LANGUAGE IN INDIA
Strength for Today and Bright Hope for Tomorrow
Volume 9 : 5 May 2009
ISSN 1930-2940

Effect of Temporal Variations on Phoneme Identification Skills in Children and Adults

A Comparative Study

S. Powlin Arockia Catherine, Ph.D. Candidate
S. R. Savithri, Ph.D.

Effect of Temporal Variations on Phoneme Identification Skills in Children and Adults
A Comparative Study

S. Powlin Arockia Catherine, Ph.D. Candidate
S. R. Savithri, Ph.D.

Abstract

The present study investigated the ability of 2 - 3 year old Kannada speaking children, and adults 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 a 21-year-old female native Kannada speaker were recorded and stored onto the computer memory.

VOT was truncated and /b – p/, and /g - k/ continuum were synthesized using Praat software which was audio recorded onto a CD.

The edited tokens were presented to thirty 2 - 3 year old children and thirty adults (18 - 28 years) individually. Each group consisted of 15 males and 15 females. Children were instructed to point to the pictures placed before them as they listen to the tokens and the investigator noted their responses on a binary forced-choice scoring sheet. Adults were asked to mark their responses on a binary forced-choice scoring sheet as they heard the synthetic token. Percent identification scores were calculated.

The results indicated that the 50% crossover from voiced to unvoiced cognate occurred in the lead VOT region for children and adults, and boundary width was wider for /g - k /continuum compared to /b - p/ continuum. Phoneme boundary width was wider in children compared to adults for /g - k/ continuum. The results of the current study can be used to compare phoneme identification skills in clinical population of same age.

Introduction

Speech perception refers to the processes by which humans are able to interpret and understand the sounds used in a language. Studies on speech perception have investigated the process involved in recognition of speech sounds and use this information to understand spoken language. Studies on infant speech perception postulate 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).

Modification of perceptual abilities takes place between infancy and adulthood. It is important to investigate and document these changes which would enlighten the perception- production relationship and to understand the perceptual abilities in clinical population. Phoneme perception ability is studied by altering the acoustic correlates that distinguishes speech sounds from one another.

Some Important Studies

In a seminal study, Lisker & Abramson, (1964) made acoustical measurements of voicing in initial stops in 11 languages. They defined voice onset time (VOT) as the time duration between the articulatory release in a stop consonant and the onset of voicing. Stop consonants in these languages were classified as long lead, short lag and long lag voicing distinctions based on VOT. In various perceptual experiments, it was found that VOT can effectively distinguish voicing in stop consonants (Yeni-Komishian, Preston & Cullen, 1967; Abramson & Lisker, 1970; Zlatin & Koenigsknecht, 1975; Savithri, 1996; Sathya, 1996) and subjects identify and discriminate differences in VOT categorically.

In the past, perception of stop consonants along VOT continuum has been studied in infants, children and adults. The early studies in children failed to see any 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.

Studies in the Indian Contexts

In the Indian context, Savithri (1996), Sathya (1996), and Catherine & Savithri (2007) reported developmental trend in VOT in Kannada and Telugu speaking children. However, the voicing contrast is language related.

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 is different.

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 and compared it with adults.

Method

Subjects: Thirty typically developing children in the age range of 2 - 3 years, and thirty adults in the age range of 18 - 28 years participated in this study. Each group included 15 males and 15 females. The children were from four play schools in Mysore and adults were bachelor degree students and volunteers. All the subjects were native Kannada speakers and from middle socio-economic status. The subjects were screened for their 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) with 16 kHz sampling rate and 16 bit rate and stored onto the computer memory. Mean voice onset time, phoneme duration and total word duration were measured 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 onset of burst were truncated from the waveform display of the words /bennu/ and /gere/ and were reduced in steps of 0.9 (factor) till the burst and then concatenated to the remaining part of the 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 for a /b - p/ VOT continuum (-77 to +40 ms) and 15 tokens for /g - k/ VOT continuum (-78 to +40 ms) were generated. The tokens were iterated thrice, randomized and recorded onto a CD. Thus a total of 90 tokens (15*2*3) formed the stimulus.

Procedure: Subjects were tested individually and were familiarized with the pictures of the words used in the experiment before the day of data collection. Experiment was carried out in a quiet room. Each participant was seated comfortably in a chair. The stimulus was audio-presented through two speakers placed at 45^o azimuth at a comfortable loudness.

Each subject was instructed to carefully listen to the stimuli. In the case of children, each child was asked to point to the respective picture out of four picture cards placed before him/her.

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 investigator noted the responses of children on an alternate forced choice scoring sheet. Adults were instructed to mark ✓ under the respective column on an alternate forced choice response sheet on listening to the tokens.

