Effect of Spectral 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 typically developing children to identify synthetic phonemes varying in second formant frequency (F_2). Two picturable (minimal pair) words with stop consonants contrasting in place of articulation (labial /p/ and retroflex /ṭ/) 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. Place of articulation continuum was prepared using Acophon 1 programme of SSL Pro3V3 software and tokens were audio recorded onto a CD. The synthesized tokens were presented to thirty (15 boys and 15 girls) 2 to 3 year old children individually. Subjects were instructed to point to the pictures placed before them as they listen to the token and the investigator noted their responses on a scoring sheet. Percent identification scores were calculated. The results indicated that the 50% crossover from labial to retroflex cognate occurred at 1692 Hz in children. Also, phoneme boundary width was wider in girls compared to that in boys. The results of the current study can be used to compare phoneme identification skills in clinical population of same age.

Key words: Identification, Synthesis, F2 transition.

Speech perception is the decoding and interpretation of speech by the listener. 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 have 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 the modification process in phoneme perception during language development period, as it would strengthen our understanding of perception-production relationship.

Of all phonemes, stop consonants are widely studied. They are produced by occluding the oral cavity and releasing the articulator after sufficient air pressure is built up. Temporal and acoustic parameters cue voicing and place of articulation in stop consonants. Cues for place of articulation include burst spectrum at consonantal release (Cooper, Delattre, Liberman, Borst & Gerstman, 1952; Winitz, Scheib & Reeds, 1972) and onset frequency of second and third formants (Delattre, Liberman, Cooper, 1955). Potter, Kopp & Green (1947) stressed the importance of second formant (F2) transition to cue place of articulation and described various pattern of transitions. Rising F2 transitions are perceived as bilabials (/b/ and /p/), F2 slight fall for front vowels and sharp fall for back vowels are perceived as alveolars (/d/ and /t/) and sharp F2 fall for front vowels and slight fall of F2 for back vowels are perceived as velars (/g/ and /k/) (Borden & Harris, 1980).

In the past, several investigations (Liberman, Delattre, Cooper and Gerstman, 1954; Liberman, Harris, Hoffman & Griffith, 1957; Sussman, 1993) have reported the importance of F_2 onset in identifying and discriminating place of articulation

[1]Junior Research Fellow, Dept. of Speech-Language Sciences, All India Institute of Speech and Hearing, Mysore, email:arockia_cath@yahoo.com, [2]Professor, Dept. of Speech-Language Sciences, All India Institute of Speech & Hearing, Manasagangothri, Mysore 570 006, email:savithri_2k@yahoo.com,

of stop consonants in adults and children. However, the F_2 varies depending on the place of articulation of the stop consonant in a language. Also, most of the studies are in English. India being a multilingual country offers great potential for research in this area. While this area has received some attention at the international level, the scene at the national level is dismal. Although the sources of data in non-Indian languages are useful, there is a pressing need for experimental evidence in various Indian languages. Both cross-sectional and longitudinal studies are needed with groups of children of sufficient size to allow at least preliminary generalizations about normative development.

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 by systematically varying the F2 onset frequency.

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 words with stop consonants in the initial position in Kannada were selected. The word initial stop consonants in the word pair contrasted in place of articulation (p- ṭ; pa:ṭa – ṭa:ṭa). These 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. From the wide-band bar type spectrograms with LPC analysis, the onset frequency of first three formants, F2 transition duration and F2 frequency of steady state of vowel for the initial vowel of the word pair was obtained. For /pa/, onset frequency of F1, F2 and F3 were 785 Hz, 1300 Hz and 2960 Hz, respectively and F2 transition duration was 37 ms followed by F2 steady state of 130ms. Formant frequencies of vowel steady state were 890 Hz (F_1), 1547 Hz (F_2) and 3030 Hz (F_3). F_2 continuum was prepared using formant based analysis by synthesis (FBAS) module of Acophon 1 programme in SSLPro3V3 software. The word /pa: ṭa/ was analyzed with a block duration of 40 ms and block shift of 10 ms. Linear predictive coefficient was kept at 18 and pre-emphasis factor was 1. Following analysis, the F_2 continuum was synthesized using edit option by selecting glottal voice source pulse shape. F_2 onset frequency of word-initial vowel was varied from 1300 Hz to 2400 Hz in steps of 100 Hz keeping the other formants constant in the initial vowel. The F_2 was interpolated between onset and onset of vowel steady state and the word was synthesized. Figure 1 illustrates F_2 at mid and end points of the continuum. A total of 12 synthetic words for a /p-ṭ/ continuum were generated. The synthetic words were iterated thrice, randomized and recorded onto a CD. Thus a total of 36 tokens (12*3) formed the stimulus.

