VOICE ONSET TIME PATTERNS IN 7-, 9- AND 11- YEAR OLD CHILDREN

S. P. Whiteside and J. Marshall

Department of Human Communication Sciences, University of Sheffield, UK.

ABSTRACT

Voice onset time (VOT) is a key temporal feature in spoken language. There is some evidence to suggest that there are sex differences in VOT patterns. The cause of these sex differences could be attributed to sexual dimorphism of the vocal apparatus. There is also some evidence to suggest that phonetic sex differences could also be attributed to learned stylistic and linguistic factors.

This study reports on an investigation into the VOT patterns for /p b t d/ in a group of thirty children aged 7 (n=10), 9 (n=10) and 11 (n=10) years, with equal numbers of girls (n=5) and boys (n=5) in each age group. Age and sex differences were examined for in the VOT data. Age, sex and age-by-sex interactions were found. The results are presented and discussed.

1. INTRODUCTION

Voice onset time (VOT) is a key temporal feature in spoken language. In the English language, stop consonants may be realised as voiced or voiceless. A number of phonemes are produced with the same manner and place of articulation but differ only along the dimension of voicing. Here voicing is the primary phonetic dimension used to distinguish these minimal pairs, for example with the plosive minimal pairs /p b/, /t d/ and /k g/, where /p t k/ are classed as voiceless and /b d g/, as voiced. Both physiological and acoustic differences have been identified for the two voicing categories. Acoustic cues which have been noted to differentiate voiceless and voiced plosives include: voice onset time [1, 2]; the amplitude of the plosive's burst [3, 4]; the frequency of the plosive's burst [5]; fundamental frequency characteristics [7]; formant frequency characteristics [5, 8]; closure duration [9]; and, the duration of the neighbouring vowel [10].

There appears to be a ‘many-to-one mapping’ of acoustic cues to voicing [11], that is, there are many acoustic cues signalling the one phonetic dimension of voicing. However, of these, voice onset time (VOT) is considered to be the primary cue for the voicing distinction as it is ‘the single most effective measure’ [12], identified in the perception and production of both word initial prevocalic stops [1, 2] and stops produced in spontaneous speech [13].

Interest in the value of VOT as a distinguishing feature for the two voicing categories began with the cross-language studies of Lisker and Abramson [1, 2]. The significance of VOT as a temporal feature can be seen by the number and range of studies which have since been carried out. VOT has been investigated with respect to, age differences [14, 15, 16] and sex differences [17, 18] for example.

VOT values exhibit intrinsic variations, relating to both their place of articulation and voicing category. VOT values increase as the place of articulation moves from anterior to more posterior, that is, alveolar stops will be produced with a longer VOT value than the corresponding labial stop [1, 2, 19]. Voiceless stops are produced with a wider distribution of VOT values than the voiced stop as the voiceless stop is the less stable member of the minimal pair. This is likely to be due to the increased complexity in timing of separately innervated supraglottal and glottal articulators which has been demonstrated through fibreoptic [20] and EMG studies [21].

It is important to consider the role of the development of motor control in the acquisition of adult-like VOT productions. In the production of a stop consonant, three physiological mechanisms are required, articulations to allow stop closure and release, articulations to isolate the nasal cavity at the velum and articulations to initiate vocal fold vibration [22]. In order for contrastive voicing production the speaker must be able to control these three separately innervated mechanisms. As with all other areas of motor development, control of laryngeal timing is continues to be refined during childhood.

There have been relatively few investigations into sex differences in VOT production. The results from adult studies are somewhat contradictory with reports of longer VOT values in males, longer VOTs in females, and in some cases no differences. Smith [23] reported longer VOTs for initial /d g/ in males. However, more studies have noted longer VOTs in females for the phonemes /t d/ [17], /p b/ [18] and across all phonemes [24]. Sweeting and Baken [25] noted no significant differences between males and females for /p b/. In general though, the results indicate that adult females produce longer VOT values than adult males. Underlying causes of this gender difference are unknown but suggestions include physiological or anatomical differences [17].