Percent identification scores were calculated and identification curves were drawn. Fifty percent 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.

Results

Fifty percent crossover: Fifty percent crossover occurred in lead VOT region for both /b - p/ and /g - k/ continuum in children and adults. In /b - p/ continuum, the shift from voiced to unvoiced occurred at shorter lead VOT in children (-8.3 ms) compared to that in adults (-26.5 ms). But, for /g - k/ continuum, the shift occurred earlier in children (-10.3 ms) compared to that in adults (-1.4 ms). Within groups, independent t test revealed significant gender difference only in adults {/b - p/ continuum [$t(28) = 1.9$; $p < 0.05$]; /g - k/ continuum [$t(28) = 1.9$; $p < 0.05$]}.

Multivariate analysis revealed significant difference in the mean scores between children and adults for /b - p/ [$F(1, 56) = 24.3$; $P < 0.001$] and /g - k/ [$F(1, 56) = 5.04$; $P < 0.05$] continuum. Table 1 shows the mean and standard deviation values of 50% crossover for children and adult. Figures 1 and 2 show average percent identification score in children and adults for /b - p/ and /g – k/ continuum.

Subjects	Children			Adults		
	Male	Female	Average	Male	Female	Average
Bilabial	-6.4 (9.2)	-10.2 (12.07)	-8.3 (10.7)	-20.4 (15.8)	-32.7 (18.2)	-26.5 (17.9)
Velar	-6.4 (14.8)	-14.3 (12.1)	-10.3 (13.9)	4.8 (11.7)	-7.6 (21.2)	-1.4 (18)
Average	-6.4 (12)	-12.29 (12.08)	-9.3 (12.3)	-7.8 (13.7)	-20.1 (19.7)	-13.9 (17.9)

Table 1: Mean (in ms) and standard deviation (in parenthesis) values of 50% crossover for VOT in children and adults.

(a) (b)

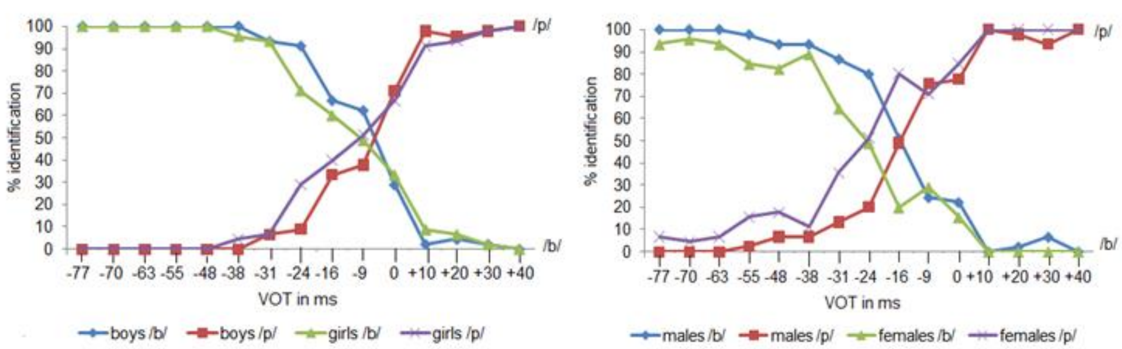

Figure 1: Percent identification scores of children (a) and adults (b) on /b/-/p/ continuum.

(a) (b)

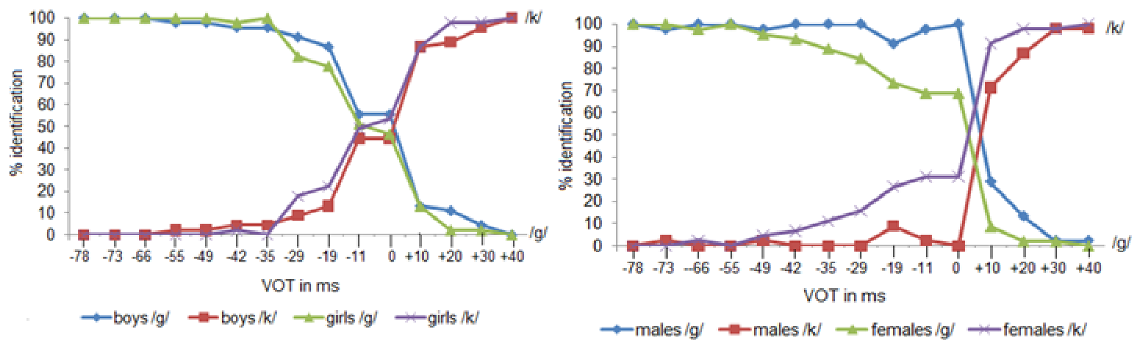

Figure 2: Percent identification scores of children (a) and adults (b) on /g/-/k/ continuum.

