NIH-PA Author Manuscript

Phonetically Governed Voicing Onset and Offset in Preschool Children Who Stutter

Richard M. Arenas, MA,
University of Iowa, Dept. of Communication Sciences and Disorders, 250 Hawkins Drive, Iowa City, IA 52242

Patricia M. Zebrowski, Ph.D., and
University of Iowa, Dept. of Communication Sciences and Disorders, 250 Hawkins Drive, Iowa City, IA 52242

Jerald B. Moon, Ph.D.
University of Iowa, Dept. of Communication Sciences and Disorders, 250 Hawkins Drive, Iowa City, IA 52242

Patricia M. Zebrowski: tricia-zebrowski@uiowa.edu; Jerald B. Moon: jerald-moon@uiowa.edu

Abstract

Phonetically governed changes in the fundamental frequency (F_0) of vowels that immediately precede and follow voiceless stop plosives have been found to follow consistent patterns in adults and children as young as four years of age. In the present study, F_0 onset and offset patterns in 14 children who stutter (CWS) and 14 children who do not stutter (CWNS) were investigated to evaluate differences in speech production. Participants produced utterances containing two VCV sequences. F_0 patterns in the last ten vocal cycles in the preceding vowel (voicing offset) and the first ten vocal cycles in the subsequent vowel (voicing onset) were analyzed. A repeated measures ANOVA revealed no group differences between the CWS and CWNS in either voicing onset or offset gestures. Both groups showed patterns of F_0 onset and offset that were consistent with the mature patterns seen in children and adults in previous studies. These findings suggest that in both CWS and CWNS, a mature pattern of voicing onset and offset is present by age 3;6. This study suggests that there is no difference between CWS and CWNS in the coordination of respiratory and laryngeal systems during voicing onset or offset.

Keywords

stuttering; voicing; respiration; coordination

1. Introduction

Developmental stuttering is a disorder that begins in early childhood. As such, the extent to which specific speech production behaviors are associated with either the onset or moment of stuttering needs to be determined from speech samples that are collected relatively close

Corresponding Author: Richard Arenas, M.A., University of Iowa, Dept. of Communication Sciences and Disorders, 250 Hawkins Drive, Iowa City, IA 52242, Phone: 319-936-1287, Fax: 319-936-1287, rick-arenas@uiowa.edu.

to the onset of the stuttering problem. The main argument for this criterion is that compared to adults who stutter (AWS), the fluent and disfluent speech of children who stutter (CWS) is less likely to be influenced by years of learned reactive and compensatory behaviors (e.g. Caruso, Conture, & Colton, 1988; Conture, Colton, & Gleason, 1988; Zebrowski, Conture, & Cudahy, 1985).

Beginning in the late 1970's, a relatively large body of literature describing the acoustic characteristics of the speech production behaviors of CWS has emerged. To a large extent, these studies have focused on acoustic measures that reflect laryngeal coordination through measures of voice initiation or laryngeal reaction time (e.g. Cross & Luper, 1979; McKnight & Cullinan, 1987; Till, Reich, Dickey, & Sieber, 1983), as well as laryngeal-supralaryngeal coordination through measures of F_0 and temporal and spatial measures that include, among other things, F_0 jitter and shimmer, voice onset time (VOT), vowel duration, and formant transition duration and rate(e.g. Adams, 1987; Chang, Ohde, & Conture, 2002; Hall & Yairi, 1992; Subramanian, Yairi & Amir, 2003; Yaruss & Conture, 1993; Zebrowski, Conture & Cudahy, 1985). Findings from this work have been equivocal, but in general the data suggest that there is a subgroup (or multiple subgroups) of children who stutter who exhibit difficulty with the rapid, precise laryngeal and laryngeal-supralaryngeal coordination required for fluent speech production.

There are several reasons for these inconsistent results, with the most likely related to the fact that the measures used were obtained from speech samples that by their very nature are subject to both linguistic influence (e.g. utterance length and syntactic complexity, prosodic variation) (Stepp, Hillman & Heaton, 2010) and the idiosyncratic dialectical pattern and articulation rate of the speaker. Taken together, these variables contribute to the frequent observation of large within- and between-group variability seen in the acoustic representations of speech production for children, and to a lesser extent, adults.

Recently, several researchers have proposed that measuring short-term phonatory changes associated with the offset (devoicing) and onset (voicing) of phonation in vowels surrounding a voiceless stop obstruent provides a measure of laryngeal function. The change and variability of change in F_0 (referred to as relative fundamental frequency, or RFF) in so-called phonetically governed devoicing and voicing gestures is thought to reflect the coordination of laryngeal and aerodynamic adjustments necessary for the production of voiceless obstruents or "true consonants" (i.e. stops, fricatives and affricates; Edwards & Shriberg, 1983) with vocal fold morphology (e.g. stiffness) and vocal tract anatomy. That is, changes in phonetically-governed devoicing and voicing F_0 are the direct consequence of the laryngeal and respiratory coordination required for specific articulatory gestures (e.g. lip closing and opening associated with bilabial stop consonants), and as such are not under overt or deliberate speaker control or the influence of language and related suprasegmental variation, (Baken & Orlikoff, 1988; Goberman & Blomgren, 2008; Robb & Smith, 2002; Stepp et al., 2010).

Recently, Robb and Smith (2002), Goberman and Blomgren (2008), Stepp et al., 2010, and Stepp, Merchant, Heaton & Hillman (2011) adopted a method developed by Watson (1998) and based on work by House and Fairbanks (1953), and Baken and Orlikoff (1988), to examine short-term RFF or phonetically governed voicing changes in both typical and atypical children and adults. For typical speakers, this work has shown consistent, age-related patterns of vowel F_0 change and variability immediately before and after the production of voiceless stop consonants (Peterson, 2001; Robb & Smith, 2002; Watson, 1998); specifically, there is a reduction in F_0 across vocal cycles in the vowel preceding the consonant, and a relatively high F_0 at voicing onset in the subsequent vowel followed by a decline across vocal cycles. Further, there is relatively high inter-subject F_0 variability for

both the devoicing and voicing gestures, and this variability decreases with age. In studies of atypical speakers, researchers have observed these same patterns (i.e. F_0 reduction across cycles preceding a voiceless obstruent, and a relatively high F_0 at voicing onset, followed by a steep decline) in individuals with Parkinson's disease (PD) and vocal hyperfunction, although qualitative differences between these populations and typical speakers have emerged. Specifically, individuals with PD show smaller relative fundamental frequency changes in phonetically-governed voicing onset and offset compared to typical speakers, and those who exhibit vocal hyperfunction show lower short-term RFF overall than their typical peers (Goberman& Blomgren, 2008; Stepp et al., 2010). Stepp et al., concluded that the differences between typical and atypical speakers (i.e PD and vocal hyperfunction) is likely due to increased laryngeal tension and/or reduced airflow secondary to rigidity of the repiratory system in the latter.

