

## Voice Onset Time (VOT) In Prelingually Deaf Malay-Speaking Children with Cochlear Implants

Cila Umat<sup>1</sup>, Badruzaman Abdul Hamid<sup>2</sup>, Azlina Baharudin<sup>2</sup>

<sup>1</sup>Audiology Program, School of Rehabilitation Sciences, Faculty of Health Sciences, University Kebangsaan Malaysia, Kuala Lumpur, Malaysia.

<sup>2</sup>Speech Sciences Program, School of Rehabilitation Sciences, Faculty of Health Sciences, University Kebangsaan Malaysia, Kuala Lumpur, Malaysia.

\*Corresponding author: Cila Umat, Audiology Program, School of Rehabilitation Sciences, Faculty of Health Sciences, University Kebangsaan Malaysia, Kuala Lumpur, Malaysia, Tel: +603-9289-5035; Fax: +603-2698-6039; E mail: cila@ukm.edu.my

Received date: Nov 09, 2015; Accepted date: Dec 16, 2015; Published date: Dec 21, 2015

Copyright: © 2015 Umat C, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### Abstract

This study aimed to investigate the acquisition of the voicing contrast among Malay-speaking prelingually deaf children with cochlear implants (CI) and compared them to a group of normal hearing (NH) children. A total of 15 Malay children with 4 to 6 years duration of cochlear implant experience participated, 5 children in each age group. Secondary data from 15 NH children aged 4 to 6 years old was utilized for comparison. Speech samples were collected using a picture-naming task. There was a significant hearing age effect within the CI group. Comparing the CI and NH groups, a significant group effect was evidenced for all plosives, age effect was significant except for /b/ but there was no significant interaction between the group and age suggesting that in general, the pattern of responses was similar across age for both the study groups. Hearing age significantly correlated with the mean VOTs but not the age at implantation. The results suggest that the ability of these CI children to produce accurately the perceived voicing cues was not at par with the NH children with similar hearing experience. Longer duration of implant experience helps to improve the production of these sounds especially for the velar plosives.

### KEY WORDS:

Voice onset time (VOT), cochlear implant, Malay children, voiced, voiceless, plosives.

### Introduction

The acquisition of voice onset time (VOT) in plosive consonants among hearing-impaired individuals has long been of interest to many researchers especially following the use of a hearing device such as the cochlear implant [1-4]. The perception of VOT during binaural dichotic listening provides an insight on auditory discrimination processes [5] while the study on developmental acquisition of the VOTs gives an insight on the mechanism of the motor control of the voicing contrast [3,6,7]. The present study reported the acquisition of voicing contrast at word-initial position by prelingually deafened Malay children with cochlear implants and compared the mean VOT values of the voiced and voiceless plosives to that of hearing-age matched normal hearing (NH) children.

Plosive sounds at word initial position are often chosen by phoneticians to study from the acoustical perspectives due to the fact that listeners have a higher tendency to divide their attention to the beginning of a sound in a word that is being uttered rather than the middle or the final position. Moreover, the initial position is often viewed as the cue for an individual with hearing loss especially those who use hearing devices to aid their hearing. The ability to detect the initial consonant from a word being uttered benefits the hearing-impaired listeners as it signifies a conversation commences, among others. Voicing is a phonological contrast which begins to develop early in the speech of children. The acquisition of voicing in normal population is an ongoing process as children's voicing capabilities tend to continue developing as they grow older [8,9].

