Effects of age, sex and syllable structure on voice onset time: Evidence from children's voiceless aspirated stops

Vickie Y. Yu,

Division of Neurology, Hospital for Sick Children, Canada; Neurosciences and Mental Health, Sick Kids Research Institute, Canada

Luc F. De Nil, and

Department of Speech-Language Pathology, University of Toronto, Canada

Elizabeth W. Pang

Division of Neurology, Hospital for Sick Children, Canada; Neurosciences and Mental Health, Sick Kids Research Institute, Canada

Abstract

Voice onset time (VOT) is a temporal acoustic parameter that reflects motor speech coordination skills. This study investigated the patterns of age and sex differences across development of voice onset time in a group of 70 English-speaking children ranging in age from 4.1 to 18.4 years and 12 young adults. The effect of the number of syllables on VOT patterns was also examined. Speech samples were elicited by producing syllables /pa/ and /pataka/. Results supported previous findings showing that younger children produce longer VOT values with higher levels of variability. Markedly higher VOT values and increased variability were found for boys at ages between 8 to 11 years, confirming sex differences in VOT patterns and patterns of variability. In addition, all participants consistently produced shorter VOT with higher variability for multisyllables than monosyllables, indicating an effect of syllable number. Possible explanations for these findings and clinical implications were discussed.

Keywords

Voice onset time; sex differences; variability patterns; syllable number

Introduction

Voice onset time (VOT) is an objective acoustic parameter that reflects subtle motor speech coordination skills. It is a salient temporal cue marking voiced and voiceless phonemic contrasts in word-initial English stops (e.g., /p/ vs. /b/, /t/ vs. /d/). In the past few decades, researchers have extensively studied VOT patterns by measuring the mean VOT values to examine the subtle differences of VOT patterns between children and adults to understand how this temporal parameter of speech component is acquired during childhood. Despite a

vast literature, there is not yet any consensus with regards to both age-related changes in VOT and differences in VOT as a function of sex. In this study, we investigated the change of VOT patterns as a function of age and sex with a wide age range of participants; moreover, we examined the role of syllable number in relation to VOT productions as we hypothesized that syllable number might be an important factor that impacts VOT stability that may reflect motor speech skills in child speech development.

Studies (e.g., Macken & Barton, 1980; Preston & Yeni-Komshian, 1967; Preston, Yeni-Komshian, Stark, & Port, 1968) of VOT acquisition have reported that English-speaking children require a longer time to master the production of voiceless stops during the early stage of motor speech development as those sounds require fine temporal coordination to delay the onset of laryngeal vibration relative to oral closure release. Research studies have reported that English-speaking children attain adult-like VOT patterns for voiceless stops at a fairly late age (e.g., Kewley-Port & Preston, 1974; Lowenstein & Nittrouer, 2008; Macken & Barton, 1980; Ohde, 1985; Zlatin & Koenigsnecht, 1976); however, the direction of differences between adults and children has not yet been defined consistently. Some authors reported that the mean VOT productions were fairly short in most children aged four to six relative to adults (e.g., Kewley-Port & Preston, 1974; Macken & Barton, 1980; Zlatin & Koenigsnecht, 1976), whereas other authors observed longer VOT productions in children compared to adults, and occasionally fairly long VOT productions in the age 4 years and younger group (e.g., Barton & Macken, 1980; Gilbert, 1977; Menyuk & Klatt, 1975; Smith, 1978).

While reports of mean VOT patterns in children's speech development have not been consistent, a common finding across studies is that children produce voiceless VOT stops with higher variability than adults (e.g., Barton & Macken, 1980; Koenig, 2000; Kewley-Port & Preston, 1974; Macken & Barton, 1980; Menyuk & Klatt, 1975; Whiteside, Dobbin, & Henry, 2003; Zlatin & Koenigsnecht, 1976). This finding was supported by evidence from studies in child speech development that have documented a decreasing variability in repeated productions of words or sounds and this trend extended into the adolescent years (e.g., Eguchi & Hirsh, 1969; Munson, 2004; Ohde, 1985; Walsh & Smith, 2002). Likewise, in the case of mean VOT, high VOT variability in children may suggest that they are in the process of mastering the coordination of vocal-fold vibration and oral release (Eguchi & Hirsh, 1969; Kewley-Port & Preston, 1974; Whiteside et al., 2003). Whiteside, Dobbin, and Henry (2003) were the first to focus mainly on observing variability patterns specifically in VOT of children with ages ranging from 5.8 to 13.2 years. Their findings suggested a decrease in variability between age 5.8 and 11.1 years, which indicates maturing motor speech skills as children approach adolescence. However, they did not address sex differences in the study. Little is known about sex differences in variability patterns during VOT acquisition.

Compared to age-correlated patterns of VOT productions, only a few studies in the literature have addressed sex differences in VOT productions with the majority of studies conducted with adult speakers. The findings from these studies are inconsistent; thus, there is a need to include an adult cohort in further studies. Some studies report longer VOT values in adult females (e.g., Koenig, 2000; Ryalls, Zipprer, & Baldauff, 1997; Swartz, 1992; Whiteside &