Based on their study of short-term RFF in normally speaking children and adults, Robb and Smith (2002) suggested that comparing F_0 change and variability associated with voicing offset and onset in children with and without speech disorders would help to uncover specific laryngeal/respiratory adjustments surrounding disrupted articulatory events (including speech disfluencies). The consistent pattern of F_0 change across short devoicing and voicing onset intervals seen for even the youngest children (four years; Robb & Smith), with associated variability decreasing with age, suggests that these measures may be particularly sensitive to subtle differences in laryngeal muscle tension and coordination that have long been suspected in CWS. As previously discussed, of special benefit is that this method offers a simple, noninvasive and nonintrusive method for observing the subtle and difficult to observe changes in laryngeal articulation (specifically, laryngeal/aerodynamic coupling necessary to facilitate consonant production) that occur in young children during normal speech development. Further, the nature of the analyzed sample should reduce the high degree of within- and between-child variability in speech production that characterizes children's speech.

With this in mind, the purpose of this study was to compare both mean F_0 change and variability (i.e. standard deviation) associated with voicing offset and onset, in V-(voiceless) C-V syllables produced by preschool children who do and do not stutter (CWNS). As discussed at the beginning of this paper, we used young children relatively close to the onset of stuttering as participants because of our assumption that any speech production anomalies that have been suspected to be associated with the onset and early development of stuttering need to be determined from speech samples that are collected relatively close to the onset of the stuttering problem.

Specifically, we hypothesized that consistent with other between-group studies of speech physiology, there would be no difference in the trajectory or patterns of RFF associated with the devoicing-voicing gestures, but that compared to CWNS, CWS would show higher within- and between-group variability, reflecting a more heterogeneous population in general, and less stable patterns of laryngeal strategies for voicing and devoicing. Further, with reference to the developmental data reported by Robb and Smith (2002), increased variability of CWS compared to same-aged CWNS may reflect immature development of phonatory coordination in this population.

2. Methods

2.1 Participants

The participants and protocols used in this study were part of a larger longitudinal study conducted by the University of Illinois. The study was reviewed and approved by Institutional Review Boards at the University of Illinois, University of Iowa, Northern

Illinois University and University of Wisconsin. Informed consent was obtained prior to participation in the longitudinal study. The data used in this study were collected during the first visits of the longitudinal study. Participants in the study were chosen from a larger sample of participants in the longitudinal study based on their classification (CWNS, CWS), age, and amount of available analyzable data. Two groups in the present study consisted of 14 children who did not stutter (CWNS), and 14 children who did stutter (CWS). The CWS group contained 9 males and 5 females, with a mean age of 42.5 months and a range from 30 to 57 months. The CWNS group had 8 males and 6 females, with a mean age of 41.9 months and a range from 30 to 57 months. The two groups of children were recruited from the general public in the Urbana-Champaign, IL, Iowa City, IA, DeKalb, IL, and Milwaukee, WI areas.

The following criteria were employed to identify children in the CWS group: (a) age five years or younger at the time of testing, (b) regarded by their parents as having a stuttering problem, (c) regarded by an investigator as having a stuttering problem, (d) stuttering severity rated by both the parent as investigator as 1 or above on an 8-point severity scale (0-1=normally fluent to 7 = very severe stuttering) and the average rating ((parent rating + investigator rating)/2) had to be at least 1.5, (e) no more than twelve months post stuttering onset. Children in the CWNS group met the following criteria: (a) age five years or younger at the time of testing, (b) regarded by parents as never having stuttered, (c) regarded by the investigators as not exhibiting stuttering, and (d) stuttering severity rated as less than or equal to 1 by both the parents and investigators. The individuals in the CWNS were further selected by age matching to individuals in the CWS group. Each age matched pair could not be greater than three months apart and the average age separation was 1.14 months. Other than stuttering in the CWS group, none of the participants were identified as exhibiting a speech, language, or hearing disorder at the time of data collection. All children demonstrated normal communication behavior as determined by their parent and the investigators.

2.2 Procedures

Each child's speech was recorded using an audio CD recorder (HHB CDR-830) as he or she produced the sentence "I see two papas" six times. The target V1- C-V2 sequences used for the analysis were the/apa/in "papa" and the /itu/in "see two". Each child was given an opportunity to practice each stimulus prior to recording. During the practice interval, the examiner used verbal and physical prompts to elicit the child's production of the stimuli. The examiners were either research assistants or graduate students that were not one the authors of this study. Physical prompts included gestures and a children's book about grandfathers. To elicit "I see two Papas," children were shown a page of the book containing two grandfathers, and instructed to repeat the phrase. During the actual recording interval, the examiner provided verbal instructions with a modeled exemplar if the child was not producing the utterance with just the physical prompt. The examiners were aware that only utterances that were spoken in the child's typical speech, were produced correctly and free of disfluencies could be used in the analysis. The goal of the examiner was to acquire six utterances that could be incorporated into the data analysis. If the participant produced an utterance that the researcher believed could not be used in the analysis the participant was asked to repeat the utterance. All productions were obtained in locations judged to be low in ambient noise.

The procedures for measuring F_0 onset and F_0 offset closely paralleled those of Watson (1998) and Robb and Smith (2002). Due to a lack of clear vocal cycle signatures, not all V-C-V sequences were able to be used in the analysis. Across both V-C-V target sequences, participants had to produce at least 6 analyzable sequences to be included in the dataset. Some utterances were excluded from analysis because ten vocal cycles before and after the

stop consonant could not be visibly identified. The CWNS group had 146 (76 /apa/ and 70 /itu/) analyzable V-C-V sequences out of 168 (86.9%), with a mean of 10.4 V-C-V sequences per participant. The CWS group had 143 (74 /apa/ and 69 /itu/) analyzable V-C-V sequences out of 168 (85.1%), with a mean of 10.2 V-C-V sequences per participant. Of the 25 V-C-V sequences that could not be analyzed, six of them were due to disfluencies in the utterance, the remaining 19 were due to a lack of consistent vocal cycle patterns.

2.3 Measures

The recorded audio signal was digitized at 44,100 Hz using CSL hardware and software (Kay CSL-4300). Within CSL, each utterance was displayed as an amplitude-by-time waveform. V-C-V segments were selected and magnified using vertical cursors. An example of a typical waveform is presented in Figure 1.