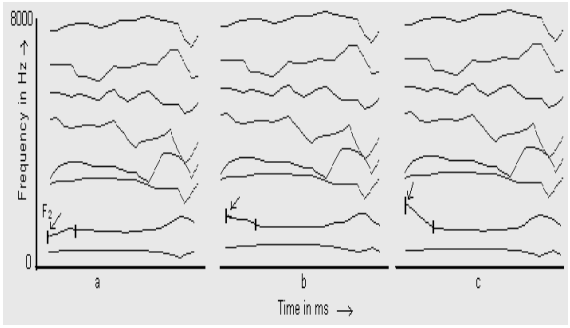

Figure1: Illustration of F_2 at 1300 Hz (a), 1800 Hz (b) and 2400 Hz (c).

Procedure: Children were tested individually. Each child was familiarized with the pictures of the words used in the experiment on the previous day of the data collection. Experiment was carried out in a quiet room. The child was seated comfortably in a chair. The stimulus was audio-presented through two speakers placed at 45° azimuth at a comfortable loudness. The child was instructed to carefully listen to the stimuli and point to the respective picture out of two picture cards placed before him/her. The investigator noted the child's responses on a scoring sheet. 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 labial (alveolar) category. Lower limit of phoneme boundary width is the point along the acoustic cue

continuum where an individual identified labial (alveolar) stop 75% of the time and upper limit of phoneme boundary width defined as the corresponding point of the identification of the labial (alveolar) cognate 75% of the time. Phoneme boundary width was determined by subtracting the lower limit from upper limit of boundary width. Identification data obtained from one child (B9) was not considered for analysis as it didn't show crossover.

Results

Result indicated that 50% crossover occurred at 1680 Hz. The mean LLPB and ULPB were at 1520 Hz and 1920 Hz, respectively. The mean PBW was 400 Hz. In girls 50% crossover, LLPB and ULPB occurred earlier compared to those in boys. PBW was wider in girls compared to that in boys. Also, the range of all the measures was wider in girls compared to those in boys. Results of independent t-test revealed significant gender difference on LLPB [t (27) = -2.176; $p<0.05$]. No significant difference was found across gender for 50% crossover [t (27) = -1.980; $p>0.05$], ULPB [t (27) = -1.754; $p>0.05$] and PBW [t (27) = 0.462; $p>0.05$]. Table 1 shows the mean values of all parameters in children. Figures 2 and 3 shows mean percent identification in boys and girls.

Parameters	Boys	Girls	Average
50%crossover (Range)	1820 (590)	1620 (630)	1680 (650)
LLPB (Range)	1630 (610)	1485 (650)	1520 (650)
ULPB (Range)	1995 (470)	1880 (610)	1920 (660)
PBW (Range)	365 (240)	395 (460)	400 (430)

Table 1: Mean and range values of measures of /p- ṭ/ continuum (in Hz).

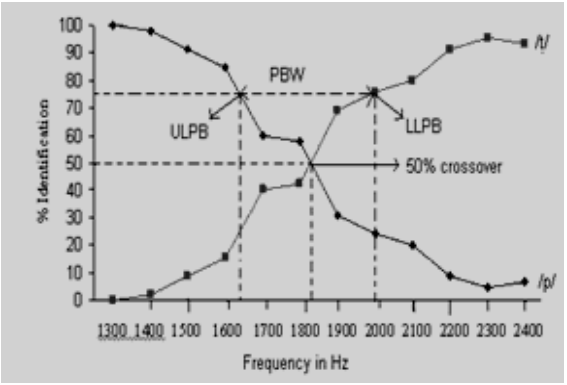

Figure 2: Mean percent identification scores in boys.