Irving, 1997), whereas Smith (1982) reported longer VOT values in adult males. On the other hand, Sweeting and Baken (1982) did not find any differences in VOT productions between adult males and females. In terms of VOT acquisition, little is known about sex differences in VOT patterns between boys and girls. To our knowledge, only two studies regarding this issue have been published, both by Whiteside and colleagues (2001; 2004). Whiteside and Marshall (2001) first measured /p, b, t, d/ VOT patterns for 30 children aged 7, 9, and 11 years. They found different VOT patterns between boys and girls, suggesting that sex differences occur during VOT acquisition. In addition, they observed marked higher VOT values for /p, t, d/ only for boys at age 9 years. They concluded that this sex difference may be related to the rapid changes of fundamental frequency (F0) in boys around 7–8 years, which may be related to anatomical changes occurring at the onset of puberty. In their later study, Whiteside and colleagues (2004) examined VOT patterns for /p, b, t, d, k, g/ in two different vowel contexts with 46 children aged 5;8 to 13;2 years (five groups: 5;8 years, 7;1 years, 9;1 years, 11;1 years, 13;3 years). They reported a markedly longer VOT in girls than boys at the age of 13;2 years. Similarly, they concluded that this sex difference may be related to changes of sexual dimorphism of the larynx and vocal tract during adolescence. This study, however, in contrast to their findings in 2001, did not show a marked increase mean VOT for boys at age 9 years. Thus, it is still unclear whether sex differences occur during VOT acquisition in children.

Based on the inconsistent findings in VOT patterns seen in the literature, the initial goals of this study were two-fold: first was to supplement the literature describing developmental patterns of VOT productions by using a wider age range of participants, including adults, and focusing on changes of VOT values as a function of both age and sex. The second goal was to address the gaps in our knowledge by examining patterns of variability in VOT in children and young adults, and to investigate whether there are any sex differences in patterns of variability during VOT acquisition. In addition, a related question that arose was the maturational time course of the mastery of gesture coordination in more complex utterances (measured as the number of syllables in one speech production). Previous research has indicated that subtle differences among different linguistic tasks (e.g., vowel context, stressed vs. unstressed syllable, words or non-word stimuli, speaking rate) may affect changes in VOT productions (e.g., Ryalls et al., 1997; Whiteside et al., 2004). During child speech development, one syllable words are usually easier and faster to master than multisyllable words (Shulman & Capone, 2010). Likewise, clinically, children with motor speech disorders tend to have more difficulties in correctly producing plosives for /pataka/ than /pa/ due to deficits of temporal and muscular coordination (Marquardt, Sussaan, Snow, & Jacks, 2002; Ogar et al., 2006; Rogers & Storkel, 1999). The findings from the previous literature and our clinical observations led to our other question of whether the number of syllables in one utterance would influence changes of VOT values and/or affect the stability of VOT productions.

In this study, we experimentally used the monosyllable /pa/ and the multisyllable sequence /pataka/ as speech stimuli to examine children's VOT patterns and their variability by measuring VOT for voiceless aspirated stops. We recruited 70 children aged 4 to 18 years and additionally recruited 12 young adults to investigate the full age range and determine when this temporal parameter would be mastered and reach maturity. We predicted that VOT

patterns would be affected by age and sex, and expected to see a decreasing pattern of variability in VOT productions that reflects increasing levels of stability in speech output as a function of maturing motor speech skills. We hypothesized that syllable number would affect stability of VOT productions. Our ultimate goal is to better understand how children acquire the subtle temporal coordination of this speech component, its refinement and mastery to adult-like patterns.

Method

Participants

Data were collected from 70 healthy children (range = 4;1 – 18;4 years) with 36 boys and 34 girls. Children were grouped into seven age groups (see Table 1 for full description). Twelve adults (range = 23.5 – 30.2; mean = 27.6 years) with equal numbers of females (range = 23.5 – 30.2, mean = 27.1 years) and males (range = 26.5 – 30.2, mean = 28.2 years) were recruited to serve as an end-point for indexing full maturity of VOT patterns.

All participants were native English speakers. Adults were healthy without any history of neurological, hearing, and speech-language difficulties. Parents of child participants were interviewed to ascertain that there were no known or suspected histories of speech, language, hearing or developmental disorders. Before the experiment, children received two standardized clinical tests: the Peabody Picture Vocabulary Test (3rd Ed.) (PPVT; Dunn & Dunn, 1997) and the Expressive Vocabulary Test (EVT; Williams, 1997), conducted by research assistants supervised by a neuropsychologist. Results confirmed that all children's scores were at or above expected scores for their ages on PPVT and EVT (see Table 2). Children's speech was also verified to show no signs of articulatory difficulties through observation.

Speech stimuli and recording procedures

Speech stimuli were the monosyllable /pa/ and the multisyllable sequence /pataka/. The intention to use /pa/ and /pataka/ was two-fold: first, these two stimuli were simple enough for our young healthy children; second, in the clinic, /pa/ and /pataka/ are parts of speech stimuli in the Diadochokinetic rate (DDK), one of the motor speech assessment tools for evaluating neuromuscular control in populations with motor speech disorders. Individuals who suffer from motor speech disorders usually have differing degrees of difficulty producing these syllables. Results from this set of data might provide helpful indications, or serve as an index, in the clinic. Prior to data acquisition, the experimenter demonstrated the productions of /pa/ and /pataka/. The stimulus /pataka/ was produced without any word-like stress patterns in a normal speaking manner. The participants were instructed to say /pataka/ at an approximately normal speaking rate, and reminded not to say it too fast and not to exaggerate each sound. The participants practiced each production before the experiment. Recordings for /pa/ and /pataka/ productions were made separately in a sound-proof room. During the recording, participants were supine on the magnetoencephalography (MEG)¹ bed

¹These data were recorded as part of a larger neuroimaging study involving magnetoencephalography (MEG). MEG was used to identify brain regions involved in production of these stimuli.

with a high-fidelity directional condenser microphone (Rode NTG-2) positioned 60 cm from the mouth. They were instructed to say the target sound once immediately when they were cued by the appearance of a small white circle displayed on the monitor. Using a white circle as a cue engaged the young children in the task. The white circle was presented for 200 ms, and the inter-stimulus interval was jittered between 2100 to 2500 ms. Speech productions were recorded via a microphone and saved as digital recordings onto a desktop PC. Each stimulus (/pa/ and /pataka/) yielded 115 trials in total for each participant.