F_0 offsets and onsets were measured from the magnified waveform in the following manner:

F_0 Offset—The last ten vocal cycles of V1 preceding the consonant were identified using the negative peaks and marked on a peak-to-peak basis as vocal cycles -10 through -1. The offset of voicing was identified as the last visible periodic vocal cycle preceding the voiceless stop consonant (vocal cycle -1). The time values for each voiced period mark were then extracted from CSL and the period for each vocal cycle was calculated. The F_0 for each vocal cycle was determined as the reciprocal of the period.

F_0 Onset—The first ten vocal cycles of V2 following the consonant were identified using the negative peaks and marked on a peak-to-peak basis as vocal cycles 1 through 10. The onset of voicing was identified as the first visible periodic vocal cycle following the voiceless stop consonant (vocal cycle 1). As for F_0 offset, the time values for each voiced period mark were then extracted from CSL and the period for each vocal cycle was calculated. The F_0 for each vocal cycle was determined as the reciprocal of the period.

If distinct vocal cycle patterns were not present for both the voicing onset and offset, the V-C-V sequence was not included in the analysis. A total of 51 V-C-V sequences were excluded from the analyses due to this reason.

Although the participants in this study were approximately the same age, it has been recommended that F_0 data be normalized when comparing across groups in order to minimize the effects of intrasubject and intersubject variability associated with intonation contours (Watson, 1998). As such, the F_0 values of the ten vocal cycles for V1 and V2 were normalized by converting them to semitone values relative to a reference vocal cycle. The reference cycle for both the F_0 offset and F_0 onset was the vocal cycle furthest from (prior to and following) the consonant (vocal cycle -10 for offset and vocal cycle 10 for onset). The equation that was used to convert the frequencies to semitones was 39.86 * LOG10 (FREQUENCY/REFERENT).

2.4 Reliability

Identification of the individual vocal cycles was considered the critical measurement procedure necessary to perform the F_0 onset and F_0 offset measurements. Reliability was examined by having a second researcher (JM) analyze data from three randomly selected participants, one participant in the CWNS group and two participants in the CWS group. A total of 36 VCV sequences were analyzed, this accounted for 12.5% of the total 289 VCV sequences used in the statistical analysis. The period of each vocal cycle was compared with the original values obtained for the same vocal cycle. The results of this comparison

indicated a mean difference across all 720 vocal cycles of 0.0123 ms (1.36 Hz). The correlation coefficient between the original and the remeasured cycles was r = .93.

3. Results

Mean and standard deviation semitone values for each participant were derived for each vocal cycle based on utterance type (/itu/ and /apa/). The mean semitone values for the V1 and V2 data sets were separately submitted to $2 \times 2 \times 10$ repeated measures analysis of variance (ANOVA) tests implemented within a general linear model (GLM) (SPSS, 19) to evaluate the interaction of change in average semitone values across the 10 vocal cycles. A similar repeated measures analysis was performed using the standard deviation values to evaluate differences of within group variability. Across all of the repeated measures analysis the between-group factor was the CWS versus CWNS distinction, and the two within-subject factors were utterance type (2 levels; /itu/ and /apa/) and vocal cycles (10 levels). The results of interest were the main effect of vocal cycle, cycle-by-utterance interaction, cycle-by-group interaction and cycle-by-utterance-by-group interaction. An a priori alpha level of .05 was used to determine statistical significance.

All of the repeated measures tests had a p-value < .001 on Machley's test for sphericity across vocal cycles. To compensate for this violation, the degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity on all of the within-group results presented below.

3.1 F_0 Offset

To provide an overall summary of change in F_0 during voicing offset, average semitone values for each of the ten vocal cycles across utterance type from the CWS and CWNS are displayed in Figure 2. Plotted data points represent the average semitone values of 143 V-C-V sequences for the CWS group and average semitone values from 146 V-C-V sequences for the CWNS group. Both groups showed a trend of decreasing semitone values moving closer to the voiceless plosive. The average semitones were submitted to a $2 \times 2 \times 10$ repeated measures analysis described above. The results showed a significant main effect for vocal cycle effect, $F(4.08, 105.98) = 4.75$, $p = .001$, partial eta square = .16. Post hoc comparisons using Tukey HSD test indicated that the mean semitone value for the vocal cycle closest to the consonant (1st vocal cycle) was significantly smaller than the referent (10th) vocal cycle. There was no significant cycle-by-utterance interaction, $F(4.7, 121.98) = 0.35$, $p = .875$, partial eta square = .01. There was no significant cycle-by-group interaction, $F(4.08, 105.98) = 1.73$, p = .147, partial eta square = .06. There was no significant cycle-by-utterance-by-group interaction, $F(4.7, 121.98) = 1.24$, $p = .297$, partial eta square = .05. There was also no significant main effect for group, $F(1, 26) = 0.01$, $p = .982$, partial eta square < .01.

The group variability associated with each vocal cycle during voicing offset is plotted in Figure 3. Plotted data points represent the standard deviation of semitone values across utterance type from 143 V-C-V sequences for the CWS group and 146 V-C-V sequences for the CWNS group. Both groups showed a trend toward increasing F_0 variability, with the largest variability associated with those cycles immediately preceding the consonant. The semitone standard deviations were submitted to a $2 \times 2 \times 10$ repeated measures analysis described above. The results showed a significant main effect for vocal cycle effect, $F(3.93, 102.13) = 35.4$, $p < .001$, partial eta square = .58. Post hoc comparisons using Tukey HSD test indicated that the mean of the semitone standard deviation value for the 4th vocal cycle from the consonant was significantly greater than the referent (10th) vocal cycle. There was no significant cycle-by-utterance interaction, $F(5.34, 138.8) = 1.15$, $p = .34$, partial eta square = .04 and there was no significant utterance-by-group interaction, $F(1, 26) = 0.001$, p

= .98, partial eta square < .01. However, there was a significant cycle-by-utterance- by-group interaction, $F(5.34, 138.8) = 2.41$, $p = .036$, partial eta square = .09, and a significant cycle-by-group interaction, $F(3.93, 102.13) = 2.94$, $p = .025$, partial eta square = .11. Due to the interaction with utterance type, the standard deviations associated with the /apa/ and /itu/ utterances are plotted separately in figures 4 and 5 respectively. There was no significant main effect for group, $F(1, 26) = 0.62$, $p = .438$, partial eta square = .02.