Voice onset time (VOT) is the length of time that is measured between the release of a closure of a plosive and the onset of voicing, see for example Lane and Perkell [6]. The early study of VOT was made by Lisker and Abramson [10]. VOT is a measurable acoustic parameter, resulting from the temporal coordination between laryngeal and supralaryngeal gestures [10,11]. It can be divided into two types which are voicing lag and voicing lead. When voicing occurs at the start right after the release of air following the glottal closure, it is referred to as voicing lag and is measured in positive VOT values. If the voicing occurs before the actual release of air prior to glottal closure, this is then referred to as voicing lead and is measured in negative VOT values. In addition VOT values also serves to distinguish between voiceless aspirated and unaspirated plosives in initial position. An aspirated sound occur where voice onset is appreciably later than the release of the occlusion and unaspirated sounds where voice onset occurs as the occlusion is released [12]. Cross linguistics, VOT values are different for different languages. English plosives, for example, have no negative VOT values but are contrasted based on short and long lags acoustic cues for the voiced and voiceless plosives, respectively. In Hebrew and Spanish, voiced plosives are characterized by voicing lead that have negative VOT values while voiceless plosives have positive values of voicing lag [13,15]. These features of negative and positive VOT values for voiced and voiceless plosives, respectively, are also seen in the Malay language as collected in our lab (manuscript in preparation).

In Asian Languages indeed there has been considerable work in VOT particularly on Korea and Hindi plosives such as in Han and Weitzman [15], Benguerel and Bhatia [16]. Other study on VOT also can be seen in Shimizu [17], where he examined the VOT value in initial plosives in six languages; Japanese, Mandarin Chinese, Korean, Burmese Hindi and Thai. VOT has been shown to be useful indication for a wide range of languages but for the mean time there have been no

systematic studies of VOT in Malay particularly in children with cochlear implants or hearing-impaired children in general.

Hearing is important for self-monitoring of the production of the plosive consonants in that in the absence of hearing, VOT anomalies (as compared to normal hearing listeners) have been reported even in postlingually deafened adults [4]. Kishon-Rabin et al., [3] for example, investigated the effect of auditory feedback on speech production of five postlingually deafened adults who were implanted with the Nucleus 22 cochlear implant device. They measured the changes in the speech production of the subjects before and post-implantation at 1, 6 and 24 months using various acoustic measurements including the VOTs. A significant increment of voiced plosives' voicing lead from positive to negative VOT values was reported over time until 2 years post-implant when values fell within the Hebrew norms. The results supported the hypothesis that restoration of hearing helps to recalibrate the mechanism of speech production as the adult CI users were able to monitor their articulation and their acoustic output. Studies done in prelingually deafened CI children [2,18] also yield significant differences between the CI and NH groups for VOT values albeit performing in a similar manner or similar trend of development as to that shown by the NH children.

The present study investigated the voicing contrast acquisition of prelingually deafened Malay-speaking children with CI who had had between 4 years to 6 years and 11 months hearing experience with the implant. The results were compared to matched groups of NH listeners with chronological age ranged from 4 years to 6 years and 11 months. Specifically, the study objectives were as followed: To investigate the (1) VOT values across the hearing age (i.e. the duration between the activation of the implant and the time of testing or duration of CI experience); (2) differences between CI and NH children in terms of their VOT values across the age groups; (3) correlations between the VOT values and the age at implantation and duration of CI experience.

## Method

### Subjects

A total of 15 CI subjects were chosen from those attending the Universiti Kebangsaan Malaysia (UKM) Cochlear Implant Program. Subjects were divided into 3 age groups: 4 years (hereafter denoted as 4;00) to 4 years and 11 months (hereafter denoted as 4;11), 5;00 – 5;11, and 6;00 – 6;11. There were 5 subjects in each hearing age group. All subjects chosen fulfill the following inclusion criteria: (1) using Malay as their dominant language; (2) prelingual deafness; (3) had a duration of CI experience for at least 4 years and a maximum of 6 years and 11 months at the time of the study; (4) used mainly oral communication mode; (5) had no other disabilities aside from hearing loss such as neurological impairment, mental retardation, autism and others; and (6) had parents' consent to participate in the study.

Data for the NH children was secondary. The data was collected as part of the VOT studies in Malay-speaking children and adults in our lab, using similar methodology as described in this paper for the CI children.