Acoustic measurement and criteria

VOT measures of aspirated voiceless /p/ and /k/ were made from the productions of monosyllable /pa/ (hereafter *mono-/p/*) and multisyllable /pataka/ (hereafter *multi-/p/ and multi-/k/*). Prior to the VOT measurement, all productions of /pa/ and /pataka/ were perceptually evaluated by a speech-language pathologist with linguistics training. For /pa/, if the production was perceived with exaggerated loudness (exaggerated vocal intensity), which was often seen in our younger participants, these productions were not included in the acoustic analysis. For /pataka/, if any of the syllables was perceived as a stressed syllable, for instance, [PAtaKa], [paTAka], [pataKA] or [PATAKA] (capitalized syllable refers to a stressed-syllable), these productions were excluded from the acoustic analysis. Acoustic measures were carried out using Praat acoustic analysis software (Boersma & Weenink, 2008). Figure 1 illustrates the spectrogram of /pa/ and /pataka/ with the measurement boundaries for VOT /p/ and /k/ marked. The measurements for these plosives /p/ and /k/ provided temporal information of VOT across age, sex and syllable structure. VOT measurements were made directly from the spectrograms by measuring the time between the release of the plosive and onset of the first formant of the following vowel. VOT was not measured in the cases where the release burst could not be identified (e.g., plosives released with affrication or background noise from body movement), or where the place of articulation did not match the target. Measurements of VOT for /t/ were not attempted since it is frequently realized as a flap sound [ɾ] as observed in our participants. To ensure consistency in VOT measurements, two months after the original measurements were made, 10% of all tokens for each participant were reanalyzed by the same experimenter following the same procedure and criteria. A Pearson's product-moment correlation analysis showed a significant correlation coefficient (mono-/p/: $r = .984$, $p < .001$; multi-/p/: $r = .974$, $p < .001$; multi-/k/: $r = .966$, $p < .001$) indicating a high level of measurement reliability.

Data analysis

Prior to statistical analysis, three calculations were carried out for each participant as a function of age: mean, standard deviation (SD) and coefficient of variance (CoV). CoV was used to represent the pattern of variability for each participant. The CoV is the ratio of the SD to the mean (in percentage, %), which is used to control for higher SDs due to larger mean values. Data were analyzed in two ways to answer our research questions: 1) mono-/p/ and multi-/p/ were used to answer our research question of whether number of syllables (syllable number) would affect VOT values and whether it would affect the stability of VOT productions as a function of age, and 2) multi-/p/ and multi-/k/ were used to examine effects of age and sex on VOT values and patterns of variability in VOT.

Results

Tables 2 and 3 summarize the mean, SD and CoV VOT values for mono-/p/, multi-/p/ and multi-/k/ by age and sex. The order of magnitude of mean VOT was as follows: the largest order of magnitude was for multi-/k/ (71 + 9.6 ms), followed by mono-/p/ (70 + 5.8 ms) and multi-/p/ (61 + 8.0 ms).

Mean VOT values

Overall, mono-/p/ (70 + 5.8 ms) exhibited longer mean VOT values than for multi-/p/ (61 + 8.0 ms); this was seen for all participants across age and sex. A two-way analysis of variance revealed a significant main effect of syllable number ($F(1, 66) = 9.880, p = .002$). Neither the age effect nor interaction effect between age and syllable number was found to be significant.

A separate two-way analysis of variance was carried out to further examine age and sex effects on mean values for multi-/p/ and multi-/k/. Only main effects of age were found to be significant for both multi-/p/ ($F(7, 66) = 3.216, p = .006$) and multi-/k/ ($F(7, 66) = 4.242, p = .001$). Figures 2 and 3 illustrate the mean VOT values as a function of age and sex for multi-/p/ and multi-/k/ respectively. Consistent with a developmental trend suggested by the figures, group 4–5 displayed a tendency toward longer VOT values compared to the older groups. Post-hoc testing showed that the 4–5 group had significantly longer mean VOT values compared to some of the older groups (i.e., 12–13, 14–15, 16–18; $p < .01$ corrected) for multi-/p/, and the 4–5 group had longer VOT values compared to all the other age groups ($p < .001$ corrected) for multi-/k/. Groups 12–13, 14–15, 16–18 had VOT values similar to those for the adult group (see Figures 2 and 3; Tables 2 and 3). In addition, with respect to sex differences, the data from groups 8–9 and 10–11 suggest longer VOT values for boys than girls. Therefore, in order to examine the effect of sex on VOT values, paired-samples *t*-tests (two-tailed) were performed separately only on these two age groups. Results indicated significant sex differences for groups 8–9 ($t = -5.048, p = .037$) and 10–11 ($t = -4.924, p = .039$) for multi-/p/, and only group 10–11 ($t = -6.538, p = .007$) for multi-/k/ with longer mean values for boys than girls in all these cases.