One of the primary purposes of the standard deviation analysis was to determine whether the CWS showed greater variability across the vocal cycles compared to the CWNS. However, the non-significant group main effect cannot be sensibly interpreted due to the significant interactions found in the analysis. Based on figure 3, the only cycle in which CWNS show greater variability is the cycle closest to the plosive (1st vocal cycle). It was possible that the lack of a significant group main effect was caused only by the 1st vocal cycle. To test this, a follow-up $2 \times 2 \times 9$ repeated measures analysis was conducted using cycles 2 though 10, excluding the cycle closest to the plosive. In this analysis, there was still no significant main effect for group, $F(1, 26) = 1.41$, $p = .244$, partial eta square = .05, and there was no longer a significant cycle-by-group interaction, $F(3.36, 86.74) = 0.87$, $p = 0.47$, partial eta square = .03. The follow up provides validity to the initial result of a non-significant group main effect. The follow up test also indicates that the initial significant cycle-by-group interaction resulted from the interaction contrast between the 1st and 2nd vocal cycles.

3.2 F_0 Onset

To provide an overall summary of change in F_0 during voicing onset, average semitone values for each of the ten vocal cycles across utterance type from the CWS and CWNS are displayed in Figure 2. Plotted data points represent the average semitone values of 143 V-C-V sequences for the CWS group and average semitone values from 146 V-C-V sequences for the CWNS group. Both groups showed a trend of decreasing semitone values moving away from the voiceless plosive. The average semitones were submitted to a $2 \times 2 \times 10$ repeated measures analysis described above. The results showed a significant main effect for vocal cycle effect, $F(2.63, 68.34) = 15.04$, $p < .001$, partial eta square = .37. Post hoc comparisons using Tukey HSD test indicated several significant contrasts in the mean semitone values across vocal cycles. The three vocal cycles closest to the consonant were significantly greater than the three vocal cycles farthest from the consonant, and the 4th and 5th vocal cycles were significantly greater than the referent vocal cycle. There was a significant cycle-by-utterance interaction, $F(4.28, 111.4) = 2.81$, $p = .026$, partial eta square = .11. Post hoc comparisons indicated that the mean semitone value for the vocal cycle closest to the consonant was significantly larger for /apa/ compared to /itu/. There was no significant cycle-by-group interaction, $F(2.63, 68.34) = 0.56$, $p = .621$, partial eta square = .02. There was no significant cycle-by-utterance- by-group interaction, $F(4.28, 111.4) = 1.08$, $p = .37$, partial eta square = .04. There was also no significant main effect for group, $F(1, 26) = 0.01$, $p = .321$, partial eta square = .01.

The group variability associated with each vocal cycle during voicing onset is plotted in Figure 3. Plotted data points represent the standard deviation of semitone values across utterance type from 143 V-C-V sequences for the CWS group and 146 V-C-V sequences for the CWNS group. Both groups showed a trend toward decreasing F_0 variability, with the largest variability associated with those cycles immediately following the consonant. The semitone standard deviations were submitted to a $2 \times 2 \times 10$ repeated measures analysis described above. The results showed a significant main effect for vocal cycle effect, $F(3.68, 95.71) = 34.7$, $p < .001$, partial eta square = .57. Post hoc tests indicated that the mean of the semitone standard deviation value for the first 6 vocal cycles closest to the consonant were greater than the referent (10th) vocal cycle. There was no significant cycle-by-utterance interaction, $F(3.83, 99.5) = 1.01$, $p = .402$, partial eta square = .04. There was no

significant cycle-by-group interaction, $F(3.68, 95.71) = 0.60$, $p = .65$, partial eta square $= .02$. There was no significant cycle-by-utterance-by-group interaction, $F(3.83, 99.5) = 2.41$, $p = .476$, partial eta square $= .03$. There was also no significant main effect for group, $F(1, 26) = 0.37$, $p = .55$, partial eta square $= .01$.

4. Discussion

Both the CWS and CWNS in this study showed a significant decrease in vowel F_0 immediately preceding a voiceless stop consonant, and higher F_0 values at voicing onset, followed by significantly decreased vowel F_0 at the tenth vocal cycle following the consonant. There were no significant between-group differences in either mean semitone values or related standard deviations for any of the measures obtained. This confirms our hypothesis that CWS show the same patterns of relative F_0 during voicing onset and offset. During both voicing offset and onset there were no significant group effects for standard deviations, this goes against our hypothesis that CWS would be more variable than the CWNS. Our findings are in agreement with previous research showing a pattern of decreasing F_0 preceding a voiceless consonant (Robb & Smith, 2002), and extend this work by providing evidence that a specific pattern of F_0 change in the vowels surrounding a voiceless obstruent is in place from a younger age than previously shown. This observation offers additional support for an immutable interaction of aerodynamic and physiologic responses associated with vowel-voiceless consonant sequences. Further, present findings show that this interaction remains stable regardless of phonetic context, while also providing new evidence that developmental stuttering close to onset is not characterized by a pervasive deficit in the coordination between the respiratory and laryngeal systems that is observable at the acoustic level.

The current study used a measure that looked at subtle changes in F_0 that are the result in laryngeal/aerodynamic coupling necessary for stop consonant production. Given that we found no significant difference between the patterns of F_0 change in the phonetic contexts of /itu/ and /apa/, it appears that the laryngeal/aerodynamic coordination is similar across phonetic contexts as long as the consonant is a stop plosive. Our measure may be sensitive to the function and stability of the individuals stored motor commands for producing a VCV sequence that is quite stable across linguistic contexts. However, based on Robb and Smith's (2002) findings that the voicing and devoicing gestures become more stable with age, it can be concluded that these stored motor plans are refined over development. Given that the participants in this study were aged matched across groups, the similar degrees of variability between the CWS and CWNS provide evidence that the two groups follow a similar developmental trajectory with regard to the acquisition of mature aerodynamic and laryngeal strategies for voicing onset and offset.

The analyses of standard deviations showed that there was no overall group difference in variability across the vocal cycles. However, in the voicing offset analysis there was a significant cycle-by-group interaction which was likely due to a single vocal cycle. We interpret this interaction to mean that there may be differences between groups in the configuration of variability across the vocal cycles, but with regard to the discussion of whether there are overall group differences in motor stability this interaction does not seem relevant. Our results showing that CWS do not differ from CWNS in overall variability do not agree with previous research showing that people who stutter are more variable than fluent speakers (Kleinow & Smith, 2000; Smith, Sadagopan, Walsh & Weber-Fox, 2010). This is likely due to the measure and level of analysis. If in fact the F_0 changes during voicing and devoicing in a VCV reflect stored motor plans for that particular phonetic sequence then it appears that there is no difference between CWS and CWNS in terms of their stored motor plans and their ability to execute those plans. The group differences in

global measures of stability are likely due to some upstream interactions between speech and language. This may suggest that if differences exist in the fluent speech of CWS and CWNS, it does not exist at the segmental level of motor execution, but rather at a more upstream suprasegmental level that involves the interaction of many factors: linguistic, cognitive and emotional (Smith & Kelly, 1997).

