Voice onset time in language acquisition: Data from Hungarian

Judit Bóna, Anita Auszmann

Eötvös Loránd University, Hungary

bona.judit@btk.elte.hu, auszmannanita@gmail.com

Abstract

VOT provides important information about the language acquisition especially how the motor speech skill is being refined. Our hypotheses were the following: 1) If the age increases, the mean and standard deviation of VOTs decrease. 2) The VOTs of Hungarian-speaking children become similar to adults' VOTs at the age of 11. 3) The differences between genders depend on the age and the types of plosives. In this study 30 Hungarian-speaking children were participated in three age groups (9-, 11- and 13-year-olds). The VOTs of voiceless plosives ([p], [t], [k]) were analyzed in spontaneous speech. Our hypotheses were only partially proved by the results. There were significant differences between VOTs of the different plosives in all age groups. There were age differences in VOTs of all three plosives: with increasing age, the means of the VOTs of [p], [k] decreased, and the mean of the VOTs of [t] increased. Gender differences in VOTs depended on the age of the speakers.

Keywords: voice onset time, voiceless plosives, Hungarian speaking children

1. Introduction

Plosives occur in practically all languages of the world but they can be rather varied as far as their articulatory and acoustic properties are concerned (Maddieson 1984; Laver 1994). The silent closure phase of voiceless plosives is followed by a burst phase and then transition to the next sound; whereas in voiced plosives, voicing may be present in the closure phase and during the burst, too (Ladefoged & Maddieson 1996; Stevens 1998).

Voice onset time (VOT) is the length of time that elapses between the burst and the onset of voicing for the next voiced segment (Lisker & Abramson 1964; Zlatin 1974; Lieberman & Blumstein 1988). VOT values may differ from one language to the other, and they are influenced by a number of factors other than the voicing of the plosive including place of articulation (Volaitis & Miller 1992), the quality of the following vowel (Pind 1999), speech rate (Baum & Ryan 1993; Pind 1995), and speakers' sex and age (Whiteside–Marshall 2001; Whiteside–Dobbin–Henry 2003), too.

The analysis of VOT in children's speech is very important, because it provides us with information on how this aspect of motor speech skill is being refined (Whiteside–Dobbin–Henry 2003). VOT indicates atypical language acquisition, thus, the measurements of typical children's speech could serve as control for analysing the speech of atypically developing children. In addition, acquisition of VOTs is a challenge to many bilingual speakers, too (Fabiano-Smith–Bunta 2012).

In childhood, VOTs are affected by language acquisition, refinement of speech perception processes, and the fact that the motor control and articulation gestures are becoming more accurate (Whiteside–Marshall 2001; Whiteside–Dobbin–Henry 2003). In English-speaking children's speech production the variability of VOT gradually decreases, and VOTs become similar to adults' VOTs by the age of 11 (Whiteside–Marshall 2001).

There is also difference between males and females in VOT; however gender differences have little attention in the literature. Whiteside and Marshal (2001) analysed 7, 9, and 11 year-old children's VOTs, and they found gender differences at age of 11. With increasing age (between the analysed 7 and 11 years of age), VOTs of [p] and [b] decrease in males, while VOT of [t] increase in females. In another study, Whiteside, Dobbin and Henry (2003) found that differences between males' and females' VOTs were higher around 13 years of age, and gender differences occurred to different degree in various phonetic contexts. In analysis of Swedish children's speech at the ages 3 and 9, Karlsson et al. (2004) found significant gender effects in the aspirated plosives in the 3-year-old children's group that were not present in the plosives produced by adults. The authors hypothesised that the effect of gender at this early childhood might have been due to the differences in trans-glottal airflow properties (Karlsson–Zetterholm–Sullivan 2004).

We have only sporadic data about Hungarian children's VOT (for example Gósy 1984). Analysing Hungarian VOTs could lead us to different results from the above mentioned languages (English and Swedish), because they have plosives with aspirated VOTs, while Hungarian plosives are not aspirated (Gósy 2004). Hungarian VOT has been investigated both in word lists and in adults' spontaneous speech (Gósy 2000; Gósy & Ringen 2009; Gráczi et al. 2009; Neuberger–Gráczi 2013; Bóna 2011). The results show that VOT in Hungarian depends on the place of articulation, the following vowel, and the type of speech.