Variability patterns in VOT

Two-way analyses of variance were further performed separately for multi-/p/ and multi-/k/ to examine age and sex effects for variability patterns in VOT. A significant main effect of age was seen for multi-/p/ ($F(7, 66) = 31.906, p < .001$) and multi-/k/ ($F(7, 66) = 39.154, p < .001$), indicating younger children had higher CoV than older children. In addition, a sex effect was found for multi-/p/ ($F(1, 66) = 7.981, p = .006$) and multi-/k/ ($F(1, 66) = 6.303, p = .015$) with higher CoV in boys than girls. No sex effect was found for the adult group. Figures 4 and 5 illustrate the CoV for VOT by age and sex for multi-/p/ and multi-/k/, respectively. With respect to a developmental trend, both figures show a similar pattern with higher CoV for younger children than older children and a tendency for higher CoV values in groups 4–5 and 6–7. Group 16–18 showed similar CoV values to the adult group (see Figures 4 and 5; Tables 2 and 3).

With respect to the sex differences, the pattern for multi-/p/ and multi-/k/ showed a consistently higher CoV pattern for boys than girls in the younger ages (see Figures 4 and 5), particularly in groups 6–7 and 8–9 for multi-/p/ and, 8–9 and 10–11 for multi-/k/. The mean differences for CoV were fairly comparable in the rest of the age groups, particularly groups 12–13, 14–15, 16–18 and adults. These sex differences were confirmed by paired-sample t-tests (two-tailed) only on these three age groups. Results indicated significant sex differences for group 8–9 ($t = -7.239$, $p = .019$) for multi-/p/; groups 8–9 ($t = -4.874$, $p = .017$) and 10–11 for multi-/k/ ($t = -7.889$, $p = .016$) with higher CoV for boys than girls.

Figure 6 illustrates the CoV patterns of VOT for mono-/p/ versus multi-/p/ as a function of age. CoV for VOT was consistently higher for multi-/p/ than mono-/p/ across age groups. Two-way analysis of variance using values of CoV for VOT revealed significant main effects of syllable number ($F(1, 66) = 102.649$, $p < .001$). Results also revealed a significant interaction effect between age and syllable number ($F(7, 66) = 3.498$, $p = .002$).

In Figure 6, the pattern shows that younger children had significantly higher CoV than older children. This pattern of CoV displayed a continuously decreasing CoV with age, except in group 10–11. When CoV differences between two syllable structures were compared, the patterns showed larger differences for younger children (12% – 9%) and a much smaller difference for older children (groups 14–15, 16–18 with 6% – 2% differences), where groups 16–18 showed nearly the same differences in their CoV patterns as the adults (2%).
(INSERT FIGURE 6 HERE)

Discussion

Our study aimed to examine the developmental trend in VOT patterns and its pattern of variability, particularly focusing on age, sex and syllable number effects. Data were recorded from a wide age range of children and adults who repeatedly produced the monosyllable /pa/ and multisyllable /pataka/. With respect to the age effect on mean VOT productions, as expected our data showed a developmental trend on VOT patterns with longer mean values for younger children than older children (Figures 2 and 3). This trend was marked especially for multi-/k/ VOT productions (Figure 3). Children in the group aged 4 to 5 years had significantly longer mean VOT for multi-/k/ than the others. This result was consistent with previous studies that young English-speaking children (age 4 or younger) produced longer VOT relative to adults (e.g., Barton & Macken, 1980; Gilbert, 1977; Menyuk & Klatt, 1975; Smith, 1978). In addition, our data indicated groups older than age 11 years did not show any mean differences from adults. Consistent with the aforementioned studies, our data indicated that children initially produce longer VOT than adults and continue to refine this pattern and gradually approach adult-like patterns after age 11.

With respect to sex differences for developmental VOT patterns, our data showed significant sex differences marked by increasing mean values for boys aged 8–11 for multi-/p/ and aged 10–11 for multi-/k/ (Figures 2 and 3). Our data are consistent with the results found in Whiteside and Marshall's study (2001), where they observed that 9-year-old boys produced larger VOT values for plosives /p, t, d/ than girls and other age groups. Another goal of this study was to examine the developmental changes in the variability pattern of VOT. As with

the age effect, as expected, our data suggested a developmental trend for variability in VOT productions. Younger children had higher variability in VOT than older children. Children aged 4–7 showed the highest variability in VOT. This pattern of variability decreased as a function of age, until age 16–18, when children exhibited no statistically significant differences in VOT production from adults (Figure 4). These patterns of decreasing variability are indicative of maturing motor speech skills as children approach adolescence. In addition, a marked sex difference was found in children aged 8–9 for multi-/p/ and 10–11 for multi-/k/ (Figures 4 and 5). In our data, boys aged 8–9 and 10–11 produced more variable patterns of VOT than girls.

The sex differences in developmental VOT found in our data may be explained by a number of factors. One highly plausible factor is related to anatomical changes with physical development. The larynx of a developing child undergoes rapid growth and structural change during adolescence and these changes may possibly result in changes in VOT during adolescence (Whiteside & Marshall, 2001). Several studies have reported that the rapid laryngeal growth for males emerges around early adolescence (e.g., Kaplan, 1971; Negus, 1962; Seikel, King, & Drumright, 2000) and that the anatomical changes in the vocal tract affect the changes in the fundamental frequency (F0) in males during this period of time (e.g., Kent, 1976; Hasek & Singh, 1980). Kent (1976) compiled data on F0 values for children from a number of studies, and reported that sex differences started to emerge before age 11 and F0 declined more gradually for boys around age 11 or 12 years. Hasek and Singh (1980) observed changes of F0 for both boys and girls and suggested the emergence of significant sex differences at around age 7–8 due to the rapid changes of F0 in boys. These studies suggested that the sudden F0 changes might be related to anatomical changes in the developing laryngeal and pharyngeal structures that occur at the onset of puberty. It is feasible to assume that sudden increases in the size of the larynx and pharyngeal structures would influence a number of acoustic parameters, such as F0 (Kent, 1976; Hasek & Singh, 1980), formant frequency (e.g., Kaplan, 1960; Kent, 1976), and VOT (Whiteside & Marshall, 2001). Our observation of marked sex differences in VOT patterns between 8 to 11 years, around the time of the onset of the puberty, suggests a strong link between the anatomy and physiology of the developing laryngeal system and VOT productions, especially as this difference was mainly driven by the VOT productions in boys and not in girls.