As discussed earlier, these patterns of devoicing and voicing onset have been shown to be present in adults and children down to four years of age (Robb & Smith, 2002). Robb and Saxman (1985) suggested that instantaneous F_0 may undergo a developmental process which emerges sometime between two and four years of age. In the present study, there was an age range from 30 to 57 months with an average of 41.9 months. Our results provide evidence that a similar pattern of instantaneous F_0 is present prior to age four. However, our results found standard deviations that exceeded the four year old data from Robb and Smith (2002) by nearly one semitone, providing even further evidence that the development of mature patterns of instantaneous F_0 begin highly variable and stabilize throughout development. It is unclear whether this development is due to morphological changes in the vocal apparatus during this period or whether it is due to improved coordination between the respiratory and laryngeal systems; it is likely to be a combination of both.

5.0 Conclusion

Since the change and variability of change in vowel fundamental frequency in phonetically governed devoicing and voicing gestures are thought to reflect the coordination of the laryngeal and respiratory systems, this study provides evidence that CWS and CWNS do not differ in regards to the coordination of these two systems, at least within the context of fluent speech. We examined perceptually fluent utterances in the present study; this is important to note given the observation by Sacco and Metz (1986, 1989), that compared to normally fluent adults, adults who stutter showed significantly greater period-by-period F_0 onset variability when utterances containing disfluencies were combined with similar fluent utterances within the same sample. The authors attributed this phenomenon to the "spread" or "vicinity" effect of stuttering on adjacent fluently produced words.

To our knowledge, there has been no investigation of "spread" or "vicinity" effects of stuttering in CWS. Given the present findings that phonetically governed voicing seems to be robust across phonetic contexts, this measurement appears well-suited as a reflection of "spread" or "vicinity" effects of stuttering in the speech of CWS. Longitudinal assessment of these voicing patterns in utterances controlled for phonetic context that contain stuttered disruptions and are produced fluently would provide an opportunity to observe these so-called "vicinity" effects as they may change over time. Determining whether "vicinity" effects are present in children would provide insight in to the development and nature of stuttering.

Acknowledgments

References

Adams M. Voice onsets and segment durations of normal speakers and beginning stutterers. Journal of Fluency Disorders. 1987; 12:133–139.

Baken R, Orlikoff R. Changes in vocal fundamental frequency at the segmental level: Control during voiced fricatives. Journal of Speech and Hearing Research. 1988; 31:207–211. [PubMed: 3398494]

Caruso A, Conture E, Colton R. Selected temporal parameters of coordination associated with stuttering in children. Journal of Fluency Disorders. 1988; 13:57–82.

Chang S, Ohde R, Conture E. Coarticulation and formant transition rate in young children who stutter. Journal of Speech, Language and Hearing Research. 2002; 45:676–688.

Conture E, Colton R, Gleason J. Selected temporal aspects of coordination during fluent speech of young stutterers. Journal of Speech and Hearing Research. 1988; 31:640–653. [PubMed: 3230894]

Cross D, Luper H. Voice reaction time of stuttering and nonstuttering children and adults. Journal of Fluency Disorders. 1979; 4:59–77.

Edwards E, Shriberg. Speaking fundamental frequency characteristics of stutterers and nonstutterers. Journal of Communication Disorders. 1983; 15:21–29.

Goberman A, Blomgren M. Fundamental frequency change during offset and onset of voicing in individuals with Parkinson's disease. Journal of Voice. 2008; 22:178–191. [PubMed: 16950600]

Hall K, Yairi E. Fundamental frequency, jitter and shimmer in preschoolers who stutter. Journal of Speech and Hearing Research. 1992; 35:1002–1012. [PubMed: 1447911]

House A, Fairbanks G. The influence of consonant environment upon the secondary acoustical characteristics of vowels. Journal of the Acoustical Society of America. 1953; 35:84–92.

Kleinow J, Smith A. Influences of length and syntactic complexity on the speech motor stability of the fluent speech of adults who stutter. Journal of Speech Language and Hearing Research. 2000; 43:548–560.

McKnight S, Cullinan W. Disfluencies, utterance length, and linguistic complexity in nonstuttering children. Journal of Fluency Disorders. 1987; 14:17–36.

Peterson, C. Unpublished master's thesis. University of Iowa; 2001. Interarticulatory timing in children who do and do not stutter.

Robb M, Saxman J. Developmental trends in vocal fundamental frequency in young children. Journal of Speech and Hearing Research. 1985; 28:421–427. [PubMed: 2981049]

Robb M, Smith A. Fundamental frequency onset and offset behavior: A comparative study of children and adults. Journal of Speech Language and Hearing Research. 2002; 45:446–456.

Sacco, P.; Metz, D. Fundamental frequency declination patterns in stutterers and nonstutterers. Paper presented at the meeting of American Speech-Language-Hearing Association; Detroit, MI. 1986.

Sacco P, Metz D. Comparison of period-by-period fundamental frequency of stutterers and nonstutterers over repeated utterances. Journal of Speech and Hearing Research. 1989; 28:430–444.

Smith, A.; Kelly, E. Stuttering: A dynamic, multifactorial model. In: Curlee, RF.; Siegel, GM., editors. Nature and treatment of stuttering: New directions. 2nd. Needham Heights, MA: Allyn & Bacon; 1997. p. 204-217.

Smith A, Sadagopan N, Walsh B, Weber-Fox C. Increasing phonological complexity reveals heightened instability in inter-articulatory coordination in adults who stutter. Journal of Fluency Disorders. 2010; 35:1–18. [PubMed: 20412979]

Stepp C, Hillman R, Heaton J. A virtual trajectory model predicts differences in vocal fold kinematics in individuals with vocal hyperfunction. Journal of Acoustical Society of America. 2010; 127:3166–3176.

Stepp C, Merchant G, Heaton J, Hillman R. Effects of voice therapy on relative fundamental frequency during voicing offset and onset in patients with vocal hyperfunction. Journal of Speech, Language and Hearing Research. 2011; 54:1260–1266.

