Effect of Temporal Variation on Phoneme Identification Skills in 2-3 Year old Typically Developing Children

[1]Powlin Arockia Catherine S. & [2]Savithri S. R.

Abstract

The present study investigated the ability of 2 to 3 year old Kannada speaking children to identify synthetic phonemes varying in voice onset time (VOT). Four picturable (two minimal pair) words with stop consonants contrasting in voicing (labial: /b/ and /p/, velar: /g/ and /k/) in word-initial position in Kannada were selected. These words as uttered by 21-year-old female native Kannada speaker were recorded and stored onto the computer memory. A VOT continuum from /b - p/, and /g - k/ were synthesized using Praat software which was audio recorded on to a CD. The edited tokens were audio presented to thirty (15 boys and 15 girls) 2 to 3 year old children individually. Subjects pointed to one of the pictures placed before them as they heard the tokens. The investigator noted their responses on a binary forced-choice scoring sheet. The results indicated that the 50% crossover from voiced to unvoiced cognate occurred in the lead VOT region and boundary width was wider for /g - k /continuum compared to /b - p/ continuum. The results of the current study can be used to compare phoneme identification skills in clinical population of same age.

Keywords: Voice onset time, synthesis, stop consonants, phoneme boundary

Speech perception refers to the processes by which humans are able to interpret and understand the sounds used in a language. Research on speech perception seeks to understand how human listeners recognize speech sounds and use this information to understand spoken language. Studies on infant speech perception postulates that the ability to perceive universal phoneme contrast is present at birth and with exposure, infants loose this ability and could perceive only the native contrasts (Werker & Tees, 1984). Also, cross language studies on adults demonstrated language specific perception patterns (Abramson & Lisker, 1970). A great deal of modification of perceptual abilities takes place between infancy and adulthood. It is important to investigate and document the modification process in phoneme perception during language development period, which would also strengthen our understanding of perception-production relationship. Phoneme perception ability is studied by altering the acoustic correlates that distinguishes speech sounds from one another.

In the past, perception of stop consonants along voice onset time (VOT) continuum has been studied in infants, children and adults. The early studies in children didn't show a developmental trend in perception. Winterkorn, MacNeilage & Preston (1967) using VOT continuum of /t - d/ reported that children aged 2.9 to 3.6 years could identify stop consonants similar to adults. Also, Yeni-Komishian, Preston & Cullen (1967) experimented 5 to 6 year old American English speaking children's ability to identify synthetic syllables (apical consonants /t/ and /d/) through imitation. The authors reported adult like perception in their subjects. However, results by Zlatin & Koenigsknecht (1975), Simon & Fourcin (1978) and Williams (1977a, 1977b) indicate a development trend in VOT in English and French speaking children. In the Indian context, Savithri (1996) (Kannada), Sathya (1996) (Telugu) and Catherine & Savithri (2007) (Kannada) reported developmental trend in VOT in children.

Studies in the past have focused on different age groups with small sample size and in different languages. The results found in one language cannot be generalized to other languages since auditory processing skills may differ with languages as the phonemic structure of languages are different and also normative research is required for clinical purposes in individual languages. In this context, the present study investigated phoneme identification skills in typically developing Kannada speaking children between the age range of 2 and 3 years. More specifically, the investigation would help in finding the cause for late talking and Specific language

[1]*Junior Research Fellow, Department of Speech-Language Sciences, AIISH, Mysore, email:arockia_cath@yahoo.com*, [2]*Professor of Speech Sciences, Department of Speech-Language Sciences, AIISH, Mysore. email:savithri_2k@yahoo.com*

impairment.

Method

Subjects: Thirty typically developing, Kannada speaking children from four play schools in Mysore participated in the study. The subjects included 15 boys and 15 girls in the age range of 2 to 3 years. All the children were from middle socio-economic status. The children were formally screened for speech, language and hearing abilities by the experimenter and those who passed the screening were included in the study.

