric population. The use of maladaptive phonatory behaviors in the pediatric population may lead to more severe vocal fold damage, or a larger perceptual change in voice, given the differences in anatomy. Future research should examine if maladaptive phonatory behaviors affect pediatric vocal folds differently from adults.

Limitations

The current study presented with several limitations that may have impacted the outcome. First, the study was limited in several ways by power. Although the study had

40 participants, when completing statistical analysis, certain measures (such as examination of potential group differences within sex) could not be conducted due to the lack of power. This study also lacked a representative sample regarding age. Children can develop VFN as early as a few months old (Shah et al. 2005), however the mean age of our sample was 8.8 years old. Additionally, some of the children in this study were past the normative age for some developmental stages of VOT (e.g., VOT average). Further limitations of this study include the averaging of VOT measurements across phonemes. Ideally, VOT measurements should be analyzed according to each phoneme as normative data for VOT measurements differs by phoneme. This study exclusively analyzed voiceless phonemes (/p/ & /k/) which likely would not have contributed much significant difference due to their similarities in normative data. It is possible there may have been sex differences between VOT values across phonemes when considering sex. Prior research has indicated that males within the sexual dipmorphism stage of development (8-11 years old), had increased VOT values than females (Whiteside & Marshall, 2001; Yu et al., 2015)

Future Directions

In addition to the already stated future research suggested in the discussion, additional future studies are justified in the area of pediatric voice, specifically in regard to VFN and its impact on VOT average, VOT variability, CPP, and auditory-motor deficits. Specific questions that would add to the growing body of research within this population include the effects of voice therapy on VOT average and VOT variability in children with and without VFN. Studies can include characteristics about each participant's VFN at baseline (via laryngoscopy imaging), monitoring change in severity/size after treatment. The relationship between individual VFN characteristics can be compared to overall dysphonia pre and post-treatment with VOT values measured at both points in treatment to track change. Another possible research question may examine the longitudinal change of VOT variability and CPP values across children with VFN. If we were to study the longitudinal change in children with VFN, would we see changes in VOT variability and CPP values. Considerations may include if a child seeks out treatment for VFN, but the direct nature of the study would not be treatment efficacy.

Conclusion

The current study was designed to study differences in average VOT and VOT variability in children with and without VFN. The purpose of this study was to provide further research and evidence for the need for tailored treatment for the pediatric population, as much of the current clinical practice for children is based on adult norms. Results from the study found that there was no group difference in VOT measures. However, it was found that VOT variability was increased in children with increased dysphonia severities within the VFN group. This current study will add to the limited body of research conducted regarding voice within the pediatric population.

5. REFERENCES CITED

Abur, D., Subaciute, A., Kapsner-Smith, M., Segina, R. K., Tracy, L. F., Noordzij, J. P., & Stepp, C. E. (2021). Impaired auditory discrimination and auditory-motor integration in hyperfunctional voice disorders. *Scientific Reports, 11*(1), 1-11.

Angelillo, N., Di Costanzo, B., Angelillo, M., Costa, G., Barillari, M. R., & Barillari, U. (2008). Epidemiological study on vocal disorders in paediatric age. *Journal of Preventative Medicine and Hygiene*, *49(1)*, 1-5.

Aydinli, F. E., Özcebe, E., & İncebay, Ö. (2019). Use of cepstral analysis for differentiating dysphonic from normal voices in children. *International Journal of Pediatric Otorhinolaryngology*, *116*, 107-113.

Behroozmand, R., & Sangtian, S. (2018). Neural bases of sensorimotor adaptation in the vocal motor system. *Experimental Brain Research*, *236*(7), 1881–1895.

Boersma, P., & Weenink D. (2022) *Praat: doing phonetics by computer* (Version 6.2.10) [Computer software]. http://www.praat.org/

Brown Jr, W. S., Morris J. R., & Rudolf, R. (1993) Comparative methods for measurement of VOT. *Journal of Phonetics, 21*(3), 329-336.

Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, *103*(6), 3153–3161.

Connor, N. P., Cohen, S. B., Theis, S. M., Thibeault, S. L., Heatley, D. G., & Bless, D. M. (2008). Attitudes of children with dysphonia. *Journal of Voice*, *22*(2), 197-209.

Gilbert, J. H., & Purves, B. A. (1977). Temporal constraints on consonant clusters in child speech production. *Journal of Child Language*, *4*(3), 417-432.