Subramanian A, Yairi E, Amir O. Second formant transitions in fluent speech of persistent and recovered preschool children who stutter. Journal of Communication Disorders. 2003; 36:59–75. [PubMed: 12493638]

Till J, Reich A, Dickey S, Sieber J. Phonatory and manual reaction times of stuttering and nonstuttering children. Journal of Speech and Hearing Research. 1983; 26:171–180. [PubMed: 6887802]

Watson B. Fundamental frequency during phonetically governed devoicing in normal and aged speakers. Journal of the Acoustical Society of America. 1998; 103:3642–3647. [PubMed: 9637045]

Yaruss J, Conture E. F2 transition during sound/syllable repetitions of children who stutter and predictions of stuttering chronicity. Journal of Speech and Hearing Research. 1993; 36:883–896. [PubMed: 8246477]

Zebrowski P, Conture E, Cudahy E. Acoustic analysis of young stutters' fluency: Preliminary observation. Journal of Fluency Disorders. 1985; 10:173–192.

Biographies

Richard M. Arenas, M.A., University of Iowa, Rick Arenas is a doctoral candidate at the University of Iowa. His research interests include mechanisms of online speech monitoring as well as biological mechanisms of learning and how individual variability in these mechanisms relates to stuttering development and treatment.

Patricia M. Zebrowski, Ph.D., University of Iowa, Tricia Zebrowski is an ASHA Fellow and Board Recognized Fluency Specialist. She has authored numerous research and clinical papers, book chapters, videos, and a manual for stuttering intervention. Her research focuses on potential risk factors in the development of different subtypes of stuttering in children and the physiological correlates of stuttering.

Jerald B. Moon, Ph.D., University of Iowa, Jerry Moon is a professor in the Department of Communication Sciences and Disorders at The University of Iowa. His research interests include issues of speech motor control and articulatory coordination.

Educational Objectives

The reader will be able to: (1) discuss the importance of investigating children who stutter close to the onset of stuttering; (2) describe the typical change in F_0 during voicing onset; and (3) discuss the potential implications of these results with regard to future research.