### Test procedure

Speech samples were collected by conducting a single data collection session with the CI subjects at their homes with the parents' consent. Recording was performed in a fairly quiet room at each home. Each session took about 10 – 20 minutes to complete. Each subject was

required to complete a picture-naming task which consists of 41 pictures with voiced /b, d, g/ and voiceless /p, t, k, / plosives at word-initial position, presented through a laptop. Subjects were asked to name the picture one by one when it was shown on the computer screen while the responses were audio-recorded using SONY IC digital audio recorder Model ICD-SX750/SX850 for a later analysis. After the subjects had completed a trial, the task was repeated with random presentations of the pictures to avoid memorization and to encourage spontaneous responses from the child. If the subject had problem naming any of the pictures or gave wrong answers, the researcher provided assistance by performing either by: (1) completing the incomplete sentence (e.g. "when it is raining, we use \_\_\_\_?"); or (2) describing the features of the objects; or (3) performing the gesture function of the object; or (4) delaying imitation in which the researcher did not collect the exact response at that time, but did it so during the second trial to promote spontaneous responses. Average VOT values from the two repeated trials for each sound category were calculated for the statistical analyses purposed.

### Data analyses

The speech samples collected were transcribed into the PRAAT software 4.3.33 [19] in which the samples were displayed in the spectrogram view. By using PRAAT, VOT values were measured in milliseconds (ms) and the average VOT values from the two trials, for each of the sounds were calculated. Further statistical analyses were conducted using the Statistical Package for Social Sciences (SPSS) version 18 [22].

Descriptive analyses were used to get the mean and the standard deviation values of the voiced and voiceless plosives across the age and study groups. A one-way ANOVA was used to determine the hearing age effect within the CI group. To compare with the NH group, a two-way multivariate analysis was performed with group and age as the fixed factors. For the correlation analyses, Pearson correlation was used to analyse the correlations between the VOT values of voiceless plosives /p, t, k/ and the age at implantation and duration of CI experience while Spearman correlation was utilized for the voiced plosives /b, d, g/ as the data distribution was not normal for these sounds.

### Ethical approval

This study was approved by the Universiti Kebangsaan Malaysia (UKM) Human Ethical Committee to be conducted on human subjects starting from July 2010 to June 2012. Consent forms and information sheet for parents regarding the study were distributed before the children participated in this study.

## Results

### VOT values across the hearing age

Within the CI group alone, there was a significant hearing age effect for all the VOT values of the voiced and voiceless plosives. Table 1 shows the F and p-values of the one-way ANOVA analysis, the mean and the standard deviations (SDs) for each of the plosive across the hearing age and the post hoc analyses. In general, it can be seen from Table 1 that the significant VOT differences were evidenced in the 4- and 6-year olds and 5- and 6-year old age groups but not between the 4- and 5-year hearing experience groups for all plosives except for /b/.

Plosives	F value (Hearing Age effect)	p value	Mean ± SDs			Post hoc tests
			4y.o.	5y.o.	6y.o.	
/p/	F (2,12) = 12.855	0.001	15.149 ± 2.139	18.155 ± 1.119	21.797 ± 2.665	4 & 5 y.o. p = 0.123; 4 & 6 y.o. p = 0.001; 5 & 6 y.o. p = 0.051
/t/	F (2,12) = 9.068	0.004	19.738 ± 1.465	20.223 ± 2.039	25.591 ± 3.339	4 & 5 y.o. p = 1.000; 4 & 6 y.o.p = 0.007; 5 & 6 y.o.p = 0.013.
/k/	F (2,12) = 42.714	0.000	22.340 ± 1.667	24.143 ± 1.177	31.318 ± 1.938	4 & 5 y.o. p = 0.315; 4 & 6 y.o. p = 0.000; 5 & 6 y.o. p = 0.000
/b/	F (2,12) = 4.457	0.036	-2.942 ± 17.309	-20.224 ± 5.174	-20.356 ± 3.356	4 & 5 y.o. p = 0.073; 4 & 6 y.o. p = 0.070; 5 & 6 y.o. p = 1.000
/d/	F (2,12) = 11.456	0.002	4.152 ± 23.786	18.978 ± 2.197	-22.550 ± 3.051	4 & 5 y.o. p = 0.353; 4 & 6 y.o. p = 0.031; 5 & 6 y.o. p = 0.001
/g/	F (2,12) = 111.156	0.000	26.356 ± 4.294	28.082 ± 5.333	-32.489 ± 10.658	4 & 5 y.o. p = 1.000; 4 & 6 y.o. p = 0.000; 5 & 6 y.o. p = 0.000

**Table 1:** The F statistics and p-values of the hearing age effect for all plosives within the CI group. Also shown are the post hoc analyses to compare the differences between the age groups.