Stimuli: Two meaningful, picturable, bisyllable word pairs with stop consonants in the initial position in Kannada were selected. The two words in a pair contrasted in voicing (p - b; pennu – bennu, k - g; kere - gere). The words as uttered thrice by a 21-year-old female, native Kannada speaker were recorded onto a computer using SSL Pro3V3 software (Voice and Speech Systems, Bangalore) and stored onto the computer memory. Mean voice onset time, phoneme duration and total word duration were obtained using waveform display. Voice onset time (VOT) continuum was synthesized using lengthen (PSOLA) module of Praat software. Waveforms of /b/ and /g/ (from onset of voicing to start of burst) were truncated from the waveform display of the words /bennu/ and /gere/ in steps of 0.9 (factor) till the burst using PSOLA module and then concatenated to the original word. Silence in steps of 10 ms was inserted between the burst and voicing of the vowel, till +40 ms. A total of 15 tokens were generated for /b - p/ VOT continuum (-77 to +40 ms) which was termed test 1. Fifteen tokens were generated for /g - k/ VOT continuum (-78 to +40 ms) and constituted test 2. The tokens were iterated thrice, randomized and recorded onto a CD. Thus a total of 90 tokens (15*2*3) formed the stimulus.

Procedure: Children were tested individually. Each child was familiarized with the pictures of the words used in test 1 and 2 before data collection. Training was provided using live voice and was for a maximum of 3 half-an hour session. Four pictures were presented to the children at a time, two were the pictures of the target stimuli (bennu and pennu/ gere and kere) and other two were distracters. Of the two distracters, one was semantically related to a target stimulus and the other was unrelated to the target stimuli. This is to reduce the chance of false positive responses. The pictures were placed in front of the child. The investigator named the pictures and the child was instructed to listen to it and point to the respective picture. Childs responses were noted by the investigator. After the child correctly responded 100% of the time to the target words, synthetic tokens were introduced. As in training session, experiment was carried out in a quiet room. The stimuli were audio-presented through two speakers placed at 45^{o} azimuth at a comfortable loudness. The child was instructed to carefully listen to the synthetic tokens and point to the respective picture out of four picture cards placed before him/her. The investigator noted the child's responses on a binary forced-choice scoring sheet. Percent identification scores were calculated and identification curves were plotted. From identification curve, 50% crossover, lower limit of phoneme boundary (LLPB), upper limit of phoneme boundary (ULPB) and phoneme boundary width (PBW) were obtained (Doughty, 1949). Fifty percent crossover is the point at which 50% of the subject’s response corresponds to the voiced (voiceless) category. Lower limit of phoneme boundary width is the point along the acoustic cue continuum where an individual identified voiced (voiceless) stop 75% of the time and upper limit of phoneme boundary width defined as the corresponding point of the identification of the unvoiced (voiced) cognate 75% of the time. Phoneme boundary width was determined by subtracting the lower limit from upper limit of boundary width. Mean 50% crossover, mean LLPB, mean ULPB and mean phoneme boundary width was calculated. Independent sample t- test and paired sample t-test were carried out to find the gender difference and across continuum difference.

Results

/b - p/ continuum: Results indicated that mean values of 50% crossover occurred at lead VOT region (-7.9 ms). Also, mean LLPB and mean ULPB were found in the lead VOT region. The mean phoneme boundary width for children was 12 ms. Independent sample t- test showed no significant difference between gender for 50% crossover [t (28) = -0.96; p>0.05], LLPB [t (28) = -1.72; p>0.05], ULPB [t (28) = -0.13; p>0.05] and phoneme boundary width [t (28) = 1.11; p>0.05]. Table 1 shows the mean values of all the parameters for /b-p/ continuum. Figures 1 and 2 show identification functions in a boy and girl for /b - p/ continuum.

Parameters	Boys	Girls	Average
50% rossover	-6.47	-10.27	-7.9
LLPB	-10.13	-17.67	-14.47
ULPB	-2.47	-3.13	-2.23
PBW	10.47	13.87	12.03

Table1: Mean values of measures for /b-p/ continuum (in ms).