Lower Limit of Phoneme Boundary (LLPB): Lower limit of phoneme boundary was in the lead VOT region in both children and adults. For /b - p/ continuum, LLPB occurred later in children (-13.9 ms) compared to adults (-30.5 ms). For /g - k/ continuum, LLPB occurred earlier in children (-15.8 ms) compared to adults (-6 ms). Independent t test within the groups revealed no significant gender difference in children for both continuum, and in adults for /g - k/ continuum.

Significant gender difference was observed in adults for /b - p/ [t (28) =2.2; $p < 0.05$] continuum.

Multivariate analysis revealed significant difference between adults and children for bilabial [F (1, 56) =19; $P < 0.001$] and velar [F (1, 56) =6.2; $P < 0.05$] continuum. Table 2 shows mean and standard deviation values of LLPB for VOT in both groups.

Subjects	Children			Adults		
	Male	Female	Average	Male	Female	Average
Bilabial	-10.1 (11.7)	-17.6 (12.1)	-13.9 (12.3)	-23.6 (16.0)	-37.4 (18)	-30.5 (18.2)
Velar	-12.3 (14.1)	-19.3 (12.2)	-15.8 (13.4)	-0.2 (10.8)	-11.9 (21.1)	-6.0 (17.5)
Average	-11.2 (12.9)	-18.4 (12.1)	-14.8 (12.8)	-11.9 (13.4)	-24.6 (19.5)	-18.2 (17.8)

Table 2: Mean (in ms) and standard deviation (in parenthesis) values of LLPB for VOT in children and adults.

Upper Limit of Phoneme Boundary (ULPB): The upper limit of phoneme boundary occurred at lead and short lag VOT for /b - p/ and /g - k/ continuum, respectively in both the groups. Upper limit of phoneme boundary occurred later in children (-2.8 ms) compared to adults (-16.7 ms) for /b - p/ continuum. Also, ULPB occurred earlier in children (3.9 ms) and in the lag VOT region compared to adults (7.7 ms) for /g - k/ continuum. Within children, no significant difference between genders was evident for either of the continuum. However, independent t test revealed significant difference between genders in adults for /g - k/ continuum [t (28) =2.3; $p < 0.05$].

Multivariate analysis indicated significant difference between groups on ULPB for /b - p/ continuum [F (1, 56) =11.7; $p < 0.05$]. Mean and SD of ULPB of VOT in children and adults are in table 3.

Subjects	Children			Adults		
	Male	Female	Average	Male	Female	Average
Bilabial	-2.4 (11.4)	-3.13 (15.1)	-2.8 (13.2)	-13.6 (19.3)	-19.8 (15.8)	-16.7 (17.6)
Velar	4.06 (15.1)	3.8 (12.6)	3.9 (13.7)	13.8 (7.2)	1.53 (19)	7.7 (15.5)
Average	0.8 (13.2)	0.3 (13.8)	0.5 (13.4)	0.1 (13.2)	-9.1 (17.4)	-4.5 (16.5)

Table 3: Mean (in ms) and standard deviation (in parenthesis) values of ULPB for VOT in children and adults.

Phoneme Boundary Width: Phoneme boundary width was almost similar in children (12.5 ms) and adults (13.6 ms) for /b - p/ continuum. For /g - k/ continuum, children had wider PBW (19.6 ms) compared to adults (12.6 ms). Multivariate analysis revealed significant difference between groups for /g - k/ continuum [F (1, 56) =5.9; $p < 0.05$]. No significant within group gender difference was evident. Table 4 shows mean and standard deviation of PBW in both groups.

Subjects	Children			Adults		
	Male	Female	Average	Male	Female	Average
Bilabial	10.4 (7.6)	14.5 (7.7)	12.5 (7.8)	11 (6.6)	16.2 (12.3)	13.6 (10.1)
Velar	18.2 (12.2)	21 (15)	19.6 (13.5)	12.8 (8.1)	12.4 (7.2)	12.6 (7.5)
Average	14.3 (9.9)	17.7 (11.3)	16 (10.6)	11.9 (7.3)	14.3 (9.7)	13.1 (8.8)

Table 4: Mean (in ms) and standard deviation (in parenthesis) of PBW for VOT in children and adults.

Discussion

The results revealed several interesting findings. Firstly, 50% cross over occurred in lead VOT region for /b-p/ and /g-k/ continuum in children (-9.3 ms) and adults (-13.9 ms). This implies that children and adults shifted their percept from voiced to unvoiced stop consonant in the lead VOT region.