Highlights

Mature pattern of voicing onset and offset is present by age 3;6. This study suggests that there is no difference between CWS and CWNS in the coordination of respiratory and laryngeal systems during voicing onset or offset.
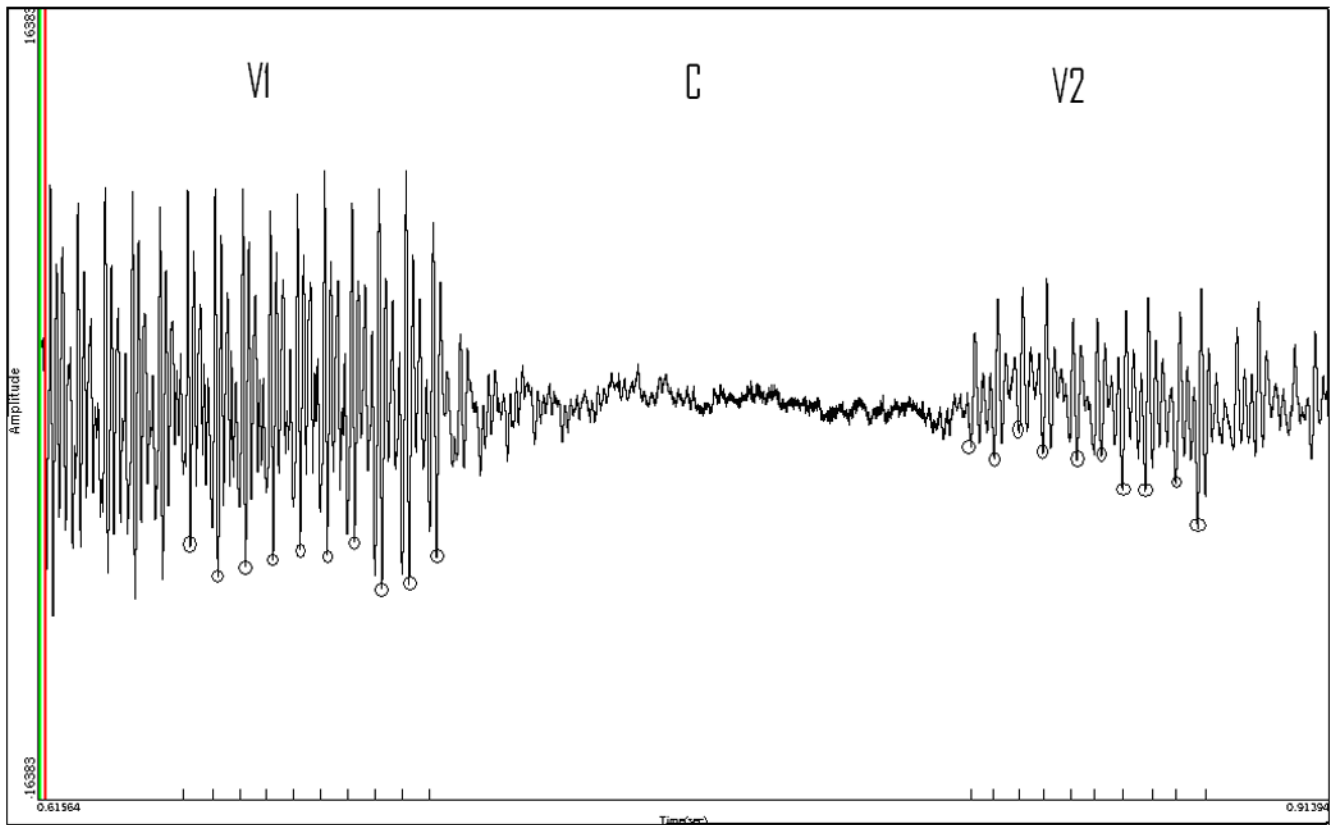
Figure 1.
Sample waveform depicting /apa/ (V1-C-V2) from the word "papa" Voice offset (V1) and voice onset (V2) vocal cycles determined from negative peak to peak basis are highlighted with circles and shown with tic marks on the x axis.
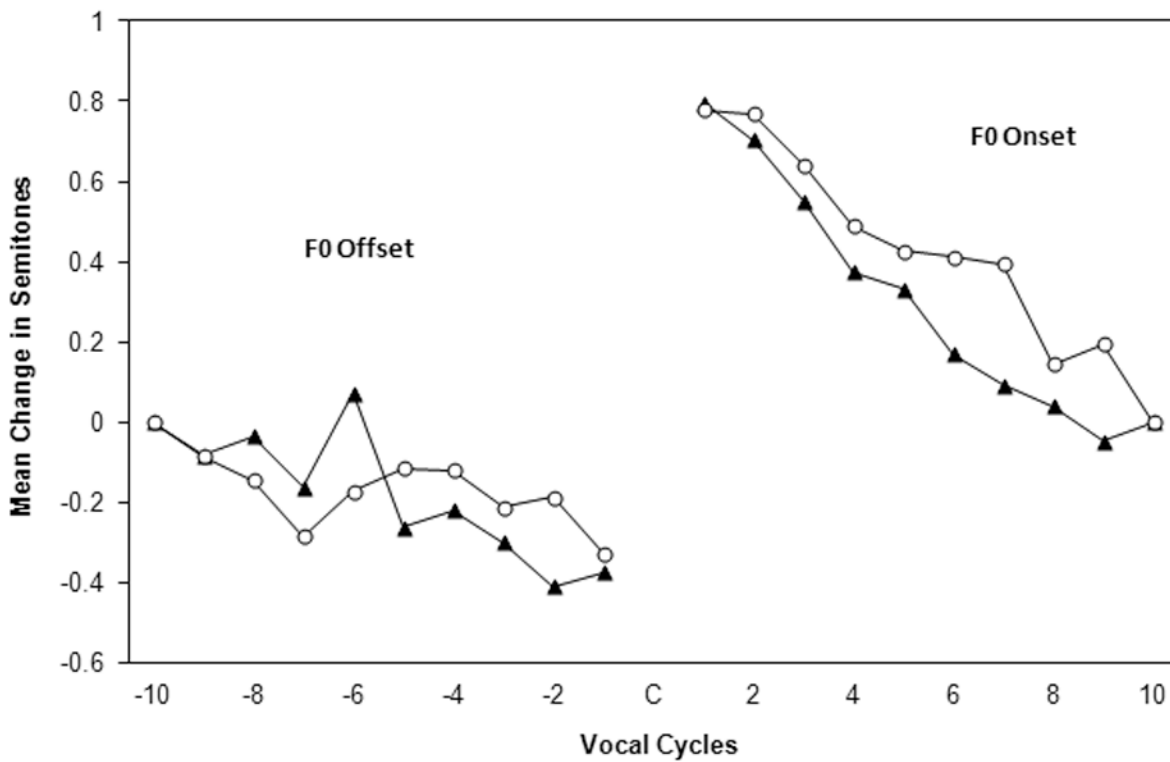
Figure 2.
Average vocal cycle semitone values for F_0 offset and F_0 onset comparing CWS and CWNS. These values are based on both the /apa/ and /itu/ utterances. Values -10 to -1 indicate the last 10 vocal cycles preceding consonant production (offset), and values 1 to 10 indicate the first 10 vocal cycles following consonant production (onset). The black triangles represent the CWS, the open circles represent the CWNS.
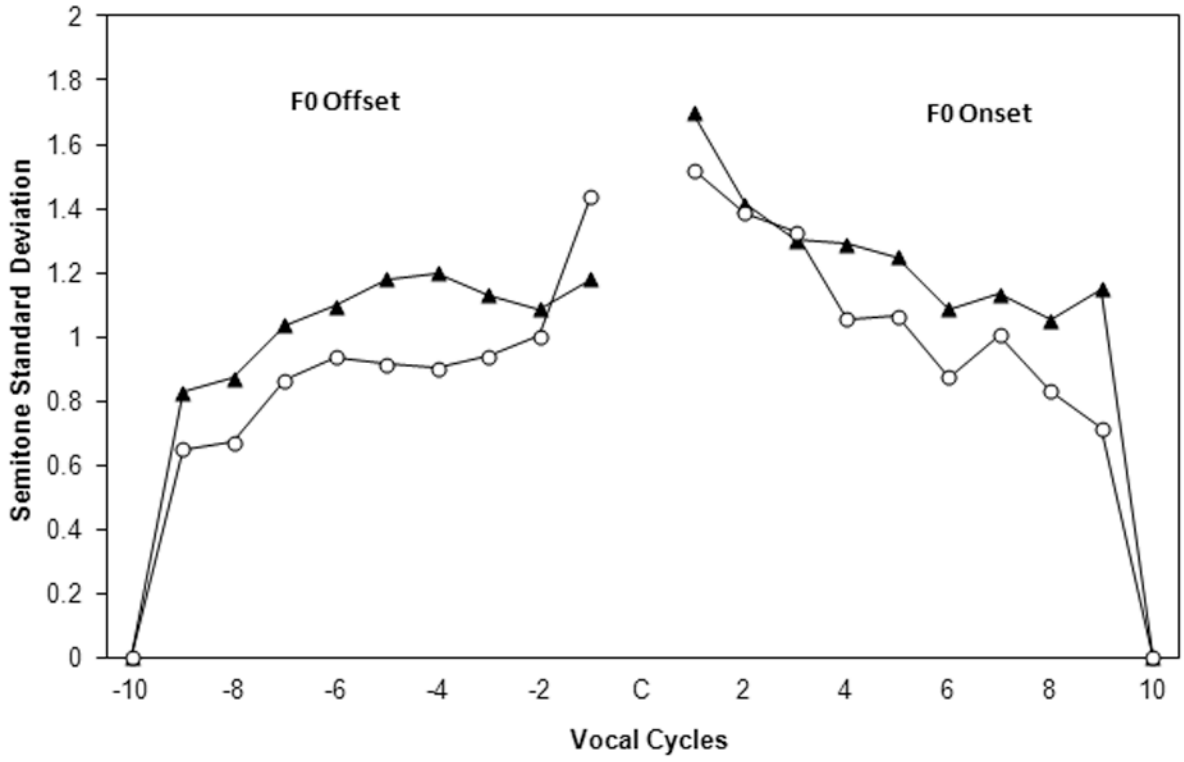
Figure 3.
