Voice onset time in Swedish children with phonological impairment

Inger Lundeborg Hammarström, Elin Nordin, Marie Zeipel-Stjerna and Anita Mcallister

Linköping University Post Print

N.B.: When citing this work, cite the original article.

This is an electronic version of an article published in:
Logopedics, Phoniatrics, Vocolgy is available online at informaworld™:
http://dx.doi.org/10.3109/14015439.2014.934276
Copyright: Taylor & Francis: STM, Behavioural Science and Public Health Titles
http://www.tandf.co.uk/journals/default.asp
Postprint available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-125843
Voice onset time in Swedish children with phonological impairment

Inger Lundeborg¹, Elin Nordin¹, Marie Zeipel-Stjerna¹, & Anita McAllister¹

Department of clinical and experimental medicine, ¹Division of speech and language pathology, Linköping University, Sweden

Corresponding author: Inger Lundeborg, Department of clinical and experimental medicine, Division of Speech and Language Pathology, Linköping University, S-501 85 Linköping, Sweden; Tel. +4613286793, e-mail: Inger.lundeborghammarstrom@liu.se

Key words

Child, Speech sound disorders, Voice onset time
**Abstract**

Mastering spatial and temporal coordination in speech production is a challenge for children. Voice onset time reflects timing in speech. The objective was to study VOT in Swedish children with a diagnosed phonological impairment and compare results with normative data. Thus 38 children 4-11 years, in three age groups were audio-recorded when producing minimal pairs with the plosives /p b t d k ɡ/. Waveforms and spectrograms were analysed. Results show that children with phonological impairment produced plosives with deviant VOT-values and greater variability compared to normative data. No developmental trend was seen with increasing age. Also, no relationship was found between VOT-values and degree of impairment measured by percentage phonemes correct. Findings were interpreted as displaying motor coordination difficulties.
Introduction

Human speech is a complex motor behaviour and the developing child needs many years to establish adult mastery of the phonemic contrasts in a given language. However, without a communicative intent intermediated by language symbols, speech has no meaning. The process of phonological acquisition in children has been described as a process of organizing and learning the mappings between the auditory speech input and the gestural control of speech output (1).

Phonological ability also improves through increased capacity to produce adult sounds and combine them into more complex phonological structures (2). Motor speech control has traditionally been regarded as separated from phonological operations, but in, for example, gestural phonology issues in phonology and articulation are addressed simultaneously (3). A relationship between oral motor control and speech and language development has been reported by several authors (4, 5, 6). However, to which extent oral motor skills influence language development requires further research.

Children with speech problems form a large, but heterogeneous group and a multitude of diagnostic terms has been used to describe speech disorders without obvious causation. One of the ongoing debates within the study of developmental speech and language disorders concerns whether children’s impaired speech is caused by incomplete linguistic knowledge, processing deficits or limited motor speech control, (7, 4). The multidimensional requirements in speech production continue to elude researchers and present problems for clinicians. Regarding the aspect of motor control different degrees of motor coordination and control are required for production of different speech sounds. Plosives are among those which demands close co-ordination of breathing and the laryngeal and the supra-laryngeal structures. In clinical practice auditory- perceptual evaluations form the basis for assessments of the voicing
aspect of plosives. However, insights of error patterns that go beyond perceptual judgements can be provided by acoustic analysis (8).

Voice onset time (VOT) is considered the most reliable acoustic cue for distinguishing between voiceless and voiced plosives (9, 10, 11) and is defined as the time interval between the release of the oral closure and the onset of voicing measured in milliseconds. Several studies have shown that VOT values are affected by speaking rate. When the speaking rate is lower, both VOT values for voiceless stop consonants and the duration of the following vowel increase (12). Typical Swedish adult values for the voiceless plosives range between 49 and 78 milliseconds and between -91 and -61 milliseconds for the voiced plosives (13). Since voice onset time is considered to reflect the co-ordination between articulatory gestures and phonation (14), VOT measures could be included in cases of suspected verbal dyspraxia and other speech motor disorders.

In two studies of typically developed Swedish children aged 3-5 years and 8-11 years, a clear developmental trend was found in the studied age-groups, both regarding length of VOT for the voiceless plosives and for the occurrence of prevoicing in voiced plosives. The studied children developed adult-like values somewhere between 9 and 10 years of age (15, 13) see table I and figure 1. Studies of VOT in English-speaking children with phonological impairment have shown atypical VOT-patterns (16). However, no study of VOT in Swedish children with speech sound disorders has been conducted. The aim of the present study is to investigate voice onset time in Swedish speaking children with phonological impairment and to compare data with the earlier obtained values from typically developed children.