Bóna (2012) analysed voice onset time in voiceless plosives produced by members of three age groups in spontaneous speech. The three groups included 9–10-year-old children, 22–31-year-old young adults and 70–90-year-old speakers. The results showed that all three plosives under discussion ([p, t, k]) exhibited significant differences in VOT between age groups. The widest scattering of VOT values was found in the children's productions for all three consonants. Gender differences occurred only with [k] in the children's group, with [t] and [k] in the young adults' group, while in the group of old subjects, VOT values for all three consonants significantly differed between male and female subjects.

The aim of this paper was to investigate the VOT of voiceless plosives ([p], [t], [k]) in Hungarian-speaking children's speech. We had three preliminary hypotheses: 1) If the age increases, the mean and standard deviation of VOTs decrease. 2) The VOTs of Hungarian-speaking children are stabilized in spontaneous speech at the age of 11, and become similar to adults' VOTs. 3) The differences between genders in the velar plosives is detectable at the age of 9, but in the alveolar plosives can only be detected at the age of 13.

2. Subjects, material, and methods

2.1. Subjects

30 children's recordings were selected from speech samples recorded for earlier studies (Bóna 2012; Auszmann 2013; Neuberger 2013) for this analysis. The subjects were divided in three age groups: 9, 11, and 13 years old. All of them were Hungarian-speaking children, and none of them reported any hearing or speech disorders. In all groups there were 5 males and 5 females.

2.2. Material

Spontaneous speech was recorded from all children in a quiet room. Subjects were tested individually. They were asked to talk about their family, school and free time activities. We recorded spontaneous speech instead of word or sentence recall or read speech, because 1) in these ages articulation rate differs in these speech styles from spontaneous speech (it is slower in recall and reading); 2) in recalls the production of the item to repeat can influence the children's speech production.

In Hungarian VOT is not influenced by the phonetic position of the plosives. In a previous study, we found that phonetic position did not significantly affect the VOT values across the positions #C̲V, VC̲V, and CC̲V (Bóna 2011). A similar result was obtained by Gósy (2010) from her analysis of the speech of young men and women with respect to [t]. Therefore, in the present study, we did not separate the VOT data by position.

We analyzed about 100–150 data from each speaker to compensate the effects of other factors based on the characteristics of spontaneous speech. Alltogether 4146 VOTs were analyzed.

2.3. Methods

The VOTs of voiceless plosives ([p], [t], [k]) were analyzed using Praat 5.0 (Boersma–Weenink 2008). We annotated all of the files by hand.

VOT measurements raise a number of methodological issues (Francis et al. 2003; Gráczi & Kohári 2012). On the one hand, the burst cannot always be seen in the spectrogram, and the consonant may switch into fricative articulation after the closure period, or multiple bursts can be attested (Gráczi & Kohári 2012). On the other hand, researchers may disagree about the end point of voice onset time. Francis et al. (2003) compared four different methods of measurement with respect to VOT: the end point can coincide with the onset of regular voicing, that of F1, or F2, or F3. They performed both acoustic and electro-glottographic measurements and found that the most "accurate" method of measuring VOT (i.e., one that yielded the least amount of variability) was based on the onset of regular voicing.

Thus, in this study voice onset time was defined as the time span between the beginning of the burst and the absolute onset of voicing as observed on the oscillogram and on the spectrogram in parallel (Beckman et al. 2011). First we compared all VOTs of the same plosives across the age groups. Then we calculated the means and standard deviations for all speakers for the comparision. VOTs were also compared between males and females. Statistical analyses (one-way ANOVA for normal distribution, and Mann-Whitney *U*-test, the non-parametric alternative to the independent samples *t*-test) were carried out by SPSS 13.0 at the 95% confidence level.

3. Results

There were significant differences between VOTs of the different plosives in all age groups. There was only one exception: there was no significant difference between the VOTs of the 9-year-old children's bilabial and alveolar plosives (at this age between [p] and [k] $Z = -9{,}262$; $p \leq 0{,}001$; between [t] and [k] $Z = -9{,}262$; $p \leq 0{,}001$). (In 11-year-old children between [p] and [t] $Z = -4{,}939$; $p \leq 0{,}001$; between [p] and [k] $Z = -8{,}991$; $p \leq 0{,}001$; and between [t] and [k] $Z = -11{,}675$; $p \leq 0{,}001$. In 13-year-old children between [p] and [t] $Z = -4{,}125$; $p \leq 0{,}001$; between [p] and [k] $Z = -9{,}908$; $p \leq 0{,}001$; and between [t] and [k] $Z = -15{,}135$; $p \leq 0{,}001$.)