**Comparing the VOT values between the CI and NH children**

A two-way multivariate analysis was performed to examine the group and age main effects on the VOT values of all the plosives. Table

2 reveals the F and p values of both the main effects of group and age and the interaction between the two independent variables.

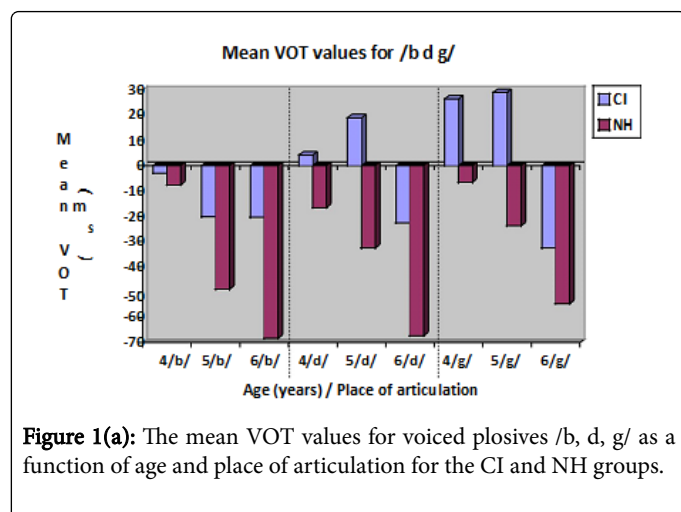
Plosives	F value (Group, GP)	p-(GP)	F (Age)	p-(Age)	F (GP*Age)	p- (GP*Age)
/p/	F(1,24) = 98.077	0.000	F(2,24)= 7.624	0.003	F(2,24)= 0.726	0.494
/t/	F(1,24) = 38.837	0.000	F(2,24)= 3.934	0.033	F(2,24)= 0.579	0.568
/k/	F(1,24) = 10.118	0.004	F(2,24)= 14.705	0.000	F(2,24)= 1.251	0.304
/b/	F(1,24) = 16.785	0.000	F(2,24)= 1.241	0.307	F(2,24)= 0.366	0.697
/d/	F(1,24) = 18.259	0.000	F(2,24)= 7.854	0.002	F(2,24)= 1.047	0.366
/g/	F(1,24) = 11.447	0.002	F(2,24)= 10.009	0.001	F(2,24)= 0.682	0.515

**Table 2:** The F statistics and p-values for the main effects of Group and Age and the interaction between the Group and Age for the all the voiced and voiceless plosives. Values in bold are significant effects.

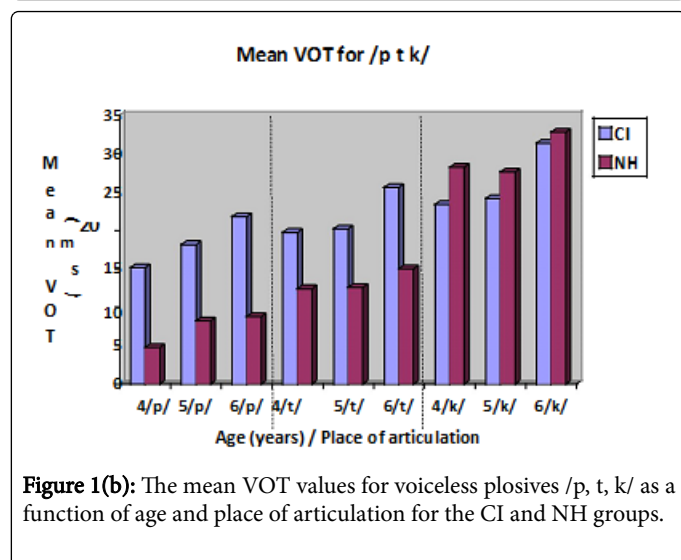
It can be seen from Table 2 that the main effect of group (i.e. comparing between the CI and NH groups) was significant for all VOTs of the voiced and voiceless plosives across all the age groups. The age effect was not significant for the VOT/b/. There was no interaction between the group and age effects suggesting that in general, the pattern of responses was similar across age for both the study groups.