Insert table I and figure 1 about here

Material and methods

Subjects
A total of 38 children, ten girls and 28 boys with phonological impairment, aged 4.2 to 11.6 years (mean age 6y and 10m,) participated in the study. The children were divided into three age groups according to Swedish pre-school and school organization; group 1, pre-school children (4.2- 6.0 years old, mean age 4y and 11m, n=18), group 2 primary school children (6.1-9.0 years old, mean age 7y and 8m n=13) and group 3 late primary and intermediate
school children (9.1-11.6, mean age 10y and 4 m, \(n=7\)). The children were recruited from speech pathology clinics and language pre-school and school units for children with speech and language impairments in six towns in the south-east and mid-south of Sweden. All children were diagnosed with phonological impairment but able to produce plosives. They all had Swedish as their first language and age appropriate intelligence levels as judged by their SLP’s.

Assessment procedure

All children participated in a small, pre-test with all Swedish plosives, to ensure that they had plosives in their phoneme inventory. A Swedish phonology test including 72 items was performed (17). Auditory discrimination was tested using pictures of nine minimal pairs with the different plosives, three pairs for each place of articulation. For the VOT measurements, picture naming using the same material as in the previous study of typically developed Swedish children (15, 13) with six minimal pairs, pil-bil (arrow-car), tennis-Dennis (tennis-Dennis) and kula-gula (ball-yellow) repeated three times. The procedure was audio-recorded using a Marantz PMD 660 Professional Recorder with a sampling frequency of 44.1 kHz and an Audio-Technica microphone on a stand directed to the speaker’s mouth at a distance of approximately 30 cm (11.8 inches).

Measures and analysis

Acoustic measures

The second production of each word was analysed using displays of waveforms and spectrograms generated through the Praat software (http:www.fon.hum.uva.nl/praat/, version 5.1.31, Paul Boersma and David Weenik, Phonetic Sciences Department, University of Amsterdam). The VOT measurements were obtained according to the method described by Lisker and Abrahamson (18) by measuring the distance between the transient burst of the stop to the first positive zero-crossing, introducing periodicity of the following vowel. VOT is expressed in milliseconds (ms), see Figure 2. The spectrograms were used for additional
support of the analysis and the measurements taken from the release of the plosive (marked by a transient in the spectrogram) to the onset of low-frequency periodical activity.

Insert figure 2 about here

**Phonological analyses and scoring**

The recordings of the phonological test were transcribed phonetically by the second and third author together using the transcription conventions of the International Phonetic Association (19). Based on the transcriptions calculations of Percentage of Phonemes Correct (20) were made.

**Reliability**

A random selection of four recordings were re-analyzed and intra-rater reliability was calculated using Cronbach’s alpha and found to be 0.97 for VOT and 1 for PPC.

**Statistical analysis**

Data were expressed with descriptive statistics for the demographics. Comparisons of VOT-values between the age groups were made using one-way analyse of variance (ANOVA). Comparisons between VOT-data and earlier obtained norm data were made using Student’s t-test. \( P \)-values <0.05 were considered statistically significant. A correlation analysis of the PPC values and VOT was performed using Spearman’s rank correlation coefficient. All VOT-measures were performed by two of the authors (M.A. and E.N.) and the values were mutually decided. A correlation analysis between the VOT values and the duration of the plosive syllables was made using the Pearson product-moment correlation coefficient. Comparisons between the age groups regarding syllable duration and VOT values in relation to the syllable duration were made using Student’s t-test.

All Statistical analyses were performed using SPSS© Windows version 19.0 (SPSS Inc. Chicago, Illinois).

VOT-values above or below one standard deviation from the mean value for the typically developed children in Larsson and Wiman, (15) and Lundeborg, et al.(13) were considered as deviant.
The study was carried out in accordance with the ethical principles for medical research of the Helsinki Declaration as revised in 2008 (21).

**Results**

All studied children had VOT-values above or below one standard deviation from the mean value for typically developed children on at least one of the plosives, see figure 3. However no general pattern could be observed and no trend towards less deviant productions with increasing age was observed.

*Insert figure 3 about here*

The mean VOT-values and values normalized with respect to the duration of the syllable for all plosives and age groups are shown in table II and III. No statistically significant differences between age groups or sex were found.