With increasing age, the means of the VOTs of [p] and [k] decreased, while the mean of the VOTs of [t] increased (Table 1). The mean of VOTs of [p] in native Hungarian speaking young adults was 18 ms (St. dev. 5,3); the mean of VOTs of [t] was 18 ms (St. dev. 5,8); and the mean of VOTs of [k] was 36 ms (St. dev. 10,7) (Bóna 2014). Standard deviations in the cases of [p] and [k] were lower in 13-year-old children than in 9-year-old children, while in the case of [t] standard deviation was higher in the older groups.

Table 1: *Means and standard deviations of VOTs.*

Age	VOT (ms)	St. dev.
[p]		
9	25	9,9
11	19	11,5
13	21	8,5
[t]		
9	23	7,9
11	25	12,2
13	25	8,9
[k]		
9	36	12,4
11	33	14,5
13	34	11,6

There were age differences in VOTs of all three plosives according to the statistical analyses (Figure 1). In case of [p], there were significant differences between 9- and 11-year-old children ($Z = -4{,}482$; $p \leq 0{,}001$), between 9- and 13-year-old children ($Z = -2{,}430$; $p = 0{,}015$), and between 11- and 13-year-old children ($Z = -2{,}374$; $p = 0{,}018$), too. In case of [t], there were significant differences between 9- and 13-year-old children ($Z = -5{,}322$; $p \leq 0{,}001$) and between 11- and 13-year-old children ($Z = -3{,}916$; $p \leq 0{,}001$), but there was no significant difference between 9- and 11-year-old children. In case of [k], there were significant differences between 9- and 11-year-old children ($Z = -5{,}804$; $p \leq 0{,}001$), between 9- and 13-year-old children ($Z = -2{,}961$; $p = 0{,}003$), and between 11- and 13-year-old children ($Z = -3{,}279$; $p = 0{,}001$), too.

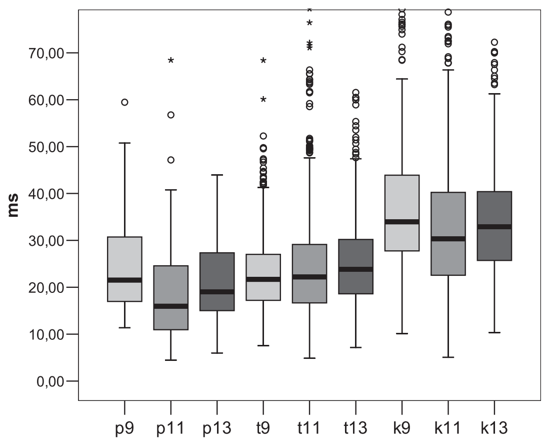

Figure 1: *Medians and ranges of VOTs.*

We calculated the means and standard deviations for all participants, and compared them between the age groups. The statistical analyses showed that there were no significant differences in the individual values between the age groups.

We compared VOTs of males and females in all plosives in each age group. In case of [p], there was no significant difference between boys and girls in any age group (Figure 2).

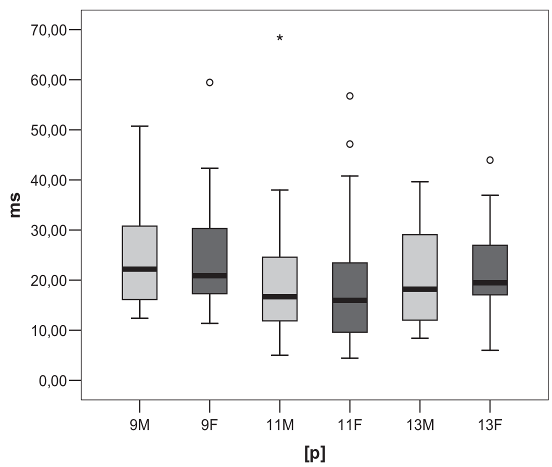

Figure 2: *Gender differences in VOTs of [p].*

In case of [t], there was no difference between the 9-year-old boys and girls, while the difference was significant between males and females in the 11-year-old group (Z = -2,362; p = 0,018) and in the 13-year-old group (Z = -2,467; p = 0,014) (Figure 3).

The results were unexpected in case of [k]: there was significant difference between the 9-year-old boys and girls (Z = -3,966; $p \leq$ 0,001), but in the other two age groups the difference was no significant between males and females (Figure 4).