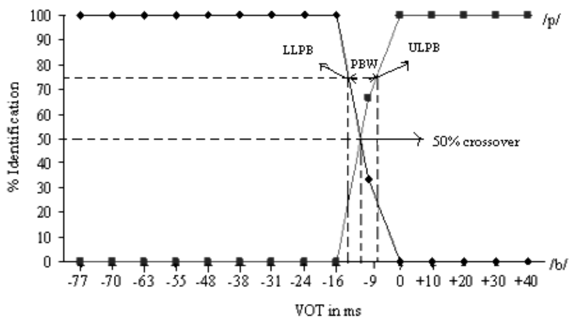

Figure 1: Percent identification scores of a boy (B3) on /b/-/p/ continuum.

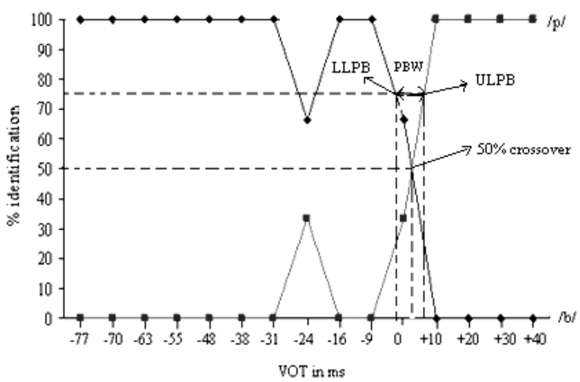

Figure 2: Percent identification scores of a girl (G7) on /b/-/p/ continuum.

/g - k/ continuum: Children changed their percept from /g/ to /k/ in lead VOT region (-9.8 ms). Lower limit of phoneme boundary and ULPB occurred in lead (-15.8 ms) and lag (5.8 ms) VOT region, respectively. Also mean phoneme boundary width was 19.6 ms. Independent sample t-test showed no significant difference across gender for 50% crossover [t (28) = -1.602; p>0.05], LLPB [t (28) = 1.44; p>0.05], ULPB [t (28) = -0.039; p>0.05] and phoneme boundary width [t (28) = 0.545; p>0.05]. Table 2 shows the mean values of all parameters for /g - k/ continuum. Figures 3 and 4 show percent identification in a boy and a girl for /g - k/ continuum.

Parameters	Boys	Girls	Average
50% crossover	-6	-14.33	-9.8
LLPB	-12.33	-19.33	-15.8
ULPB	4.07	3.87	5.83
PBW	18.27	21	19.63

Table2: Mean values of measures for /g-k/ continuum (in ms).

Results of paired t-test indicated significant difference between continuums for ULPB [t (29) = 2.73; p< 0.05] and phoneme boundary width [t (29) = 2.94; p< 0.05]. Upper limit of phoneme boundary and phoneme boundary width were significantly higher in /g - k/ continuum compared to /b - p/ continuum.

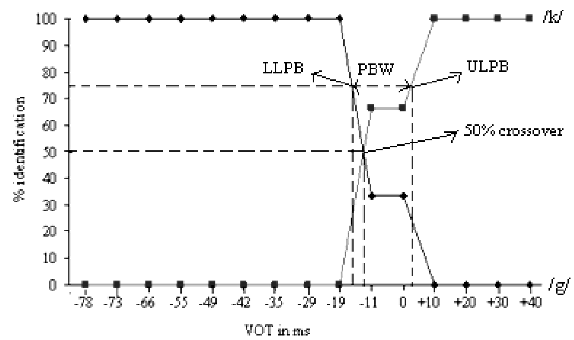

Figure 3: Percent identification scores of a boy (B15) on /g/-/k/ continuum.