In general, for the voiced plosives /b, d, g/, the mean VOT values decrease (more negative values) as a function of age (Figure 1a) while for the voiceless plosives /p, t, k/, as the age increases, subjects in both groups showed longer mean VOTs (Figure 1b). For the CI children with 4 and 5 years hearing age, the mean VOT for voiceless /k/sound

was similar to the mean VOT for voiced /g/ sound (~25 ms) suggesting these velar sounds were perceived as similar. That is, there is an effect of substitution of the voiced for the voiceless sounds. The separation of the cognate VOTs was observed in the 6-year hearing experience with the CI group. For the voiceless plosives /p, t/, the CI children consistently showed longer mean VOTs than the NH children across the age group but not for the /k/ sound in which the mean VOTs for CI were shorter than the NH children across the age groups (see Figure 1b). The following figures show the mean VOT values for the voiced (1a) and voiceless (1b) plosives as a function of age and place of articulation for both the CI and NH groups.



**Figure 1(a):** The mean VOT values for voiced plosives /b, d, g/ as a function of age and place of articulation for the CI and NH groups.



**Figure 1(b):** The mean VOT values for voiceless plosives /p, t, k/ as a function of age and place of articulation for the CI and NH groups.

### Correlation analyses

The VOT values were submitted for simple bivariate analyses with the age at implantation and duration of CI experience (i.e. the hearing age of the CI subjects). Table 3 shows the correlation coefficients and the p-values obtained in these analyses.

	VOT /p/	VOT /t/	VOT /k/	VOT /b/	VOT /d/	VOT /g/
<b>Age at implantation</b>	r = 0.040 p = 0.887	r = 0.206 p = 0.461	r = 0.501 p = 0.057	r = 0.002 p = 0.994	r = -0.40 p = 0.13	r = -0.546 p = 0.035
<b>Hearing age</b>	r = 0.824 p = 0	r = 0.699 p = 0.004	r = 0.885 p = 0	r = -0.624 p = 0.013	r = -0.54 p = 0.03	r = -0.661 p = 0.007

**Table 3:** The correlation analyses between the age at implantation and hearing age of the CI children and the mean VOT values for all the plosives. Pearson correlation was used for the /p, t, k/ sounds while the

non-parametric Spearman correlation was utilized for the voiced /b, d, g/ correlations. Values in bold are significant correlations.

### Discussion

The present study investigated the VOT values of voiced and voiceless plosives produced by Malay-speaking prelingually deafened CI children with duration of CI experience ranged between 4;00 to 6;11 years. The results were then compared to a group of NH children with chronological age similar to the hearing age of the CI children.

Our first main finding was that within the CI group, hearing age effect was significant especially between the 4- and 6-year old groups and 5- and 6-year old groups. The results suggest that in general, 6 years hearing experience through the cochlear implant device seemed to be the ‘cut-off’ point for a leap in the ability of the Malay CI children to monitor their production of the plosive sounds in terms of its voicing contrast. This was evidenced from our data as the 4- and 5-year old groups were not significantly different in terms of their VOTs. The fact that there was a significant effect of hearing age in that as the experience of hearing through the CI device increases, the mean VOTs of the voiceless plosives became longer and more negatives for the voiced plosives, support the findings by others [2,3,8,14] that show incremental changes of VOT over time with hearing. The presence of hearing ability allows the auditory feedback loop to operate and enable the CI children to monitor production of the speech sounds.

