Article:

Whiteside, S.P., Dobbin, R. and Henry, L. (2003) Patterns of variability in voice onset time: a developmental study of motor speech skills in humans. Neuroscience Letters, 347 (1). pp. 29-32. ISSN 0304-3940

https://doi.org/10.1016/S0304-3940(03)00598-6

Published paper
Whiteside, S.P., Dobbin, R. and Henry, L. (2003) *Patterns of variability in voice onset time: a developmental study of motor speech skills in humans.* Neuroscience Letters, 347 (1). pp. 29-32.

Neuroscience Letters 347 (2003) 29-32.
Accepted 15 May 2003.

Patterns of variability in voice onset time: a developmental study of motor speech skills

Sandra P. Whiteside[a)], Rachel Dobbin[b)] and Luisa Henry[b)]

[a)]Department of Human Communication Sciences, University of Sheffield, Sheffield S10 2TA, United Kingdom. Tel: +44 (0)114 2222447, Fax: +44 (0)114 2730547.

Email: s.whiteside@sheffield.ac.uk. *Corresponding Author*.

[b)]Present address: Speech and Language Therapy Department, Hull and East Riding Community NHS Trust, Victoria House, Park St, Hull HU2 8TD.

Neuroscience Letters 347 (2003) 29-32.
Accepted 15 May 2003.

Abstract

This study investigated the developmental patterns of variability in the speech parameter voice onset time (VOT) in forty six children. Five groups of children participated in the study as follows: i) Group 1 - aged 5 years 8 months (n=6); ii) Group 2 - 7 years 10 months (n=10); iii) Group 3 - 9 years 10 months (n=10); iv) Group 4 - 11 years 10 months (n=10), and v) Group 5 - 13 years 2 months (n=10). Coefficient of variation (COV) values were examined for the VOT values of both "voiceless" (/p t k/) and "voiced" (/b d g/) plosives to determine patterns of variability. Significant effects of age were revealed for both the voiceless and voiced plosives, and levels of variability leveled off for Group 4. The data suggest that although variability in VOT decreases with age, the presence of residual variability may be a prerequisite for the further refinement of motor speech skills.

Keywords: Voice onset time, variability, speech development, motor speech skills, pre-adolescence, adolescence, phonetic context

Voice onset time (VOT) is a salient acoustic temporal cue which signals the "voiced" or "voiceless" status of a plosive in English [6]. In CV/CVC words, where C is a plosive and V is a vowel, VOT is defined as the interval between the release of an oral constriction (acoustically marked by a noise burst) and the onset of voicing (acoustically marked by the onset of periodic voicing). The VOT patterns of "voiced" plosives are characterised by either a voicing lead, no voicing lead or a short voicing lag. Conversely, VOT patterns of "voiceless" plosives are characterised by a longer voicing lag compared to their "voiced" counterparts [6,11,15]. Patterns of VOT also vary according to phonetic context of the speech material being examined [11,12].

Voice onset time (VOT) plays a key role in both speech production and perception and has been the focus of a number of investigations in speech perception [11,14,16,17], and speech production [2,3,4,6,7,8,9,15]. Furthermore, this temporal parameter has also been investigated in a range of motor speech disorders as it provides information on disruptions to the coordination between laryngeal and supralaryngeal mechanisms which govern the integrity of this speech parameter [1,5], or how individuals with motor speech disorders compensate for their underlying motor deficits [1].

The examination of VOT patterns in children provides us with information on how this temporal parameter of speech is acquired and mastered as a function of development [2,3,9,15]. Furthermore, because VOT is a microtemporal component of speech, patterns of variability in VOT during preadolescence and adolescence give us some indication of how this aspect of motor speech skill is continually being refined and mastered by school age children spanning a wide age range. The aim of this study was therefore to investigate the patterns of variability in the VOT patterns of voiceless and voiced plosives as a function of both age in forty six male and female preadolescent and adolescent children aged between 5;10 years and 13;2 years, and assess the relevance and importance of variability in speech production within a developmental framework of the acquisition and mastery of motor speech behavior.