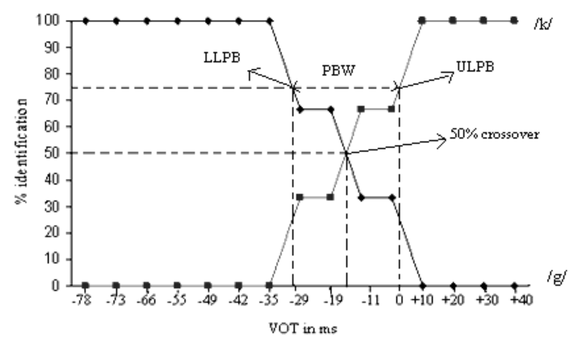

Figure 4: Percent identification scores of a girl (G7) on /g/-/k/ continuum.

Discussion

The results revealed several interesting findings. Firstly, 50% cross over occurred in lead VOT region for /b - p/ (-7.9 ms) and /g - k/ (-9.8 ms) continuum. That is, the children shifted their percept from voiced to unvoiced stop consonants in the lead region of VOT. This result is in consonance with those of Abramson & Lisker (1965) in Thai language (-20 ms), Savithri (1996) in Kannada speaking children (-16.8ms) and Sathya (1996) in Telugu speaking children (-10 ms), and Catherine & Savithri (2007) in Kannada speaking children (-22 ms) who have reported 50% crossover in lead VOT region. The result is in contrast with those of Winterkorn, MacNeilage & Preston (1967), Yeni-Komishian, Preston & Cullen (1967), Zlatin & Koenigsknecht (1975), Simon & Fourcin (1978) Williams (1977a, 1977b) and Flege & Eefting (1986) who reported 50% crossover in lag VOT region. The difference in result may be due to the phonemic structure of the two languages. In English, stops have two way (i.e. p, b) classification where as Kannada has a four way (e.g. p, ph, b, bh) classification. The distinction between unvoiced aspirated and unvoiced unaspirated stop is phonetic in English but phonemic in Kannada (Savithri, 1996). As in Spanish (Williams, 1977b), listener in the present study gave greater weight to prevoicing as a cue to voiced stop consonants than English listeners. This indicates that phonetic processing of speech sounds are governed by acoustic properties of stops found in a language (Aslin & Pisoni, 1980). Cross language studies by Simon & Fourcin, (1978), Flege & Eefting (1986) also suggest that speakers of different languages may learn to perceive stop consonants differently as they are exposed to different kinds of stops.

Secondly, phoneme boundary was within ±20 ms. This finding is in consonance with the reports of Abramson & Lisker (1965), Savithri (1996), Sathya (1996) and Catherine & Savithri (2007) who also reported phoneme boundary within ±20 ms region. Aslin & Pisoni (1980) asserts commenting on the experiment by Pisoni (1977) on perception of non speech stimuli that a psychophysical process is probably responsible for the categorical-like discrimination within this region. Also, such narrow region of high discriminability had resulted due to general sensory constraints on mammalian auditory system to resolve small differences in temporal order and not because of phonetic categorization.

Third, phoneme boundary width for /b - p/ continuum (12 ms) and /g - k/ continuum (19.63 ms) were narrower in the present study compared to those reported by Savithri (1996) for both /b - p/ (25 ms) and /g - k/ continuum (34 ms) and Catherine & Savithri (2007) for /b - p/ continuum (36 ms). This difference in boundary width may be attributed to smaller subject size that is, 6 subjects per age group (4 - 6 year old children) in Savithri's (1996) study and 10 subjects (2 - 3 year old children) in Catherine & Savithri's (2007) study. Also adult's phoneme boundary width was wider, 13 ms for /b/-/p/ and 25 ms for /g/-/k/ continuum (Savithri, 1996) compared to 2 - 3 year old children in the present study.

Fourth, there was no significant difference in the identification scores of boys and girls implying that phonetic processing abilities are similar for both the gender. Zlatin & Koenigsknecht (1975) also reported no significant difference between genders.