Taken together with the observed trajectory of the VOT pattern across age, the findings indicate that VOT productions are developmentally sensitive. The developmental changes in pattern of variability of the VOT data provide evidence to suggest that the levels of stability in speech output is a function of maturing motor speech skills (Whiteside et al., 2003). Children, in their early years, produced markedly longer mean VOTs with more variable patterns, indicating that their motor speech skills have not yet fully developed. Children continue to develop, refine, and control their motor speech skills while constantly undergoing anatomical and physiological development toward maturity. The transition between stages is limited by other factors, such as anatomy and physiology, and therefore cannot be achieved simply by practice and repetition. However, repetition and practice may help expedite development within a stage during the maturation process; alternatively, it may help facilitate the acquisition of compensatory skills in situations where factors exist that

cause deviation from the mastery of typical speech production (such as motor speech disorders, articulation disorders, etc.). In the current study, boys between 8 to 11 years may require more and continuous practice to learn to make adjustments in the timing and phasing of oral motor coordination in order to accommodate the dramatic anatomical and physiological changes in the laryngeal systems.

On the other hand, in our data, the marked increase of VOT values, seen for boys at age 8–11 years, was not seen on the girls' VOT data, and may be due to the less dramatic vocal tract changes that occur in females during puberty. However, in contrast to the finding of longer VOT for girls than boys at age 13.2 found in Whiteside et al.'s study (2004), and other previous studies that reported a significantly longer VOT value in adult females than adult males (Koenig, 2000; Ryalls et al., 1997; Swartz, 1992), our data did not show sex differences for children after age of 11 years. Possibly, this difference can be attributed to the fact that different speech materials (i.e., vowels, CV structures, words vs. nonwords, phrases vs. sentences) were examined or different sample size were involved in the study.

While the marked sex differences in VOTs in this study seems to have a strong link to the anatomy and physiology of the laryngeal system, findings on the age and staging of pubertal changes, as well as how changes in the voice correlate with other pubertal developments are somewhat controversial. Several studies suggested the onset of adolescent voice changes happen at later ages. For example, Pederson et al. (1985) found that the phonetogram (i.e., "Voice Range Profile" or a visual display of vocal intensity versus F0) range in boys begins to shrink after age 12. Hollien, Green, and Massey (1994) estimated that, for most boys, the onset of adolescent F0 changes ranges from 12.5 to 14.5 years. In addition, anatomical studies have indicated that supraglottal vocal tract differences between boys and girls do not become evident until about 14 or 15 years of age (Fitch & Giedd, 1999; Vorperian et al., 2005; 2011). Given the varied findings from these physiological and anatomical studies, the possibility is raised that the sex differences in VOT development observed in this study may be influenced by, or include, other factors. A few studies examining the acoustic features involved in speech production have suggested that sex differences in voice uses may be attributed in part to sociocultural / sociophonetic convention (Hasek & Singh, 1980; Van Bezooijen, 1995) or geographical differences (Stoddart, Upton, & Widdowson, 1999; Whiteside & Irving, 1997; Whiteside & Marshall, 2001). Although all participants in this study were recruited from the same geographical region, this region is a large urban centre with a rich multi-cultural and multi-lingual environment. Sex differences in VOT development found particularly in boys between aged 8 and 11 years could be associated with the influence of different cultural backgrounds or/and language environments. Clearly further research is needed to investigate the possibility of sociophonetic and cultural factors on the developmental VOT in this specific region.

With regard to the effect of syllable number on VOT productions, our study found a significant syllable number effect across all participants. Our data showed that the number of syllables in one speech production influenced mean VOT productions across all age groups. All participants consistently produced longer mean VOT values for plosives in monosyllables than multisyllables. This could be attributable to differences in the need for respiratory support or possibly coarticulation during speech production tasks. In our study,

participants were instructed to produce /pa/ or /pataka/ in one utterance. Normal speech sounds are produced during exhalation. Producing /pa/ may allow participants to have higher lung volumes, larger capacities, and higher expiratory flows than when producing /pataka/. Alternatively, there may be differences in production due to coarticulation. In /pataka/, participants were likely preparing to produce the second syllable while producing the first, which could reduce VOT. Further research is warranted to investigate this. In terms of stability of the productions in monosyllables vs. multisyllables, as we expected, a significant syllable number effect was found with higher stability in producing monosyllables than multisyllables, particularly for young children (Figure 6). This effect was attributed to the increase in the number of syllables. As aforementioned, in child speech development, simple words (CV, CVC) are first and easier for children to produce successfully whereas multisyllable words usually are mastered in later ages, as they require more mature motor speech skills. In the current study, all participants across all age groups produced less variable VOT patterns for monosyllabic than multisyllabic speech stimuli. The variability for both syllable structure decreased as a function of age, which suggests an increase in the stability of motor speech skills.