Guenther, F. H., & Hickok, G. (2016). Neural models of motor speech control. In G. Hickock & S. Small (Eds), *Neurobiology of Language* (pp. 725-740). Academic Press.

Hartnick, C. J., Rehbar, R., & Prasad, V. (2005). Development and maturation of the pediatric human vocal fold lamina propria. *The Laryngoscope*, *115*(1), 4-15.

Hasek, C., Singh, S., & Murray, T. (1980). Acoustic attributes of children's voices. *Journal of the Acoustical Society of America*, *68*, 1262–1265.

Heller Murray, E. S., & Chao, A. (2021). The relationships among vocal variability, vocal-articulatory coordination, and dysphonia in children. *Journal of Voice.* Available Online.

Heller Murray, E. S., Hseu, A. F., Nuss, R. C., Harvey Woodnorth, G., & Stepp, C. E. (2019). Vocal pitch discrimination in children with and without vocal fold nodules. *Applied Sciences*, *9*(15), 3042.

Heller Murray, E. S., Segina, R. K., Woodnorth, G. H., & Stepp, C. E. (2020). Relative fundamental frequency in children with and without vocal fold nodules. *Journal of Speech, Language, and Hearing Research*, *63*(2), 361-371.

Heman-Ackah, Y. D., Heuer, R. J., Michael, D. D., Baroody, M. M., Ostrowski, R., Hillenbrand, J., Horman, M., & Sataloff, R. T. (2003). Cepstral peak prominence: a more reliable measure of dysphonia. *Annals of Otology, Rhinology & Laryngology*, *112*(4), 324-333.

Hitchcock, E. R., & Koenig, L. L. (2013). The effects of data reduction in determining the schedule of voicing acquisition in young children. *Journal of Speech, Language, and Hearing Research, 56*(2), 441-457.

Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech, Language, and Hearing Research*, 37(4), 769-778.

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech, Language, and Hearing Research*, *39*(2), 311-321.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., & Vaughan, C. (1989). Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech, Language, and Hearing Research*, *32*(2), 373-392.

Infusino, S. A., Diercks, G. R., Rogers, D. J., Garcia, J., Ojha, S., Maurer, R., Bunting, G., & Hartnick, C. J. (2015). Establishment of a normative cepstral pediatric acoustic database. *JAMA Otolaryngology–Head & Neck Surgery*, *141*(4), 358-363.

JMP® Pro (version 16.1.0). SAS Institute Inc., Cary, NC, 1989-2021.

Karlsson, F., Zetterholm, E., & Sullivan, K. P. (2004). Development of a gender difference in voice onset time. *Proceedings of the 10th Australian International Conference on Speech Science & Technology,* 316 - 321.

Kempster, G. B., Gerratt, B. R., Abbott, K. V., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *American Journal of Speech-Language Pathology*, *18*(2), 124-132.

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech, Language, and Hearing Research*, *19*(3), 421-447.

Keough, D., & Jones, J. A. (2009). The sensitivity of auditory–motor representations to subtle changes in auditory feedback while singing. *The Journal of the Acoustical Society of America*, *126*(2), 837–846.

Kiliç, M. A., Okur, E., Yildirim, I., & Güzelsoy, S. (2004). The prevalence of vocal fold nodules in school age children. *International Journal of Pediatric Otorhinolaryngology*, *68*(4), 409-412.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, *18*(4), 686-706.

Koenig, L. (2000). Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *Journal of Speech, Language, and Hearing Research*, *43*(5), 1211-1228

Leeper Jr, H. A., Leonard, J. E., & Iverson, R. L. (1980). Otorhinolaryngologic screening of children with vocal quality disturbances. *International Journal of Pediatric Otorhinolaryngology*, *2*(2), 123-131.

Lien, Y. A. S., & Stepp, C. E. (2014). Comparison of voice relative fundamental frequency estimates derived from an accelerometer signal and low-pass filtered and unprocessed microphone signals. *The Journal of the Acoustical Society of America*, *135*(5), 2977-2985.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384-422.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, *7*(1), 41-74.

Markova, D., Richer, L., Pangelinan, M., Schwartz, D. H., Leonard, G., Perron, M., Pike, B. G., Veillette, S., Chakravarty, M. M., Pausova, Z., & Paus, T. (2016). Age-and sex-related variations in vocal-tract morphology and voice acoustics during adolescence. *Hormones and Behavior, 81*, 84-96.