*Insert table II and III about here*

**The voiceless plosives**

All three age-groups, with the exception of one child and the /t/-sound, had positive VOT values (with voicing lag) when producing the voiceless plosives. An increase of values as the place of articulation moved posteriorly (p<t<k) was found in the youngest age group, but not in the two older age groups (t<p<k). When comparing the voiceless plosives in the study group as a whole, the differences between /t/ and /k/ and between /p/ and /k/ were significant (p<.05) with longer voicing lag for /k/.

**The voiced plosives**
Individuals with prevoicing were found in all age groups but there was no clear trend with more values with increasing age. However, the proportion of children with prevoicing increased with increasing age for /b/ and /d/ but not for /ɡ/, see figure 4. Prevoicing for all voiced plosives was only found in one single participant in age group 2 (6.1-9.0 years). An increase of values as the place of articulation moved posteriorly was found in the oldest age group, whereas the middle group had values d<ɡ<b and the youngest group had d<b<ɡ.

Insert figure 4 about here

Syllable duration and VOT

When correlating the VOT values with the duration of the syllables moderate positive correlations were found both for the voiceless and the voiced plosives (0.46 and 0.40). No significant differences were found between the age groups regarding the relation between VOT and syllable duration.

Regarding duration of the syllables where VOT was measured, there were no differences between the one- and two syllables words. When comparing the syllable duration between the age groups, a significant difference (p<0.05) was found between the youngest and the middle age group regarding the /d/-syllable and between the youngest and oldest age group regarding the /k/-syllable. No other differences were found.

Auditory discrimination

Total number of errors on the auditory discrimination task was zero or one error for most children (n=23). However only four children made more than one error within the same category, two each in the two youngest age groups. No significant correlation was found between number of errors and number of deviant VOT.

Phoneme production

Mean percentage of phonemes correct (PPC) were 75.7, 83.7 and 97.6 for the youngest to the olders group respectively, see table IV.

When correlating auditory discrimination errors with PPC, a significant negative correlation was found (r=-0.54, p<0.001)

Insert table IV about here
VOT and phonological ability

Apart from a significant correlation ($r = -.47$) between a low PPC and a high positive VOT-value for /d/, there was no relationship between VOT-values and PPC.

Discussion

The result of this study demonstrate that children with speech sound disorders have a wide range of VOT-values for all the plosives and that degree of speech impediment according to PPC did not seem to have any relationship with the development of VOT apart from the odd correlation between a low PPC and a high positive VOT for /d/. The magnitude of the variability of obtained values widely exceeds the variability found in typically developed children. This is in line with the findings by Tyler and Saxman in 1991 (21). No child had VOT values within one standard deviation from the mean value for the typically developed children in Larsson and Wiman (15) and Lundeborg, et al. (13) for all plosives. Furthermore most children had deviant values for both voiceless and voiced plosives. Thus, there was a great scatter both regarding over- or undershoot and number of deviant VOT-values clearly indicating atypical VOT-patterns. Also, no trend towards less deviant productions with increasing age was seen. In typically developing children adult-like VOT values were established between 9 and 10 years (10, 15, 13).

Only one of the studied children had prevoicing throughout all voiced plosives and contrary to findings in typically developing children, no increase in prevoicing with increasing age was observed. Furthermore, in typically developed children, when prevoicing was well established it was present for all plosives in most cases (13).

Another observed difference to previous VOT-values in typically developing children, was that the studied children did not exhibit increased VOT-values as place of articulation moved posteriorly for voiceless plosives. All these differences could be interpreted as a sign of less
mature or deviant speech motor control in a subgroup of children with speech sound disorders.

Several studies have shown that speech rate in children increases with age and is higher in longer utterances (22). No such trend could be seen in this material in the measurement of the syllables. This could also be interpreted as the result of an immature speech motor control.

The number of syllables in a word is regarded as an important factor for consonant duration (23). However, in the present study no significant differences between the duration of the initial syllables in one- and two syllable words were found.

Apart from a significant correlation (r = -.47) between a low PPC and a high positive VOT-value for /d/, there was no relationship between VOT-values and PPC. This shows that VOT deviances are not perceived perceptually. Perhaps this is due to the fact that the perceptual evaluations were made by native listeners with experience from working with children with speech- and language difficulties.