Conclusion and Future Directions

The current study adds to the body of data describing developmental trends of VOT productions. In summary, we found evidence to support previous findings that younger children produce longer VOT values than adults. Our data confirm sex differences in VOT values for plosives. Boys aged 8–11 exhibited markedly higher VOT values than girls across all of the age groups, suggesting emerging sex differences. Furthermore, our data suggested that young children aged 4–8 produced more variable VOT productions relative to adults; the pattern of variability in VOT continually decreased as a function of age until 16 years, when it achieved a more adult-like pattern. These patterns of decreasing variability indicate increasing levels of stability in motor speech skills as children approach adolescence. The sex differences in VOT development in this study may be explained by the physiological and anatomical differences and changes during puberty; however, conclusions should be drawn cautiously as there is not yet extensive consensus in the literature. Some studies link the changes of VOT production to sub- and supra-glottal pressure, other aerodynamic laryngeal-related factors (Koenig, 2000), and to changes in F0 (McCrea, & Morris, 2005). While Koenig (2000) did not find any significant correlations between VOT and oral airflow and intraoral pressure, they observed only adult males, females and five-year-old children. It may be that they missed the age window (approximately 8–11 years) when change is most pronounced. We speculate an association between the change of VOT values and the aerodynamics of the laryngeal mechanism during this age period. Studies have shown that the size of the larynx and vocal folds affect the aerodynamic characteristics of speech, where adult male voice exhibits a higher subglottal pressure and different patterns of the valving efficiency of the glottis during phonation (Sapienza, 1996; Weinrich, Salz, & Hughes, 2005). Thus, future study is warranted to investigate the relationship between the aerodynamic parameters in the developing larynx and VOT production. On the other hand, a strong relationship between F0 and VOT was found in McCrea and Morris' study (2005), where speech productions with high F0 displayed significantly shorter VOTs than productions at

low or mid F0s. The conclusion drawn from these studies may suggest a robust relationship between VOT and F0. However, since these findings are based on adults and/or children within a limited age range, it is still unclear how these relationships, F0 and VOT, and/or the aerodynamic mechanism and VOT, interact when males are undergoing the rapid growth of their larynx and vocal tracts. Future studies should aim to replicate the findings of the above studies but include with a wider age range, particularly with emphasis on the 8 to 11 year old group.

Sociophonetic or geographical influence might be another contributing factor for the sex differences found in our study. Further investigation is warranted on this topic. We also found evidence that the number of syllables in one utterance influenced VOT values and the stability of productions in VOT. Syllable-initial plosives had longer VOT values for monosyllables, which may be attributed to differences in the need for respiratory support or possibly coarticulation during speech production. Accordingly, a higher variability in VOT was observed for multisyllables. This could be due to the increased number of syllables or increased complexity in syllable structure. This finding requires further investigation in a large group of children, of both sexes, across a wide range of age but with particular emphasis on the age range from 8 to 11 years. Furthermore, these studies should attempt to elicit speech data using different test stimuli such as real words and/or different paradigms (e.g., using carrier phrases “say ____ again”; picture naming) with measures of VOT and possibly other speech parameters (e.g., duration and/or rate of syllables, words, or utterances) to gain a greater insight into the planning and control mechanisms involved in motor speech development. Our data provide evidence of a developmental trend and sex differences in VOT patterns and the data indicate that changes in variability may be an indicator of biological developmental changes. The implications of these findings can be utilized as a reference point when evaluating children’s motor speech skills in clinical practice. As children approach adolescence, boys (and possibly girls) may need to adopt different strategies for producing adult-like VOT patterns until such time as their motor speech systems truly reach maturity.

Acknowledgments

The data reported here were recorded as part of a larger neuroimaging study where brain regions involved in production of these stimuli were also measured. The study was supported by a Canadian Institutes of Health Research operating grant (CIHR MOP-89961) to the last two authors (LDN and EWP). The authors would like to thank Matt MacDonald, Anna Oh, Gordon Hua and Marc Lalancette for acquiring the speech data as part of the neuroimaging study. The authors would like to thank Dr. Darren Kadis for consultation on neuropsychology test measures and for supervising neuropsychology testing, and Dr. Paul Ferrari for assisting with experimental design and equipment setup. Thanks to all the parents and children who participated.

References

- Barton D, Macken M. An instrumental analysis of the voicing contrast in word-initial stops in the speech of four-year-old English-speaking children. *Language and Speech*. 1980; 23:159–169.
- Boersma, P., Weenink, D. Praat: doing phonetics by computer [Computer program]. 2007. Version 4.6.22, from <http://www.praat.org/>. date last viewed 4/16/2012
- Dunn, LM., Dunn, LM. Peabody Picture Vocabulary Test. 3. Circle Pines: MN: American Guidance Service; 1997.