McCrea, C. R., & Morris, R. J. (2005). The effects of fundamental frequency level on voice onset time in normal adult male speakers. *Journal of Speech, Language, and Hearing Research*, *48*(5), 1013-1024

McKenna, V. S., Hylkema, J. A., Tardif, M. C., & Stepp, C. E. (2020). Voice onset time in individuals with hyperfunctional voice disorders: Evidence for disordered vocal motor control. *Journal of Speech, Language, and Hearing Research*, *63*(2), 405-420.

Murton, O., Hillman, R., & Mehta, D. (2020). Cepstral peak prominence values for clinical voice evaluation. *American Journal of Speech-Language Pathology*, *29*(3), 1596-1607.

Nuss R. C., Ward J., Huang L., Volk M., & Woodnorth G. H. (2010). Correlation of vocal fold nodule size in children and perceptual assessment of voice quality. *Annals of Otology, Rhinology & Laryngology*, *119*(10), 651–655.

Ongkasuwan, J., & Friedman, E. M. (2013). Is voice therapy effective in the management of vocal fold nodules in children? *The Laryngoscope*, *123*(12), 2930-2931.

Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec J. G., & Hillman, R. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology, 27*(3), 887-905.

Petrosino, L., Colcord, R. D., Kurcz, K. B., Yonder, R. J. (1993). Voice onset time of velar stop productions in aged speakers. *Perceptual and Motor Skills*, *76*(1), 83–88.

Seikel, J. A., Drumright, D. G., & Hudock, D.R. (2021). *Anatomy and physiology for speech, language, and hearing* (6th ed.). Plural Publishing Inc.

Shah, R. K., Woodnorth, G. H., Glynn, A., & Nuss, R. C. (2005). Pediatric vocal nodules: correlation with perceptual voice analysis. *International Journal of Pediatric Otorhinolaryngology*, *69*(7), 903-909.

Sonbay Yılmaz, N. D., Afyoncu, C., Ensari, N., Yıldız, M., & Gür, Ö. E. (2021). The effect of the mother's participation in therapy on children with vocal fold nodules. *Annals of Otology, Rhinology & Laryngology*, *130*(11), 1263-1267.

Spazzapan, E. A., de Castro Marino, V. C., & Fabbron, E. M. G. (2022). Smoothed cepstral peak analysis of brazilian children and adolescents speakers. *Journal of Voice*. Available online

Stepp, C. E., Lester-Smith, R. A., Abur, D., Daliri, A., Pieter Noordzij, J., & Lupiani, A. A. (2017). Evidence for auditory-motor impairment in individuals with hyperfunctional voice disorders. *Journal of Speech, Language, and Hearing Research*, *60*(6), 1545-1550.

Tavares, E. L. M., Brasolotto, A., Santana, M. F., Padovan, C. A., & Martins, R. H. G. (2011). Epidemiological study of dysphonia in 4-12 year-old children. *Brazilian Journal of Otorhinolaryngology*, *77*(6), 736-746.

Tezcaner, C. Z., Ozgursoy, S. K., Sati, I., & Dursun, G. (2009). Changes after voice therapy in objective and subjective voice measurements of pediatric patients with vocal nodules. *European Archives of Oto-Rhino-Laryngology*, *266*(12), 1923-1927.

Watson, B., & Alfonso, P. (1982). A comparison of LRT and VOT between stutterers and nonstutterers. *Journal of Fluency Disorders*, *7*(2), 219-241

Whiteside, S. P., & Marshall, J. (2001). Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica*, *58*(3), 196-210.

Whiteside, S. P., Henry, L., & Dobbin, R. (2004). Sex differences in voice onset time: A developmental study of phonetic context effects in British English. *The Journal of the Acoustical Society of America*, *116*(2), 1179-1183.

Yu, V. Y., De Nil, L. F., & Pang, E. W. (2015). Effects of age, sex and syllable number on voice onset time: evidence from children's voiceless aspirated stops. *Language and Speech*, *58*(2), 152-167.

Zeitels, S. M., & Healy, G. B. (2003). Laryngology and phonosurgery. *New England Journal of Medicine*, *349*(9), 882-892.

Zlatin, M. A., & Koenigsknecht, R. A. (1976). Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research*, *19*(1), 93-111.