The overall aim of this study was to investigate both age and sex differences in the production of VOT of /p b t d/ in a fixed phonetic context across the ages of seven, nine and eleven years. A group of thirty children, with equal numbers of males and females in each age category participated in the study. The results will be presented and discussed within a developmental and sociophonetic framework.

2. METHOD

2.1. Subjects

Five males and five females within the age groups 7, 9 and 11 years were selected from a primary school in Sheffield. The ages of subjects in each group at the time of testing were within 9 months of each other. All subjects had a similar regional accent and have lived in Sheffield all their lives, spoke English as their first language and had no speech, language or hearing problems.

2.2. Speech stimuli

All subjects produced five repetitions of each of the phonemes /p b t d/. The data was elicited in a syllable initial position by producing the phrases: 'silver /pi/ ', 'silver /bi/ ', 'silver /ti/'and 'silver /di/ '. The phonemes were elicited in syllable initial prevocalic position for two reasons. Firstly, as this is the most easily measurable position for VOT and also to allow comparison with other studies where the data was also obtained within a similar context. Vowel context was standardised in all cases with the phoneme preceded by schwa /ə/ and followed by /i/. All phonemes were elicited in the same manner by asking the children to name silver (aluminium foil) shaped letters. Prior to recording the task was explained to each subject to familiarise them with the materials and procedure. A set of practice materials was used prior to the recording session to familiarise the subjects with the task. In cases of inaccurate articulation the subject was asked to repeat the test item to ensure five samples of each phoneme were available to be analysed, from each subject.

2.3. Recording and analysis

The speech data was recorded in a quiet room. All subjects were recorded on the same day to reduce the possibility of environmental differences altering the recording. Data was recorded directly onto a DAT (Digital Audio Tape) recorder (Sony, model TCD-D3). Speech samples from each subject were the digitized onto a Kay Computerised Lab (CSL) model 4300 using a sampling rate of 10 kHz. From this digital information, speech pressure waveforms and spectrograms were used in the voice-onset time (VOT) analysis. VOT was derived directly from speech waveforms by measuring the distance between the release of the plosive to the onset of voicing. The point of closure release was taken as the transient burst of high-intensity noise. Measures of VOT were validated using fast Fourier transform spectrograms. The onset of voicing was taken as the first vertical striations at a low frequency level. In cases where both the speech waveform and the spectrogram were referred to, the VOT measurement was taken from the same data source. In the cases where VOT was unclear e.g. plosives having an affricated release or the presence of background noise, the speech sample was discarded. The total amount of omitted data points was 10, which amounted to 1.67%.

To ensure consistency in the VOT measurements, a test of reliability was carried out. For this, 20% of the data was reanalysed by the same investigator, one week following the completion of the initial analyses. The data were selected by randomly choosing one subject from each of the six age-gender groups. A Pearson's product-moment correlation was used to calculate this intrarater reliability. A significant coefficient (r=0.99, p<0.001) demonstrated a high level of reliability.

3. RESULTS

The mean and standard deviation VOT values for the voiceless bilabial plosive /p/ are illustrated in Figure 1. When these data were subjected to a two way (by age and sex) ANOVA, there were no significant age differences. There were however, significant sex differences (F(1, 142)=21.0, p<.001) and a significant age-by-sex interaction (F(2, 142)=9.8, p<.001). The age-by-sex interaction was due to marked fall in the VOT values of the males between the ages of 9 and 11 years.

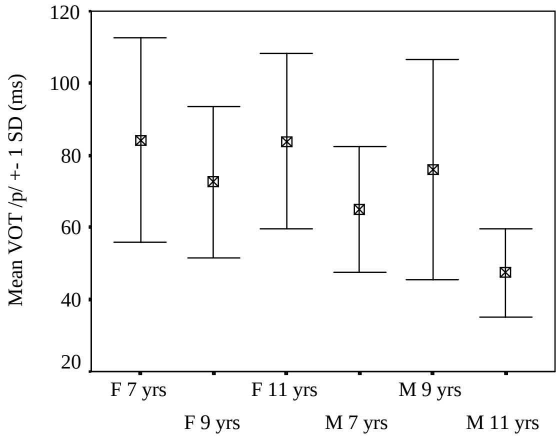

Figure 1: Mean and standard deviation VOT values (ms) for the voiceless bilabial plosive /p/ given by sex and age.