- Eguchi S, Hirsh IJ. Development of speech sounds in children. *Acta Oto-Laryngologica - Supplement*. 1969; 257:1–51.
- Fitch WT, Giedd J. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*. 1999; 106:1511–1522. [PubMed: 10489707]
- Gilbert JHV. A voice onset time analysis of apical stop production in 3-year-olds. *Journal of Child Language*. 1977; 4:103–113.
- Hasek C, Singh S, Murray T. Acoustic attributes of children's voices. *Journal of the acoustical Society of America*. 1980; 68:1262–1265. [PubMed: 7440849]
- Hollien H, Green R, Massey K. Longitudinal research on adolescent voice change in males. *Journal of the Acoustical Society of America*. 1994; 96:2646–2654. [PubMed: 7983270]
- Kaplan, HM. *Anatomy and physiology of speech*. New York: McGraw-Hill; 1960. p. 154-246.
- Kent RD. Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech and Hearing Research*. 1976; 19:421–447. [PubMed: 979206]
- Kewley-Port D, Preston MS. Early apical stop production: a voice onset time analysis. *Journal of Phonetics*. 1974; 2:194–210.
- Koenig LL. Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *Journal of Speech, Language, and Hearing Research*. 2000; 43:1211–1228.
- Lowenstein JH, Nitttrouer S. Patterns of acquisition of native voice onset time in English-learning children. *Journal of the Acoustical Society of America*. 2008; 124:1180–1191. [PubMed: 18681606]
- Macken MA, Baron D. The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*. 1980; 7:41–74. [PubMed: 7372738]
- Marquardt TP, Sussman HM, Snow T, Jacks A. The integrity of the syllable in developmental apraxia of speech. *Journal of Communication Disorders*. 2002; 35:31–49. [PubMed: 11949971]
- McCrea CR, Morris RJ. The effects of fundamental frequency level on voice onset time in normal adult male speaker. *Journal of Speech, Language, and Hearing Research*. 2005; 48:1013–1024.
- Menyuk P, Klatt M. Voice onset time in consonant cluster production by children and adults. *Journal of Child Language*. 1975; 2:223–231.
- Munson B. Variability in /s/ production in children and adults: Evidence from dynamic measures of spectral mean. *Journal of Speech, Language, and Hearing Research*. 2004; 47:58–69.
- Negus, VE. *The comparative anatomy and physiology of the larynx*. New York: Hafner; 1962.
- Ogar J, Willock S, Baldo J, Wilkins D, Ludy C, Dronkers N. Clinical and anatomical correlates of apraxia of speech. *Brain and Language*. 2006; 97:343–350. [PubMed: 16516956]
- Ohde RN. Fundamental frequency correlates of stop consonant voicing and vowel quality in the speech of preadolescent children. *Journal of the Acoustical Society of America*. 1985; 78:1554–1561. [PubMed: 4067069]
- Pedersen MF, Møller S, Krabbe S, Munk E, Bennett P. A multivariate statistical analysis of voice phenomena related to puberty in choir boys. *Folia phoniatrica*. 1985; 37:271–278. [PubMed: 4093088]
- Preston MS, Yeni-Komshian G, Stark RE, Port DK. Developmental studies of voicing in stops. *Haskins Laboratories Status Report on Speech Research*. 1968; SR13/14:181–184.
- Preston MS, Yeni-Komshian G. Studies on the development of stop-consonants in children. *Haskins Laboratories Status Report on Speech Research*. 1967; SR11:49–52.
- Ryalls J, Zipprer A, Baldauff P. A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language and Hearing Research*. 1997; 40:642–645.
- Rogers MA, Storkel HL. Planning speech one syllable at a time: the reduced buffer capacity hypothesis in apraxia of speech. *Aphasiology*. 1999; 13:793–805.
- Sapientza CM. Glottal airflow: Instrumentation and interpretation. *Florida Journal of Communication Disorders*. 1996; 16:3–7.
- Seikel, JA., King, DW., Drumright, DG. *Anatomy and physiology for speech, language, and hearing*. San Diego, California: Singular Publishing Group; 2000. p. 243-248.

- Shulman, B., Capone, NC. Language development: Foundation, and Clinical Applications. Canada: Jones and Bartlett Publishers; 2010.
- Smith BL. Effects of place of articulation and vowel environment on voiced stop consonant production. *Glossa*. 1980; 12:163–175.
- Stoddart, J., Upton, C., Widdowson, JDA. Sheffield dialect in the 1990s: revisiting the concept of NORMs. In: Foulkes, Docherty, editor. *Urban Voices: Accents studies in the British Isles*. Arnold; London: 1999. p. 72–89.
- Swartz BL. Gender differences in voice onset time. *Perceptual and Motor Skills*. 1992; 75:983–992.
- Sweeting PM, Baken RJ. Voice onset time in a normal-aged population. *Journal of Speech and Hearing Research*. 1982; 25:129–134. [PubMed: 7087415]
- Van Bezooijen R. Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*. 1995; 38:253–265. [PubMed: 8816084]
- Vorperian HK, Kent RD, Lindstrom MJ, Kalina CM, Gentry LR, Yandell BS. Development of vocal tract length during early childhood: A magnetic resonance imaging study. *Journal of the Acoustical Society of America*. 2005; 117:338–350. [PubMed: 15704426]
- Vorperian HK, Wang S, Schimek EM, Durtschi RB, Kent RD, Gentry LR, Chung MK. Development of sexual dimorphism of the oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of Speech, Language, and Hearing Research*. 2011; 54:995–1010.
- Walsh B, Smith A. Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research*. 2002; 45:1119–1133.
- Weinrich B, Salz B, Hughes M. Aerodynamic measurements: Normative data for children ages 6:0 to 10:11 years. *Journal of Voice*. 2005; 19:326–339. [PubMed: 16102661]
- Whiteside SP, Irving CJ. Speakers' sex differences in voice onset time, some preliminary findings. *Perceptual and Motor Skills*. 1997; 85:459–463. [PubMed: 9347529]
- Whiteside SP, Marshall J. Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica*. 2001; 58:196–210. [PubMed: 11423786]
- Whiteside SP, Dobbin R, Henry L. Patterns of variability in voice onset time: a developmental study of motor speech skills in humans. *Neuroscience Letters*. 2003; 347:29–32. [PubMed: 12865134]
- Whiteside SP, Henry L, Dobbin R. Sex differences in voice onset time: a developmental study of phonetic context effects in British English. *Journal of Acoustical Society of America*. 2004; 116:1179–1183.
- Williams, KT. Expressive Vocabulary Test. Circle Pines, MN: American Guidance Services; 1997.
- Zlatin MA, Koenigsknecht RA. Development of the voicing contrast: a comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research*. 1976; 19:93–111. [PubMed: 1271805]

Figure 1. Spectrograms for /pa/ (left) and /pataka/ (right; only /p/ and /k/ were measured) with the measurement boundaries for VOT mono-/p/ and multi-/p/ and multi-/k/.