Forty six preadolescent and adolescent children participated in the study. The children fell into 5 age groups: i) Group 1 ("5-year olds": Mean = 5 years; 8 months, SD = 3 months, n=6); ii) Group 2 ("7-year olds": Mean = 7 years;10 months, SD = 2.6 months, n=10); iii) Group 3 ("9-year olds": Mean = 9 years;10 months, SD = 2.7 months, n=10); iv) Group 4 ("11-year olds": Mean = 11 years;10 months, SD = 2.9 months, n=10); and v) Group 5 ("13-year olds": Mean = 13 years;2 months, SD = 3.2 months, n=10). Equal numbers of boy and girls participated in each age group. All subjects: i) were from the same geographical region (Sheffield), and spoke with a similar regional accent; ii) had lived in the same speech community all their lives; iii) had age-appropriate intelligence levels as judged by their class teacher; iv) were monolingual speakers of English; and v) had no speech, language or hearing problems; vi) volunteered to participate in the study with parental consent.

In English /p t k/ are classed as "voiceless" plosives, and /b d g/, as their respective voiced cognates. Each subject produced five repetitions of each of the following plosive consonants /p b t d k g/ in a syllable initial position in a target word within the frame "Say ______ again". The target words used were *pea, bee, tea, Dee, key, ghee, part, Bart, tart, dart, card,* and *guard*. This gave a possible total of 60 speech samples (5 repetitions x 12 target words) for each subject. The phrases and target plosives were elicited using a repetition task. In order to obtain the full range of voiced and voiceless plosives from the children, a repetition task was chosen in preference to either a reading task, or a picture naming task for a number of reasons. Firstly, because the study included very young children with a range of reading abilities, this would have more than likely, had an effect on the children's VOT values. Secondly, it would not have been possible to elicit correct renditions of the word "ghee" (/gi/) in a reading paradigm, as the lexical item in question is both borrowed (Hindi), occurs infrequently in both spoken and written English, and would probably be novel to most school age children. Prior to recording, the task was explained to subjects in order to familiarize them with the procedure of the repetition paradigm. Where possible, five samples of each plosive were elicited from every child, therefore giving a possible total of 60 samples per subject, and a possible overall total of 2760 samples.

The speech data were recorded in a quiet room. Data were recorded directly onto a DAT (Digital Audio Tape) recorder (Sony, model TCD-D3). Subsequently, speech samples from each subject were digitized onto a Kay Elemetrics Computerized Lab (CSL) model 4300 using a sampling rate of 16 kHz. From this digital information, sound pressure waveforms and wideband (146 Hz) FFT spectrograms were generated and displayed. VOT measurements were made directly from the spectrograms by measuring the distance between the release of the plosive and the onset of voicing (marked by the first visible sign of low frequency periodic acoustic activity in the spectrograms). The point of closure release was taken as the transient burst of the plosive's release. In the cases where measures of VOT needed validation, sound pressure waveforms were used. In the cases where both the speech waveform and the spectrogram were referred to for validation, the VOT measurement was taken from the same data source. In the cases where VOT was unclear (e.g. plosives being released with affrication or the presence of background noise), the speech sample was discarded. All VOT measurements were taken in milliseconds.

To ensure consistency in the VOT measurements, a test of inter-rater reliability was carried out. The inter-reliability measures were conducted by randomly selecting one subject from each of the 5 age groups for reanalysis by a second rater. A Pearson’s product-moment correlation was used to calculate the level of inter-rater reliability. A significant correlation coefficient (r=0.978, p<0.0001) demonstrated that there was a high level of inter-rater reliability.

The means and standard deviations for the VOT were calculated for each individual subject to assess patterns of intrasubject variability as a function of age for the voiceless plosives (/p t k/), and the voiced plosives (/b d g/). These values were then used to calculate coefficient of variance values (COV) for each subject as follows: (SD/Mean) * 100, in order to examine the patterns of variability in the VOT values of the children's speech samples. COV was chosen as the index of variability in order to ensure that standard deviations could be expressed as a function of the mean, and therefore control for higher standard deviations which may have been the product of larger mean values. The COV values for the voiceless plosives and voiced plosives were analyzed separately using a one way (age) analysis of variance (ANOVA).