This result is in accordance with the findings of Abramson & Lisker (1965) in Thai speaking children (-20 ms), Williams (1977) in Spanish speaking children (-4 ms), Savithri (1996) in Kannada children (-8 ms) and adults (-10.7 ms), Sathya (1996) in Telugu children (-10 ms) and adults (-20 ms) and Catherine & Savithri (2007) in Kannada speaking children (-22 ms).

The result contradicts that in English language [Winterkorn, MacNeliage & Preston (1967), Yeni-Komshian, Preston & Cullen (1967), Simon (1974), Zlatin & Koenigsknecht (1975), Simon & Fourcin (1978), Williams (1977a, 1977b) and Flege & Eefting (1986)] which reported 50% crossover in lag VOT region.

This difference in the result may be due to the distinct phonemic structure and acoustic features of stop consonants of the two languages. Stop consonants in English have two way (i.e. p, b) classifications, whereas Kannada has four way (i.e. p, ph, b, bh) classification.

The distinction between unvoiced aspirated and unvoiced unaspirated stop is phonetic in English but phonemic in Kannada (Savithri, 1996). And spectrographically, comparison of stop consonants of English [as explained by Williams (1980)] and Kannada (Savithri, 1996) indicated that aspiration and F1 cutback in lag VOT region cues bilabial unvoiced stop consonant in English but not in Kannada, which might be responsible for the 50% crossover in lag VOT region in English.

Also, Kannada listeners gave greater weight to prevoicing as a cue to voiced stop consonants that was similar to Spanish, Thai and Telugu listeners. Cross language studies by Williams (1977b), Simon & Fourcin (1978), and Flege & Eefting (1986) indicate that speakers of different languages may perceive stop consonants differently as they are exposed to different kinds of stops.

Secondly, the average phoneme boundary was within $\pm$20 ms in children (-9.3 ms) and adults (-16.3 ms). Similar results are reported in Thai (Abramson & Lisker, 1965), Spanish (Williams, 1977), Kannada (Savithri, 1996; Catherine & Savithri, 2007) and Telugu (Sathya, 1996). According to Pisoni (1977), it is the psychophysical process which is probably responsible for categorical discrimination within $\pm$20 ms and not due to phonetic categorization. The probable reason for this high discrimination within this narrow region of ±20 ms may be due to general sensory constraints on the auditory system to resolve small difference in temporal order (Aslin & Pisoni, 1980).

Third, phoneme boundary occurred earlier in /b - p/ continuum compared to /g - k/ continuum in both groups. This trend probably reflects the pattern of phoneme acquisition in which bilabial stop consonants are acquired prior to velar stop consonants. This pattern was also reported by Zlatin & Koegnigsknecht (1975), and Savithri (1996).

Fourth, phoneme boundary width decreased in adults compared to children for velar continuum. This result is in consonance with Zlatin & Koegnigsknecht (1975) and Savithri (1996). This probably indicates more confusion in children compared to adults.

Fifth, there were no significant gender differences in children; but gender difference was significant in adults. Females had earlier 50% crossover, LLPB and ULPB compared to males. This implies that phonetic processing abilities are similar in boys and girls in childhood and might differ in adults. The mechanism responsible for such a difference in phonetic processing abilities in males and females (adults) are yet to be investigated. This contradicts Zlatin & Koenigsknecht (1975) Savithri (1996), and Sathya (1996) who reported no significant difference between genders in adults.

To conclude, typically developing Kannada speaking children in the age range of 2-3 years and adults shifted their perception from voiced to unvoiced stop consonant in the lead VOT region. Future research on older age groups to study the developmental pattern of speech perception in Kannada and other languages is warranted.

Conclusions and implications

The present study compared the phoneme identification skills in 2 - 3 year old Kannada speaking typically developing children, and adults. It provided basic knowledge on the pattern of phoneme identification skills in children and adults. 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. 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.

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.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in word-initial stops: Acoustic measurements. *Word*, 20, 384-422.

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. (1974). *Some aspects of development of speech production and perception in children. Speech Communication Seminar*, Stockholm, John Wiley and Sons, London.

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.

Williams, L. (1980). In G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (Eds.), Child Phonology, Vol.2: Perception, Academic press: NewYork

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, 47-48.

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

Colophon

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 participated in this study.

S. Powlin Arockia Catherine, Ph.D. Candidate
Department of Speech-Language Sciences
All India Institute of Speech and Hearing
Mysore - 570 006, India
arockia_cath@yahoo.com

S. R. Savithri, Ph.D.
Department of Speech-Language Sciences
All India Institute of Speech and Hearing
Mysore - 570 006, India