The mean and standard deviation VOT values for the voiced bilabial plosive /b/ are illustrated in Figure 2. When these data were subjected to a two way (by age and sex) ANOVA, there were no significant sex differences or age-by-sex interactions. There were however, significant age differences (F(2, 138)=5.4, p<.01) and a subsequent post-hoc test of Least Significant Difference indicated that there were significant age differences (p<.05) between the data of the 7 year olds and the 9 year olds, and between the data of the 7 year olds and the 11 year olds, with the 7 year olds having comparatively lower VOT values for both comparisons. No significant differences were found between the data of the 9- and 11-year olds.

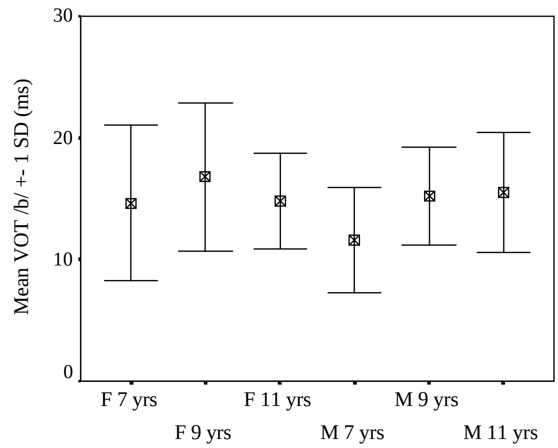

Figure 2: Mean and standard deviation VOT values (ms) for the voiced bilabial plosive /b/ given by sex and age.

The mean and standard deviation VOT values for the voiceless alveolar plosive /t/ are illustrated in Figure 3. A two way (by age and sex) ANOVA showed that there were no significant age or sex differences. There was however, a significant age-by-sex interaction ($F(2, 141)=5.5$, $p<.01$), with the males showing a dramatic fall in VOT values between the ages of 9 and 11 years, a pattern which was also found in the VOT data for /p/.

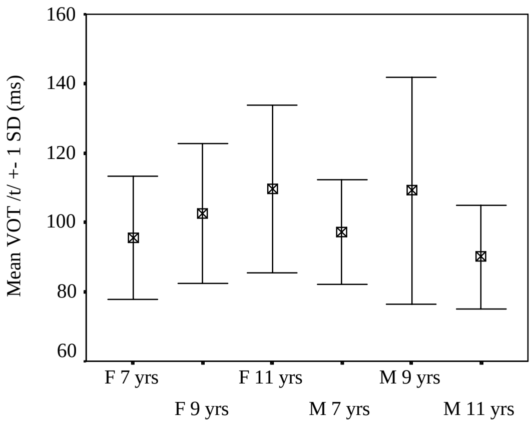

Figure 3: Mean and standard deviation VOT values (ms) for the voiceless alveolar plosive /t/ given by sex and age.

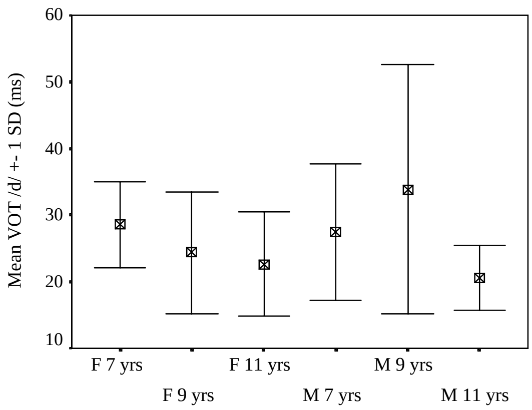

Figure 4: Mean and standard deviation VOT values (ms) for the voiced alveolar plosive /d/ given by sex and age.