The means and standard deviation values for the COV values for the voiceless plosives are illustrated in Fig. 1 by age group. A one way (age) ANOVA for the voiceless plosives revealed a significant main effect of age ($F(4, 271) = 11.433$, $P < .0001$). Multiple comparisons with Bonferroni adjustment were carried out for the factors of age and plosive. Multiple comparisons for age revealed that the 5-, 7- and 9-year olds had significantly larger COV values than the 11- and 13-year olds ($P < .05$). The smallest standard deviations for COV are also observed for the 11- and 13-year olds (Fig. 1). The means and standard deviation values for the COV values for the voiced plosives are illustrated in Fig. 2 by age group. A one way (age) ANOVA for the voiced plosives revealed a significant main effect of age ($F(4, 271) = 11.04$, $P < .0001$) and plosive ($F(2,261) = 7.078$, $P < .002$). Multiple comparisons with Bonferroni adjustment were carried out for the factors of age and plosive. Multiple comparisons for age revealed that the 5-, 7-, and 9-year olds had significantly ($P < .05$) larger COV values than the 11-

Neuroscience Letters 347 (2003) 29-32.
Accepted 15 May 2003.

and 13-year olds. The smallest standard deviations for COV are also observed for the 11- and 13-year olds (Fig. 2).

The results of this study show that variability in VOT decreased between the data for the 5-year olds and those of the 11-year olds. Furthermore, this pattern was found for both the voiceless and voiced plosives. These data corroborate earlier findings on VOT for /b p t/ [2]. The developmental patterns of variability in the VOT data can be interpreted as evidence for increased levels of stability in speech output as a function of the maturing motor speech skills. They also reflect the increased levels of proficiency in the motor learning of preadolescent children, who continue to develop, refine, control and adapt their motor speech skills within a biological system which is constantly undergoing anatomical and physiological maturation. The lack of significant age differences between the 11- and 13-year olds suggests that levels of variability in VOT may begin to level off at this stage within the human lifespan. Additional data is necessary to confirm this suggestion in light of evidence which shows that the leveling of variability occurs earlier for some children; around the age of 7 or 8 at which point adult levels of variability are achieved [2]. Furthermore, how much these age-related variations are a consequence of the differences in the methods used to elicit the speech data (e.g. read speech [2] or via repetition as used in this study), remains open to question.

If we adopt the view that the variability of VOT reaches adult levels, and stabilizes around the age of 11 years, we are left with evidence for residual levels of variability (see Figs. 1 & 2), which may remain at this level, at least for the earlier part of the adult lifespan [13], after which we may expect changes in variability to occur as a response to further developmental biological changes. These residual levels of variability require further investigation in older children (postadolescent) and adults. However, their presence may be indicative of the requisite levels of instability which are necessary for the organization of motor behavior, the development of motor schemas, skilled actions and their adaptive control [10,13]. Speech motor skills are no exception to this, and the speech motor system needs to allow some level of variability and motor flexibility in order to accommodate, and adapt to some of the factors which influence speech production. These include the following, which are but a few examples: differences in speech style (e.g. casual and formal), the level of novelty of an utterance, lexical/utterance frequency, different speech rates (e.g. fast or slow), different types of speech material which involve different underlying processing mechanisms (e.g. read speech, samples elicited via repetition, or spontaneous conversation), and phonetic context (e.g. the nature of plosive or vowel context).

In summary, the results of this study indicated a developmental trend in the variability patterns in VOT, which was characterized by a decrease in variability between age 5;8 and 11;10 years. These patterns of declining variability are indicative of maturing motor speech skills as children approach adolescence, while the residual levels of variability could be interpreted as a prerequisite for the organization of motor speech behavior, and the development of motor speech schemas.

[1] Auzou, P., Özsancak, C., Morris, R. J., Jan, M., Eustache, F, and Hannequin, D., Voice onset time in aphasia, apraxia of speech and dysarthria: a review. Clin. Ling. Phonet., 14, (2000) 131-150.