Fifth, ULPB occurred in the lead VOT region for /b-p/ continuum and in lag VOT region for /g - k/ continuum. Also, boundary width for /g - k/ continuum was wider compared to /b - p/ continuum. This trend was also reported by Zlatin & Koenigsknecht (1975) and Savithri (1996). This trend probably reflects the pattern of phoneme acquisition in which labial stop consonants are acquired prior to velar stop consonants.

To conclude, typically developing Kannada speaking children in the age range of 2-3 years had adult-like identification skills on /g-k/ and /b-p/ continuum. Future research on older age groups to study the developmental pattern of speech perception in Kannada and other languages are warranted.

Conclusions

The present study provides a basic knowledge on phoneme identification skills in 2-3 year old Kannada speaking typically developing children. The phoneme identification skills of normal children can be compared with clinical population including late-talking children, children with hearing impairment, mental retardation, seizure disorder and high-risk children. More specifically, late talking children between 2 to 3 years who exhibit language disorder in the absence of specific causes may be impaired in phoneme identification. Also, using the findings of this study as baseline, perception training program for this age group can be devised.

References

Abramson, A., & Lisker, L. (1965). Voice onset time in stop consonants: Acoustic analysis and synthesis. *Proceedings of the 5th International Congress of Acoustics*. Liege: Imp. G. Thone.

Abramson, A., & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the sixth International congress of phonetic sciences*, Academia, Prague, pp 569-573.

Aslin, R.N., & Pisoni, D. B. (1980). Some developmental process in speech perception. In G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (Eds.), Child Phonology, vol.2: Perception, Academic press: NewYork, 1980.

Catherine, S. P. A., & Savithri, S. R. (2007). Effect of temporal variation on phoneme identification skills in typically developing children. *Proceedings of the National Symposium on Acoustics*, pp 193-198.

Doughty, J. (1949). In Zlatin, M., & Koenigsknecht, R. (1975). Development of the voicing contrast: Perception of stop consonants. *Journal of Speech and Hearing Research, 18*, 541-553.

Flege, J. E., & Eefting, W. (1986). Linguistic and developmental effects on the production and perception of stop consonants. *Phonetica*, 43, 155-171.

Pisoni, D. B. (1977). Identification and discrimination of the relative onset of two component tones: Implication for the perception of voicing in stops. *The Journal of the Acoustical Society of America*, 61, 1352-1361.

Sathya, K. (1996). *Development of auditory perceptual processing in children.* Unpublished doctoral thesis submitted to University of Mysore, Mysore.

Savithri, S. R. (1996). Speech perception in children: Temporal aspects. Research project funded by the Department of Science and Technology, Ministry of Science and Technology, New Delhi, India.

Simon, C., & Fourcin, A. (1978). Cross language study of speech pattern learning*. The Journal of the Acoustical Society of America*, 63, (3), 925-935.

Werker, J. H., & Tees, R.C. (1984). Cross-language speech perception. Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, 7, 49-63.

Winterkorn, J., MacNeilage, P., & Preston, M. (1967). Perception of voiced and voiceless stop consonants in three-year-old children. *Haskins Laboratries: Status Report on Speech Research*, SR-11, 41-45.

Williams, L. (1977a). The voicing contrast in Spanish. *Journal of Phonetics,* 5, 169-184.

Williams, L. (1977b). The perception of stop consonant voicing by Spanish-English bilinguals. *Perception and Psychophysics,* 21, 289-297.

Yeni-Komshian, G., Preston, M., & Cullen, J. (1967). A brief report on the study of imitation of synthetic speech stimuli by children. *Haskins Laboratries: Status Report on Speech Research*, SR-11, 48-49.

Zlatin, M.., & Koenigsknecht, R. (1975). Development of the voicing contrast: Perception of stop consonants. *Journal of Speech and Hearing Research, 18, 541-553.*

Acknowledgements

Authors would like to thank Dr. Vijayalakshmi Basavaraj, Director, All India Institute of Speech and Hearing for permitting us to conduct this study. We thank all the subjects who participated in this study.