The mean and standard deviation VOT values for the voiced alveolar plosive /d/ are illustrated in Figure 4. A two way (by age and sex) ANOVA indicated that there were no significant sex differences. There were however significant age differences ($F(2, 144)=7.4$, $p<.001$) and a post hoc test of Least Significant Difference indicated that there were significant differences ($p<.05$) between the data of the 11-year olds and the 9-year olds, and between the data of the 11-year olds and the 7-year olds. These significant age differences indicated that the VOT values were lower for the 11-year olds for both comparisons. In addition to this, a significant age-by-sex interaction ($F(2, 144)=4.7$, $p<.05$) was found for the /d/ VOT values, with the males showing a dramatic fall in VOT values between the ages of 9 and 11 years. This pattern replicates that which was found for the VOT data for /p/ (see Figure 2).

4. DISCUSSION

The results reported in this study showed that there were significant sex differences in the VOT data of /p/, with the female subjects having larger VOT values than their male peers. This finding replicates the findings of an adult study that was carried out on subjects from the same geographical region [18]. In addition to this, significant age differences were found for /b/ and /d/, where in the case of /b/, 7-year olds displayed smaller VOT values than either the 9- or 11-year olds. In the case of /d/ however, the 11-year olds had the lowest VOT values.

Perhaps most interesting in this study was the significant level of interaction that was noted between age and gender. Further analysis showed that this was due to significant age effects in males between nine and eleven years. Greater VOTs were produced by nine year old males for the phonemes /p t d/, a pattern which was not mirrored in females or males at any other age. Possible support for such dramatic changes in VOT production at age nine may come from other developmental acoustic studies. A study on fundamental frequency (Fo) by Hasek et al [27] reported the emergence of a significant difference between the average Fo of males and females at around age seven to eight years due to a rapid decline in Fo in males. These sudden changes in male Fo may be accompanied by changes in temporal features of speech, for example, in VOT. Possible explanations for the sudden Fo change include anatomical changes that occur at the onset of puberty, or maybe even sociophonetic influences. If the anatomical change viewpoint is taken, it seems feasible that sudden increases in vocal tract size would influence a number of acoustic parameters, for example, Fo, formant frequencies and VOT. The child might be attempting to stabilise their productions with a changing vocal tract configuration, which may lead to adjustments in timing control. With increased age, adaptations would be made, allowing control of timing to stabilise. These patterns in males at age nine years are not reflected in females. This may be due to the less dramatic vocal tract changes that occur at puberty in females. The effect of anatomical changes on decreasing VOT with age is proposed here. However, without further investigation into anatomical dimensions of children of each sex at each age, no interaction of these two variables can be assumed.

A factor not previously reviewed in the literature to account for decreasing VOT values between children and adults is sociophonetic influences. The gradual move towards adult values in a child may be due to increasing awareness of adult models, followed by gradual alterations in motor patterns to effect a change in VOT values. The results of the present study are unable to support any one theory as to causes of a decrease in VOT with age. However, the development of a more marked categorisation of the voiced/voiceless distinction for /p b/ and /t d/ in the data of the female subjects, could be interpreted as evidence for some contributory sociophonetic factors. It therefore proposed that an interaction of factors, which include motor development, anatomical changes, and sociophonetic factors are responsible for the VOT patterns reported here. Further investigation into anatomical variables and sociophonetic factors are therefore needed before any conclusive comments can be made.

5. ACKNOWLEDGMENTS

Our thanks go to St. Mary's School in Walkley, Sheffield.

6. REFERENCES

1. Lisker, L., and Abramson, A. S. "A cross-language study of voicing in initial stops: acoustic measurement", *Word,* 20, 1964, 384 - 422.

2. Lisker, L., and Abramson, A. S. "Some effects of context on voice onset time in English stops", *Language and Speech,* 10, 1967, 1 - 28.

3. Abramson, A. S., and Lisker, L. "Voice-timing perception in Spanish word-initial stops", *Journal of Phonetics,* 1, 1973, 1 - 8.

3. Repp, B. H. "Relative amplitude of aspiration noise as a cue for syllable-initial stop consonants", *Language and Speech,* 22, 1979, 173 - 189.