[2] Eguchi, S., and Hirsh, I.J., Development of speech sounds in children. Acta Oto-laryngol., Suppl. (1969) 257.

[3] Kewley-Port, D., and Preston, M.S., Early apical stop production: a voice onset time analysis. J. Phon., 2, (1974) 195 - 210.

[4] Koenig, L. L., Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. J. Sp. Lang. Hear. Res., 43, (2000) 1211-1228.

[5] Kurowski. K. M., Blumstein, S. E., and Mathison, H., Consonant and vowel production of right hemisphere patients. Brain Lang., 63, (1998) 276-300.

[6] Lisker, L., and Abramson, A.S., A cross-language study of voicing in initial stops: acoustical measurements. Word, 20, (1964) 384 - 422.

[7] Löfqvist, A., Acoustic and aerodynamic effects of interarticulator timing in voiceless consonants. Lang. Speech, 35, (1992) 15-28.

[8] Löfqvist, A., Koenig, L. L., and McGowan, R. S., Vocal tract aerodynamics in /aCa/ utterances: measurements. Speech Commun., 16, (1995) 49-66.

[9] 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. J. Child Lang., 7, (1979) 41 - 74.

[10] Manoel, E., de J, and Connolly, K. J., Variability and the development of skilled actions, Int. J Psychophysiol., 19 (1995) 129-147.

[11] Nearey, T. M., and Rochet, B. L., Effects of place of articulation and vowel context on VOT production and perception for French and English stops, J. Int. Phonetic Assoc. 24, (1994) 1-19.

[12] Port, R. F., and Rotunno, R. Relation between voice-onset time and vowel duration. J. Acoust. Soc. Am., 66, (1979) 654 - 662.

[13] Schmidt, R. A., Motor Control and Learning: A Behavioral Emphasis, 2nd Edition, Human Kinetics, Champaign IL, 1988.

[14] Simos, P. G., Diehl, R. L., Breier, J. I., Molis, M. R., Zouridakis, G., and Papanicolaou, A. C., MEG correlates of categorical perception of a voice onset time continuum in humans. Cog. Brain Res., 7, (1998) 215-219.

[15] Whiteside, S. P., and Marshall, J., Developmental trends in voice onset time: some evidence for sex differences. Phonetica, 58, (2001) 196-210.

[16] Zlatin, M.A., and Koenigsknecht, R.A., Development of the voicing contrast: perception of stop consonants. J. Speech Hear. Res., 18, (1975) 541-553.

[17] Zlatin, M.A., and Koenigsknecht, R.A., Development of the voicing contrast: a comparison of voice onset time in stop perception and production. J. Speech Hear. Res., 19, (1976) 93 - 111.

Neuroscience Letters 347 (2003) 29-32.
Accepted 15 May 2003.