Another possible explanation is that measuring only VOT is insufficient. Maybe also other aspects of speech motor control than VOT must be included when studying children with speech sound disorders. A study also including electropalatography could help to further elucidate the articulatory patterns in these children. There is also a need for establishing more low- tech clinical tools to be able to distinguish between children with more motor based, as compared to language based speech sound disorders, since the clinical hypothesis regarding etiology is the basis for choice of intervention. One such tool may be auditory discrimination. Several studies have shown a relationship between perception and production (24, 25). In the present study no such clear relationship was found. No significant correlation was found between number of errors on the auditory discrimination task and the number of deviant VOT. However, auditory discrimination errors correlated with PPC (r = -0.54, p<0.001). This contrast may reflect a relationship or simply the different measures used, VOT and perceptual assessment of production by native listeners.

Another aspect regarding intervention is that both effectiveness and efficacy needs to be considered. The oldest group of children participating in the present study was 9 years old and older. However these children still had difficulties with speech production despite extensive intervention. It is possible that the coordination and timing deficit revealed in the deviant VOT-values have not been addressed in therapy.
To be able to make comparisons, the same methodology as in the studies of typically developed children was used. Also, the children were grouped with respect to the developmental stages for the acquisition of VOT summarized by Whiteside and Marshall in 2001 (10) to enable further comparison.

It would clearly have been interesting if the children in the present study had been asked to repeat the word several times to be able to make multiple measures to study individual consistency.

**Conclusion**

Children diagnosed with phonological impairment have deviant VOT-patterns with much more variability compared to typically developed children. However, the variability could not be explained by differences in degree of the speech difficulties. These findings underlines the importance of finding reliable tools for identifying children with incomplete linguistic knowledge and/or processing deficits from those with limited motor speech control.

**Declaration of interests**

The authors report no declarations of interest.

**References**


**Tables and figure captions**

**Figure 1.** Incidence of prevoicing in 3-5 years (Larsson and Wiman 2010) and 8-11 years and adults (Lundeborg et al, 2012)

**Figure 2.** Example of how the VOT values for the voiceless (to the left) and the voiced plosives (to the right) were defined and measured. The arrows in the voiceless example mark the plosive phase, the arrows in the voiced example indicate the prevoicing occurring during the plosive phase.

**Figure 3.** Proportion of participating children with deviant VOT (defined as 1 standard deviation below or above the mean for typically developed children in corresponding ages.

**Figure 4.** Percentage of individuals across age groups having prevoicing when producing /b d g/.

**Table I.** Mean (standard deviation) for VOT expressed in milliseconds in Swedish typically developed children aged 3-5 years (Larsson and Wiman, 2010) and 8-11 years + adults (Lundeborg et. al. 2012).

**Table II.** Mean (standard deviation) and range for VOT expressed in milliseconds for the children with speech sound disorders by age group.

**Table III.** Mean and standard deviation for VOT values in relation to syllable duration for the children with speech sound disorders by age group.

**Table IV.** Range of and mean (SD) percent phonemes correct (PPC) for the tree age groups.
### Table I

Mean (standard deviation) for VOT expressed in milliseconds in Swedish typically developed children aged 3-5 years (Larsson and Wiman, 2010) and 8-11 years + adults (Lundeborg et al. 2012)

<table>
<thead>
<tr>
<th></th>
<th>3 years, n=15</th>
<th>4 years, n=37</th>
<th>5 years, n=31</th>
<th>8 years, n=32</th>
<th>9 years, n=39</th>
<th>10 years, n=43</th>
<th>11 years, n=36</th>
<th>Adults, n=36</th>
</tr>
</thead>
<tbody>
<tr>
<td>/P/</td>
<td>83.6 (52.8)</td>
<td>72.9 (31.2)</td>
<td>63.2 (21.6)</td>
<td>49.8 (12.7)</td>
<td>50.4 (14.5)</td>
<td>46.2 (8.1)</td>
<td>44.8 (9.5)</td>
<td>46.3 (9.7)</td>
</tr>
<tr>
<td>/t/</td>
<td>82.4 (39.5)</td>
<td>73.9 (31.7)</td>
<td>61.4 (17.8)</td>
<td>60.8 (13.9)</td>
<td>55.8 (13.6)</td>
<td>55.8 (8.4)</td>
<td>52.8 (9.5)</td>
<td>53.8 (9.2)</td>
</tr>
<tr>
<td>/k/</td>
<td>100.6 (30.34)</td>
<td>83.2 (24.3)</td>
<td>79.5 (18.9)</td>
<td>77.9 (17.6)</td>
<td>72.6 (12.1)</td>
<td>66.9 (8.8)</td>
<td>66.4 (12.8)</td>
<td>70.1 (11.9)</td>
</tr>
<tr>
<td>/b/²</td>
<td>17.6 (10.2)</td>
<td>14.1 (6.7)</td>
<td>13 (4.3)</td>
<td>-38.6 (67.6)</td>
<td>-39.8 (58.5)</td>
<td>-62.3 (44.2)</td>
<td>-72.7 (38.2)</td>
<td>-69.4 (57.6)</td>
</tr>
<tr>
<td>/d/²</td>
<td>22.3 (10.2)</td>
<td>18.6 (5.4)</td>
<td>17.5 (9)</td>
<td>-19.6 (53.9)</td>
<td>-12.2 (49.2)</td>
<td>-47.1 (50.3)</td>
<td>-61.9 (34.3)</td>
<td>-45.4 (56.3)</td>
</tr>
<tr>
<td>/g/²</td>
<td>35.2 (13.3)</td>
<td>28.3 (13.1)</td>
<td>26.4 (10.8)</td>
<td>-9.5 (48.3)</td>
<td>+1.6 (45.2)</td>
<td>-21.1 (41.5)</td>
<td>-37.2 (43.5)</td>
<td>-35.2 (50)</td>
</tr>
</tbody>
</table>