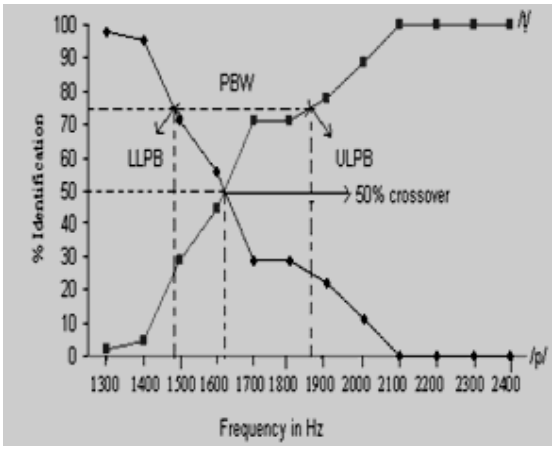

Figure 3: Mean percent identification scores in girls.

Discussion

The results revealed interesting findings. First, it was found that F2 onset frequency was an important cue to identify the place of articulation of stop consonants (/p/ and /ṭ/) in Kannada. Lower F_2 onset was perceived as bilabial /p/ and the percept changed to retroflex /ṭ/ at higher F_2 onset frequencies. This finding is in consonance with Liberman, Delattre, Cooper & Gertsman (1954) and Liberman, Harris, Hoffman & Griffith (1957) in adult English speakers.

Second, children shifted their percept from bilabial /p/ to retroflex /ṭ/ at 1680 Hz. In English speakers the shift from bilabial to alveolar place occurred at a much earlier frequency, 1199 Hz for 4 year old children, 1120 Hz for 5 – 6 year old children and 1144 Hz for adults (Sussman, 1993). This difference in shift of percept can be attributed to differences in place of articulation of stop consonants or methodological difference. In the present study the extreme frequencies of the continuum were between 1300 Hz and 2400 Hz but in Sussman's study they were 543 Hz and 1620 Hz. Also, stimulus in the current study was bisyllable meaningful Kannada words uttered by 21 year old female native Kannada speaker and F_2 continuum was prepared by manipulating the F_2 onset in initial vowel using FBAS module of Acophon 1 programme in SSLPro3V3 software. In Sussman's study the stimulus was synthetically generated using Klatt synthesis package using digital signal processing board (Data Translation model 2821).

Third, mean phoneme boundary width was 400 Hz in children and it was wider in girls (395

Hz) compared to boys (365 Hz). This can be attributed to more number of girls having wider PBW compared to that in boys.

To conclude, typically developing Kannada speaking children in the age range of 2-3 years were able to shift their percept from bilabial /p/ to retroflex /ʈ/ when F2 onset frequency increased. Future research on older age groups and adults to study the developmental pattern of speech perception in Kannada and other languages are warranted.

Conclusions

The present study provides data 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 like 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 children in this age group can be devised.

References

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.

Borden, G. J., & Harris, K. S. (1980). *Speech science primer- Physiology, acoustics and perception of speech.* Williams & Williams, Baltimore, USA, 171-214.

Cooper, F. S., Delattre, P.C., Liberman, A. M., Borst, J., & Gerstman, L. J. (1952). Some experiments on the perception of synthetic speech sounds. *The Journal of the Acoustical Society of America*, 24, 597-606.

Delattre, P. C., Liberman, A. M., & Cooper, F.S. (1955). Acoustic loci and transitional cues for consonants. *The Journal of the Acoustical Society of America*, 27, 769-773.

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.

Liberman, A. M., Delattre, D. C., Cooper, F.S., & Gerstman, L.J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. Psychological monographs, 68, 1-13.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358-368.

Potter, R. K., Kopp, G. A., & Green, H. (1947). *Visible speech*, NewYork: Van Nostran.

Sussman, J.E. (1993). Perception of formant transition cues of place of articulation in children with language impairments. *Journal of Speech and Hearing Research*. 36, 1286-1299.

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.

Winitz, H., Scheib, M. E., & Reeds, J. A. (1972). Identification of stop and vowels for the burst portion of /p,t,k/ isolated from conversational speech. *The Journal of the Acoustical Society of America*, 51, 1309-1317.

Acknowledgements

This study is an outcome of the doctoral research of the corresponding author. The authors are thankful to Dr. Vijayalakshmi Basavaraj, Director, All India Institute of Speech and Hearing for granting permission to carry out this study.