5. Jongman, A., and Blumstein S. E. "Acoustic properties for dental and alveolar stop consonants: a cross-language study", *Journal of Phonetics,* 13, 1985, 235 - 251.

5. Abramson, A. S., and Lisker, L. "Discriminabitity among the voicing continuum: cross-language tests", *Proceedings of the Sixth International Congress of Phonetic Sciences,* 1970, 569-573, Prague: Academia.

6. Haggard, M. P., Ambler, S., and Callow, M. "Pitch as a voicing cue", *Journal of the Acoustical Society of America,* 47, 1970, 613 - 617.

8. Stevens, K. N., and Klatt, D. H. "The role of formant transitions in the voiced-voiceless distinction for stops", *MIT Research Laboratory Progress Report,* 101, 1971, 188 - 197.

9. Lisker, L. "Closure duration and the intervocalic voiced-voiceless distinction in English", *Language,* 33, 1957, 42 - 49.

10. Klatt, D. H. "Interaction between two factors that influence vowel duration", *Journal of the Acoustical Society of America,* 54, 1973, 1102 - 1104.

11. Lisker, L., Liberman, A. M., Erickson, D. M., Dechovitz, D., and Mandler, R. "On pushing the voice onset time (VOT) boundary about", *Language and Speech,* 20, 1977, 209 - 216.

12. Lisker, L., and Abramson, A. S. "Distinctive features and laryngeal control", *Language,* 47, 1971, 767 - 785.

13. Abramson, A. S., and Lisker L. "Voice timing for stop classification in conversational English", *Proceedings of the XIIIth International Congress of Phonetic Sciences,* 4, 1995, 128 - 131. Stockholm.

14. Eguchi, S., and Hirsh, I. J. "Development of speech sounds in children", *Acta Otolaryngologica Suppl.* 257, 1969, 5 - 51.

15. Macken, M. S., and Barton, D. "The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants", *Journal of Child Language*, 7, 1979, 41 - 74.

16. Barton, D., and Macken, M. A. "An instrumental analysis of the voicing contrast in word-initial stops in the speech of four-year-old English-speaking children", *Language and Speech,* 23(2), 1980, 159 - 169.

17. Swartz, B. L. "Gender differences in voice onset time", *Perceptual and Motor Skills,* 75, 1992, 983 - 992.

18. Whiteside, S. P., and Irving, C. J. "Speakers' sex differences in voice onset time: some preliminary findings", *Perceptual and Motor Skills,* 85, 1997, 459-463.

19. Fischer-Jørgenson, E. "Sound duration and place of articulation", *Zeitschrift für Phonetiksprachwissenschaft und Kommunikationsforschung,* 17, 1964, 175 - 207.

20. Sawashima, M., Abramson, A. S., Cooper, F. S., and Lisker, L. "Observing laryngeal adjustments during running speech by use of a fibreoptics system", *Phonetica,* 22, 1970, 193 - 201.

21. Hirose, H., and Gay, T. "The activity of the intrinsic laryngeal muscles in voicing control", *Phonetica,* 25, 1972, 140 - 164.

22. Kewley-Port, D., and Preston, M. S. "Early apical stop production: a voice onset time analysis", *Journal of Phonetics,* 2, 1974, 194 - 210.

23. Smith, B. L. "Effects of place of articulation and vowel environment on voiced stop consonant production", *Glossa,* 12, 1978, 163 - 175.

24. Ryalls, J., Zipprer, A., and Baldauff, P. "A preliminary investigation of the effects of gender and race on voice onset time", *Journal of Speech, Language and Hearing Research,* 40, 1997, 642 - 645.

25. Sweeting, P. M., and Baken, R. J. "Voice onset time in a normal-aged population", *Journal of Speech and Hearing Research,* 25, 1982, 129-134.

26. Mattingly, I. G. "Speaker variation and vocal tract size", *Journal of the Acoustical Society of America* 39, 1966, S1219A.

27. Hasek, C. S., Singh, S., & Murry, T. Acoustic attributes of preadolescent voices, *Journal of the Acoustical Society of America,* 68 1262-1265 (1980).