Figure 2. Mean multi-/p/ VOT as a function of age group and sex. Significant mean VOT differences ($*p < .05$) between males and females for age groups 8–9 and 10–11.

Figure 3.

Mean multi-/k/ VOT as a function of age group and sex. Significant mean VOT differences ($*p < .01$) between males and females for age group 10–11.

Figure 4. CoV for VOT for multi-/p/ by age group and sex. Significant CoV VOT differences ($*p < .05$) between males and females for age group 8–9.

Figure 5.
CoV for VOT for multi-/k/ by age group and sex. Significant mean CoV VOT differences (* $p < .05$) between males and females for age groups 8–9 and 10–11.

Figure 6.

CoV for VOT for mono-/p/ and multi-/p/ a function of age group. The vertical dash lines (- -) represent the values of the difference between mono-/p/ and multi-/p/.

Table 1
Age range and means by sex and age group for child participants (N = 70; 36 males, 34 females)

Age group	4 ~ 5	6 ~ 7	8 ~ 9	10 ~ 11	12 ~ 13	14 ~ 15	16 ~ 18
range	4.1 – 5.7	6.1 – 7.7	8.1 – 9.7	10.0 – 11.6	12.4 – 13.7	14.0 – 15.8	16.2
mean	4.7	6.8	9.1	10.8	13.2	14.8	17.1
Males	N = 7	N = 6	N = 6	N = 5	N = 5	N = 4	N = 3
range	4.1 – 5.5	6.0 – 7.2	8.1 – 9.6	10.2 – 11.5	12.6 – 13.7	14.0 – 15.6	16.2 – 18.4
mean	4.7	6.6	9.1	10.8	13.3	14.8	17.0
Females	N = 5	N = 4	N = 5	N = 7	N = 6	N = 4	N = 3
range	4.1 – 5.7	6.0 – 7.6	8.1 – 9.5	10.0 – 11.6	12.4 – 13.7	14.0 – 15.8	16.9 – 18.0
mean	4.9	6.7	8.9	10.8	13.1	14.7	17.4
Total	N = 12	N = 10	N = 11	N = 12	N = 11	N = 8	N = 6

Table 2

Mean scores and standard deviations for each standardized test by age group and sex.

Age group		4 ~ 5	6 ~ 7	8 ~ 9	10 ~ 11	12 ~ 13	14 ~ 15	16 ~ 18
Overall	PPVT	118.9 (10.3)	108.7 (9.2)	113.1 (11.1)	116.2 (8.6)	116.7 (9.9)	118.8 (10.3)	122.9 (7.6)
	EVT	118.6 (11.7)	105.6 (5.1)	110.8 (8.2)	114.3 (10.3)	110.6 (10.2)	117.6 (7.3)	122.7 (9.7)
Males	PPVT	117.9 (9.5)	112.7 (6.2)	110.0 (4.2)	116.3 (7.2)	116.2 (8.0)	117.8 (11.5)	119.0 (8.2)
	EVT	117.6 (10.9)	108.0 (4.8)	114.0 (4.5)	112.8 (10.6)	108.6 (8.1)	117.7 (5.1)	122.3 (6.0)
Females	PPVT	120.3 (10.9)	101.8 (10.0)	115.6 (12.1)	116.1 (9.8)	117.2 (9.9)	120.3 (9.6)	128.0 (7.7)
	EVT	119.8 (12.5)	101.5 (2.4)	108.2 (8.1)	113.6 (10.0)	111.7 (10.5)	117.5 (10.8)	123.3 (15.0)

Table 3

Mean, standard deviation (SD), CoV for VOT values (in ms) for mono-/p/ by sex and age group.

Age group	4-5		6-7		8-9		10-11		12-13		14-15		16-18		Adult	
	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
/p/																
Mean	77.0	74.1	72.6	68.1	68.4	88.7	64.4	69.8	61.5	75.8	66.5	68.4	58.9	68.3	59.4	64.2
SD	18.2	17.0	12.3	10.8	8.4	9.8	8.0	7.4	5.9	6.0	6.5	4.9	4.0	3.9	2.5	2.4
CoV (%)	23.6	22.9	16.9	15.9	12.3	11.1	12.4	10.6	9.6	7.9	9.8	7.2	7.0	5.7	4.3	3.8

* M = males; F = females

Table 4

Mean, standard deviation (SD), CoV for VOT values (in ms) for multi-/p/ and multi-/k/ by sex and age group.

Age group	4-5		6-7		8-9		10-11		12-13		14-15		16-18		Adult	
	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
/p/																
Mean	74.9	72.3	64.4	66.2	73.0	62.9	70.0	60.2	57.8	58.4	53.8	55.9	51.5	52.6	50.7	53.0
SD	25.5	23.4	19.4	18.1	16.4	12.1	15.5	12.8	10.4	9.7	7.4	7.6	4.6	4.5	3.1	3.0
CoV (%)	34.1	32.4	30.2	27.4	22.8	19.2	22.1	21.2	18.0	16.6	13.8	13.7	9.0	8.5	6.1	5.8
/k/																
Mean	89.2	92.7	75.0	77.3	80.2	75.0	73.8	65.4	64.8	66.0	64.3	65.3	63.2	65.8	62.0	64.6
SD	24.2	22.8	20.3	19.2	16.5	12.5	14.3	10.5	11.9	12.3	8.3	8.9	5.3	5.8	3.6	4.2
CoV (%)	27.1	24.6	27.1	24.9	20.6	16.6	19.4	16.1	18.4	18.7	12.9	13.6	8.4	8.8	5.8	6.5

* M = males; F = females