Variability (standard deviations) of semitone values for F_0 offset and F_0 onset comparing CWS and CWNS. These values are based on both the /apa/ and /itu/ utterances. Values -10 to -1 indicate the last 10 vocal cycles preceding consonant production (offset), and values 1 to 10 indicate the first 10 vocal cycles following consonant production (onset). The black triangles represent the CWS, the open circles represent the CWNS.
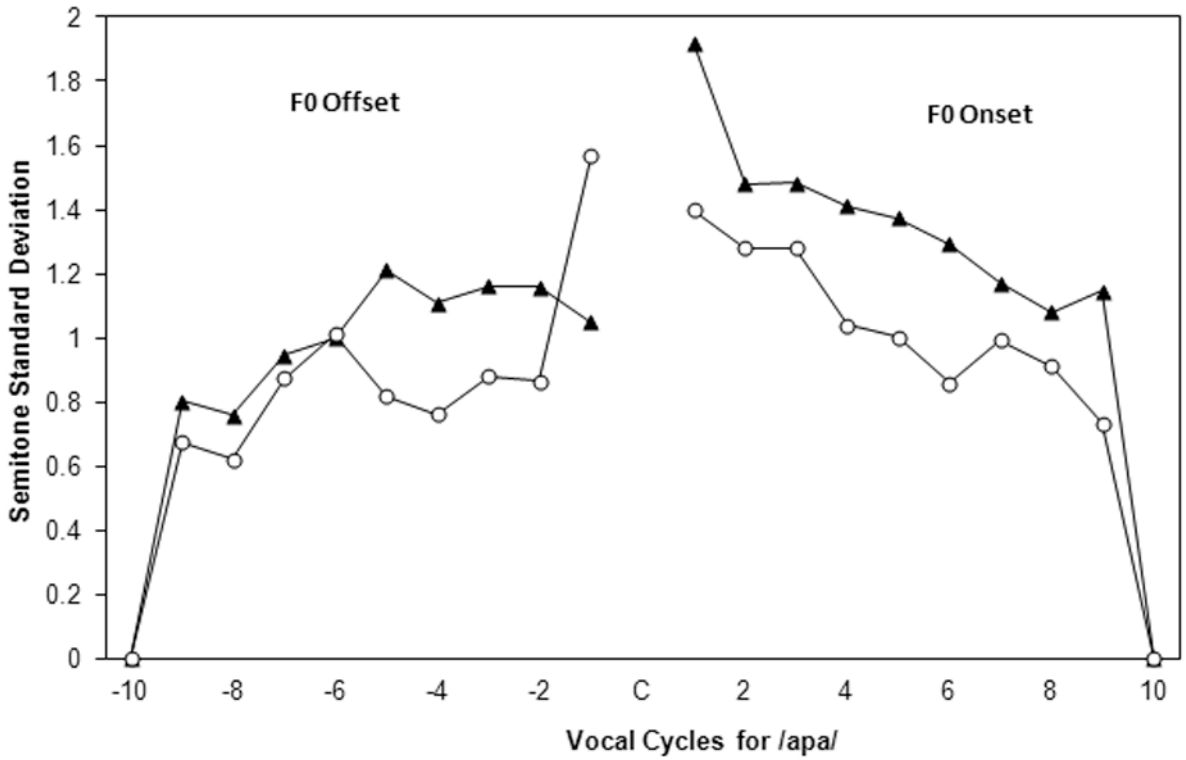
Figure 4.
Variability (standard deviations) of semitone values for F_0 offset and F_0 onset comparing CWS and CWNS. These values are based on only the /apa/ utterances. Values -10 to -1 indicate the last 10 vocal cycles preceding consonant production (offset), and values 1 to 10 indicate the first 10 vocal cycles following consonant production (onset). The black triangles represent the CWS, the open circles represent the CWNS.
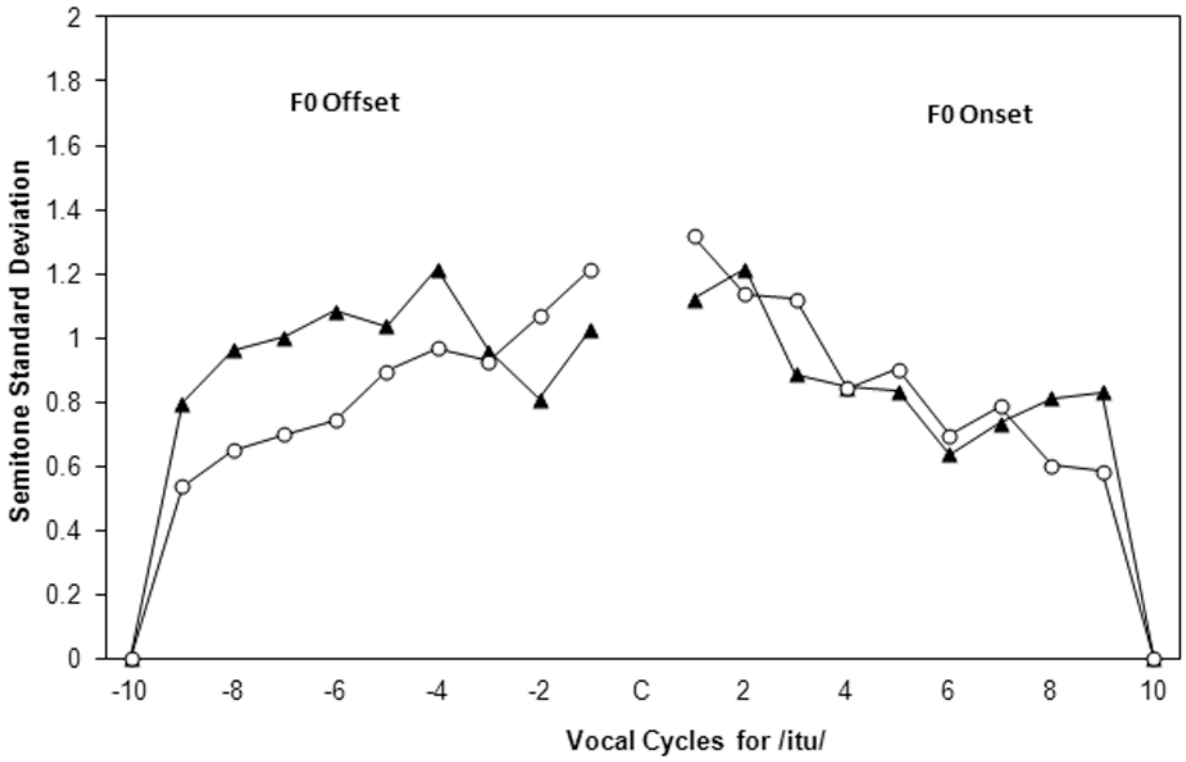
Figure 5.
Variability (standard deviations) of semitone values for F_0 offset and F_0 onset comparing CWS and CWNS. These values are based on only the /itu/ utterances. Values -10 to -1 indicate the last 10 vocal cycles preceding consonant production (offset), and values 1 to 10 indicate the first 10 vocal cycles following consonant production (onset). The black triangles represent the CWS, the open circles represent the CWNS.