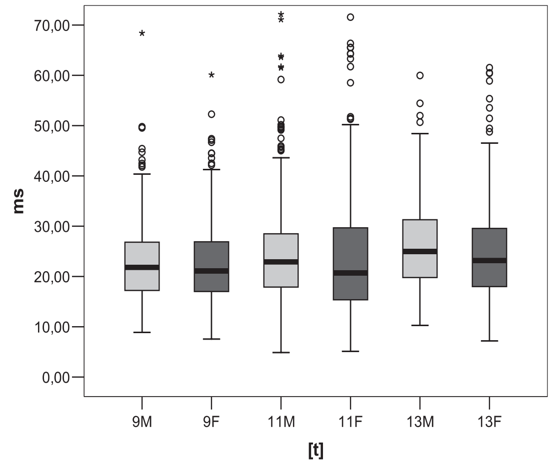

Figure 3: *Gender differences in VOTs of [t].*

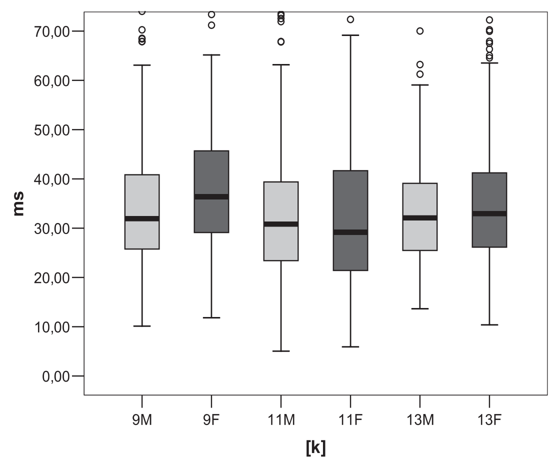

Figure 4: *Gender differences in VOTs of [k].*

4. Discussion and conclusion

In this paper, we analysed voice onset times in voiceless plosives produced by 9-, 11- and 13-year-old Hungarian children in spontaneous speech. We had three initial hypotheses, which were only partially confirmed.

Our first hypothesis was that with increasing age, the mean and standard deviation of VOTs decrease. This were confirmed only for [p] and [k], but in the case of [t] we observed opposite trend: the mean of the VOTs of [t] increased in the two older group compared to the 9-year-old children’s VOTs.

The second hypothesis was that VOTs of Hungarian-speaking children become similar to adults' VOTs at the age of 11. This hypothesis was partly confirmed, too. The VOTs of [k] were similar to the values measured in adults’ speech at the age of 9, while in the case of [p] at the age of 11. The VOTs of [t] were longer also at the age of 13 than adults’ VOTs (Bóna 2014).

The confirmation of our third hypothesis also depended on the type of plosives. The differences between genders in the velar plosives was detectable at the age of 9, but unexpectedly there was no difference between males and females in the other two age groups. In the alveolar plosives gender differences could be detected at the age of 11 and 13.

The results show that the place of articulation determines the age-related changes of VOT. In the analyzed age groups there are great individual differences between the children (as the statistics also showed that). It seems that the stabilization of the articulation of plosives is not completed at the age of 13. Our results show that the acquisition of voiceless plosives and VOT is language-specific.

5. Acknowledgements

The authors would like to thank Tilda Neuberger for the speech samples provided available for this research.

6. References

Auszmann, A. 2013. "A magánhangzók akusztikai szerkezete kisiskolások spontán beszédében" [Acoustic properties of vowels in children's spontaneous speech]. Paper presented at the 15th Summer School of Psycholinguistics, Balatonalmádi, 26–30 May 2013.

Baum, S. R., and L. Ryan (1993). "Rate of speech in aphasia: Voice onset time". *Brain and Language* 44, pp. 431–445.

Beckman, J., P. Helgason, B. McMurray, and C. Ringen (2011). "Rate effects on Swedish VOT: Evidence for phonological overspecification". *Journal of Phonetics* 39, pp. 39–49.

Boersma, P., and D. Weenink (1998). "Praat: doing phonetics by computer (Version 5.0.1)". http://www.fon.hum.uva.nl/praat/download_win.html.

Bóna, Judit (2011). 2A [p, t, k] mássalhangzók zöngekezdési ideje idősek és fiatalok spontán beszédében és felolvasásában" [VOT in [p t k] in spontaneous and read speech by young and old speakers]. *Beszédkutatás* 2011, pp. 61–72.

Bóna, Judit (2012). "A zöngekezdési idő életkori sajátosságai" [Age-realated variation in VOT]. Paper presented at the 14th Summer School of Psycholinguistics, Balatonalmádi, 20–24 May 2012.