Comparing the mean VOTs of all plosives between the CI and NH groups reveals significant group and age effects but no interaction between the two main effects. The fact that there was no interaction between the two suggests that the pattern of responses were similar across the age groups for both the study groups. The results indicate that while the use of CI helps to give significant access to sounds to enable auditory feedback monitoring, thus, developing voicing contrast ability, these prelingually deaf CI children with hearing age comparable to the NH children, likely to perceive less sharp-tuned perception of the voicing contrast especially for the children with 4- and 5 years hearing experience with the implant and for the more posterior place of articulation sounds (i.e. /k/ and /g/).

The more anterior voiceless plosives /p, t/ consistently show longer VOT values for the CI as compared to the NH children as a function of hearing age except for the /k/ sound. For the voiceless velar /k/, the CI subjects had shorter mean VOTs than the NH group across all hearing ages. It could be possible that while reorganization of the coordination and timing between laryngeal and oral articulation takes place with the presence of hearing, the fact that this sound is at the back of the articulatory organ, makes it more difficult for the CI children to perceive the voicing cue. It is known that place of articulation cue relies a lot on frequency information and that this cue is available in the cochlear implant system through different activation of the electrode place. It has been shown that CI users who had difficulties to differentiate different electrode place were associated with difficulties to discriminate speech [21]. This hypothesis is further supported by examining our VOT data for the voiced velar /g/ which showed striking differences especially for the younger age groups. The subjects with shorter duration of CI experience had positive VOT values as compared to the NH subjects who consistently had negative VOT values. The results which show similar mean VOTs for voiced and voiceless velar sounds for the CI children with 4- and 5-years hearing experience, which was around 25 ms, suggest overlapping perception of these sounds. That is, there is an effect of substitution of the voiced

for the voiceless sounds or vice versa. The separation of the cognate VOTs was observed in the 6-year hearing experience with the CI as negative mean VOT was obtained for the voiced /g/. Our present finding support in part an earlier study which found that prelingually deaf children who had worn their CIs for a longer time were more likely to produce the place feature correctly [20]. Other studies have reported that as the hearing experience increases, hearing-impaired speakers will increase their speaking rate and variability in speaking rate and vocal loudness have also been shown to affect the VOT values [18]. In addition, speech breathing mechanism in prelingually deafened speakers showed anomalies in a more severe form than postlingually deafened speakers that affect the production of the voicing contrasts by these subjects [6].

In the present study, we did not find that age at implantation significantly correlated with the mean VOTs for all plosives except for the voiced velar /g/. Hearing age however, significantly correlated with all the mean VOTs. As has been discussed earlier, several studies have shown that hearing experience for the hearing-impaired speakers help to recalibrate the mechanism for motor control of the voicing contrast that lead to more accurate speech production. Earlier age at implantation did not necessarily mean better speech production suggesting acquisition of the voicing contrast ability is a developmental process and hearing experience is important to allow the child to learn to produce the necessary upper vocal tract and laryngeal gestures and to coordinate them with very precise timing according to the rule of the language [8].

## Conclusion

The prelingually deafened Malay CI children showed similar pattern of acquisition of the voicing contrast as their NH peers albeit not at par with them despite similar hearing experience. Hearing ability obtained through their CI device helps to activate the auditory feedback loop in these hearing-impaired children with a leap in performance observed at 6 years hearing age. Velar sounds seemed to be more difficult to differentiate in which overlapping perception of the voiced and voiceless sounds of /k/ and /g/ was suspected as the mean VOTs for these sounds for the 4-and 5- year old groups were similar at around positive value of 25 ms. A longitudinal study is recommended to continue monitoring the CI children acquisition of the voicing contrast to determine at what hearing age the difference would be insignificant between the CI and NH groups and whether similar developmental trend would continue. This will help to confirm the model of acquisition that focuses on the mastery of gestural coordination with hearing rather than segmental contrasts in the CI children.