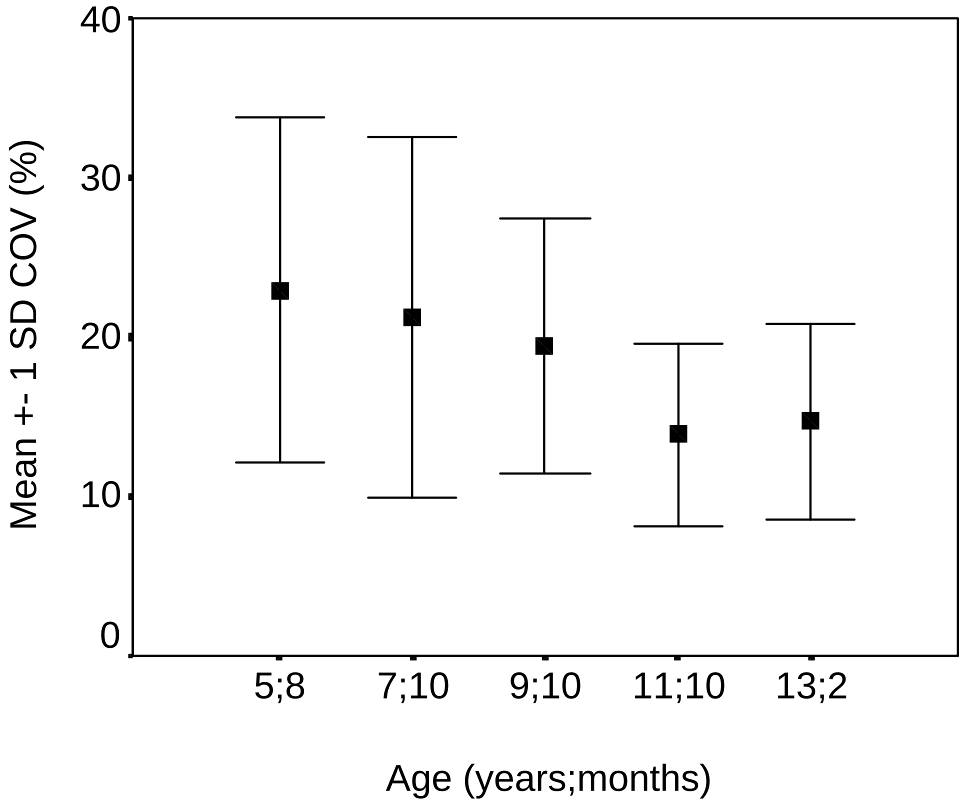

Fig.1. Mean COV values (%) for the voiceless plosives by age group. Error bars indicate +- 1 SD values.

Neuroscience Letters 347 (2003) 29-32.
Accepted 15 May 2003.

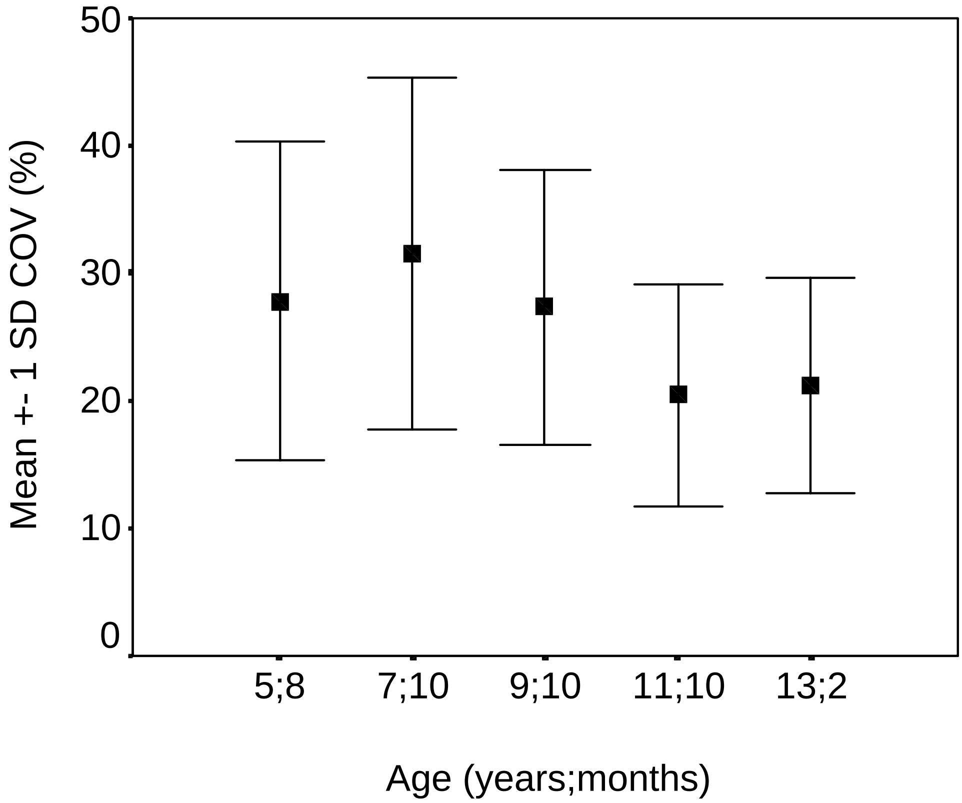

Fig.2. Mean COV values (%) for the voiceless plosives by age group. Error bars indicate +- 1 SD values.