Bóna, Judit (2014). "Voice onset time and speakers' age: Data from Hungarian". *Clinical Linguistics and Phonetics*. Accepted paper readable in Early Online.

Cho, T., and P. Ladefoged (1999). "Variation and universals in VOT: evidence from 18 languages". *Journal of Phonetics* 27, pp. 207–229.

Fabiano-Smith, L., and F. Bunta (2012). "Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers". *Clinical Linguistics and Phonetics* 26.2, pp. 148–163.

Francis, A. L., V. Ciocca, and J. M. C. Yu (2003). „Accuracy and variability of acoustic measures of voicing onset". *Journal of the Acoustical Society of America* 113.2, pp. 1025–1032.

Gandour, J., and R. Dardarananda (1984). "Voice onset time in aphasia: Thai II. Production". *Brain and Language* 23, pp. 177–205.

Gósy, M. (1984). "*Hangtani és szótani vizsgálatok hároméves gyermekek nyelvében*" [Phonetic and lexical studies on the language of three-year-olds]. *Nyelvtudományi Értekezések 102.* Akadémiai Kiadó, Budapest.

Gósy, M. (2000). "The voice onset time of the Hungarian voiceless plosives in words and in spontaneous speech". *International Journal of Speech Technology* 3–4, pp. 155–164.

Gósy, Mária (2004). "*Fonetika, a beszéd tudománya*" [Phonetics, the science of speech]. Budapest: Osiris Kiadó.

Gósy, M., and C. O. Ringen (2009). "*Everything you always wanted to know about VOT in Hungarian"*. Paper presented at the Ninth International Conference on the Structure of Hungarian, Budapest, 1 September 2009. http://icsh9.unideb.hu/pph/handout/Ringen_ Gosy_handout.pdf (downloaded 30 August 2012).

Gráczi, T. E., and A. Kohári (2012). "A zöngekezdési idő egy módszertani kérdés függvényében" [Voice onset time in a methodological perspective]. In Markó, A. (ed.): *Beszédtudomány. Az anyanyelv-elsajátítástól a zöngekezdési időig* [Speech science: from first language acquisition to voice onset time]. Budapest: ELTE Bölcsészettudományi Kar – MTA Nyelvtudományi Intézet, pp. 228–248.

Karlsson, F., E. Zetterholm, and K. P. H. Sullivan (2004). "Development of a gender difference in voice onset time". In: *Proceedings of the 10th Australian International Conference on Speech Science and Technology*. Canberra, ACT: Australian Speech Science and Technology Association, Incorporated, pp. 316–321.

Ladefoged, P., and I. Maddieson (1996). *Sounds of the world's languages*. Oxford: Blackwell.

Laver, John (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

Lieberman, P., and S. Blumstein (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge: Cambridge University Press.

Lisker, L., and A. S. Abramson (1964). "A cross-language study of voicing in initial stops: Acoustical measurements". *Word* 20, pp. 384–422.

Lisker, L., and A. S. Abramson (1967). Some effects of context on voice onset time in English stops. *Language and Speech* 10, pp. 1–28.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Neuberger, T. (2013). "Temporal patterns of children's spontaneous speech". *The Phonetician* 107/108, pp. 68–85.

Neuberger T., and T. E. Gráczi (2013). "Az alveoláris zöngétlen explozíva variabilitása". *Beszédkutatás* 2013, pp. 160–172.

Pind, J. (1995). "Speaking rate, voice-onset time and quantity: The search for higher-order invariants for two Icelandic speech cues". *Perception & Psychophysics* 57, pp. 291–304.

Pind, J. (1999). "The role of F1 in the perception of voice onset time and voice offset time". *Journal of the Acoustical Society of America* 106, pp. 434–437.

Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: The MIT Press.

Volaitis, L. E., and J. L. Miller (1992). "Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories". *Journal of the Acoustical Society of America* 92, pp. 723–735.

Whiteside, S. P., R. Dobbin, and L. Henry (2003). "Patterns of variability in voice onset time: a developmental study of motor speech skills in humans". *Neuroscience Letters* 347.1, pp. 29–32.

Whiteside, S. P., and J. Marshall (2001). "Developmental trends in voice onset time: some evidence for sex differences". *Phonetica* 58.3, pp. 196–210.

Zlatin, M. A. (1974). "Voicing contrast: Perceptual and productive voice onset time characteristics of adults". *Journal of the Acoustical Society of America* 56, pp. 981–994.