² Merged VOT values (both positive and negative)
Table II. Age and gender distributions of participants by age group

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age Group</td>
<td>4;2- 6;0</td>
<td>6;1-9;0</td>
<td>9;1-11,6</td>
</tr>
<tr>
<td>n</td>
<td>18</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Gender</td>
<td>♀4 ♂14</td>
<td>♀4 ♂9</td>
<td>♀2 ♂5</td>
</tr>
<tr>
<td>Age (m) (SD)</td>
<td>4y11m (5.3m)*</td>
<td>7y8m (9.7m)*</td>
<td>10y4m (8m)*</td>
</tr>
</tbody>
</table>

*Mean (SD)
Table III. Mean (standard deviation) and range for VOT expressed in milliseconds for the children with speech sound disorders by age group

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=18</td>
<td>n=13</td>
<td>n=7</td>
</tr>
<tr>
<td>/p/</td>
<td>80.4 (41.4)</td>
<td>62.7 (50.5)</td>
<td>66.2 (22.3)</td>
</tr>
<tr>
<td></td>
<td>9.9-177.6</td>
<td>7.6-208.5</td>
<td>34.5-104.2</td>
</tr>
<tr>
<td>/t/</td>
<td>68.8 (37.8)</td>
<td>57.3 (65.1)</td>
<td>66.1 (26.2)</td>
</tr>
<tr>
<td></td>
<td>9.6-135.3</td>
<td>-102.3-209.8</td>
<td>29.9-115</td>
</tr>
<tr>
<td>/k/</td>
<td>100.7 (46.9)</td>
<td>82.6 (25)</td>
<td>99.2 (36.7)</td>
</tr>
<tr>
<td></td>
<td>36.8-237.1</td>
<td>39.6-141.4</td>
<td>62.8-167.2</td>
</tr>
<tr>
<td>/b/</td>
<td>10.6 (42.8)</td>
<td>-30 (67.5)</td>
<td>-34.9 (52)</td>
</tr>
<tr>
<td></td>
<td>-92.4-83.4</td>
<td>-173-26.7</td>
<td>-113.8-16.9</td>
</tr>
<tr>
<td>/d/</td>
<td>5.8 (72.8)</td>
<td>21 (49.2)</td>
<td>-22.8 (61.9)</td>
</tr>
<tr>
<td></td>
<td>-137.2-93.1</td>
<td>-90.1-110.9</td>
<td>-149.7-28.5</td>
</tr>
<tr>
<td>/ɡ/</td>
<td>22.3 (42.4)</td>
<td>-2.7 (12.8)</td>
<td>14.2 (43.9)</td>
</tr>
<tr>
<td></td>
<td>-103.6-99.7</td>
<td>-189.6-40.9</td>
<td>-78.4-56.2</td>
</tr>
</tbody>
</table>

1 Merged VOT values (both positive and negative)
Figure 1. Incidence of prevoicing in 3-5 years (Larsson and Wiman 2010) and 8-11 years and adults (Lundeborg et al, 2012)
Figure 2. Example of how the VOT values for the voiceless (to the left) and the voiced plosives (to the right) were defined and measured. The arrows in the voiceless example mark the plosive phase, the arrows in the voiced example indicate the prevoicing occurring during the plosive phase.
Figure 3. Proportion of participating children with deviant VOT (defined as 1 standard deviation below or above the mean for typically developed children in corresponding ages.)
Figure 4. Percentage of individuals across age groups having prevoicing when producing /b d g/