## Acknowledgement

The authors would like to thank all CI children who participated in this study and their parents for giving the consent to involve in the research study.

## REFERENCES

1. Hudgins CV (1934) A comparative study of the speech coordinations of deaf and normal subjects. *J Genetic Psychology* 44: 1-48.

2. Kant AR, Patadia R, Govale P, Rangasayee R, Kirtane M (2012) Acoustic analysis of speech of cochlear implantees and its implications. *Clin Exp Otorhinolaryngol* 5: S14-18.

3. Kishon-Rabin L, Taitelbaum R, Tobin Y, Hildesheimer M (1999) The effect of partially restored hearing on speech production of postlingually deafened adults with multichannel cochlear implants. *J Acoust Soc Am* 106: 2843-2857.

4. Lane H, Wozniak J, Perkell J (1994) Changes in voice-onset time in speakers with cochlear implants. *J Acoust Soc Am* 96: 56-64.

5. Arciuli J, Rankine T, Monaghan P (2010) Auditory discrimination of voice-onset time and its relationship with reading ability. *Laterality* 15: 343-360.

6. Lane H, Perkell JS (2005) Control of voice-onset time in the absence of hearing: a review 48: 1334-1343.

7. Lane H, Wozniak J, Matthies M, Svirsky M, Perkell J (1995) Phonemic resetting versus postural adjustments in the speech of cochlear implant users: an exploration of voice onset time. *J Acoust Soc Am* 98: 3096-3106.

8. Lowenstein JH, Nittrouer S (2008) Patterns of acquisition of native voice onset time in English-learning children. *J Acoust Soc Am* 124: 1180-1191.

9. Lundeborg I, Larsson M, Wiman S, McCallister AM (2012) Voice onset time in Swedish children and adults. *Logopedics Phoniatrics Vocology* 37: 117-1122.

10. Lisker A, Abramson A (1964) A cross language study of voicing in initial plosive: Acoustical measurement. *Words* 20: 384-422.

11. Weismer G (1984) Control of the voicing distinction for intervocalic plosives and fricatives. *J Phonetics* 8: 427-438.

12. Clark J, Yallop C (1996) *An introduction to Phonetics and Phonology* (2nd edn). Oxford: Blackwell.

13. Fabiano-Smith L, Bunta F (2012) Voice onset time of voiceless bilabial and velar plosives in 3-year old bilingual children and their age-matched monolingual peers. *Clinic Linguist Phonet* 26: 148-163.

14. Taitelbaum-Swead R, Hildesheimer M, Kishon-Rabin L (2003) Effect of voice onset time (VOT), plosive burst and vowel on the perception of voicing in Hebrew plosives: preliminary results. *J Basic Clin Physiol Pharmacol* 14: 165-176.

15. Han MS, Weitzman RS. Acoustics features of Korean /P,T,K/, /p,t,k/, and / p,t,k/

16. Bengueural AP, Bhatia KT (1980) Hindi Plosives Consonants: An Acoustics and Fiberscopic Study. *Phonetica* 37: 134-148.

17. Shimizu K (1996) A cross-language study of voicing contrast of plosive consonants in Asian languages. Seibido Publishing, Tokyo.

18. Knuttila EL (2011) The effects of vocal loudness and speaking rate on voice-onset time in typically developing children and children with cochlear implants. University of Alberta, Canada.

19. Boersma P, Weeninnk D (2005) PRAAT: Doing Phonetics by Computer (version 4.3.33).

20. Tye-Murray N, Spencer L, Gilbert-Bedia E (1995) Relationships between speech production and speech perception skills in young cochlear-implant users. *J Acoust Soc Am* 98: 2454-2460.

21. Henry BA, McKay CM, McDermott HJ, Clark GM (2000) The relationship between speech perception and electrode discrimination in cochlear implantees. *J Acoust Soc Am* 108: 1269-1280.

22. Carver RH, Gradwohl J (2011) *Doing data analysis with SPSS, version 18* (5th edn.) Duxbury Press, California.