Association between Bandwidths of Vowel Formants and Age, Gender and Consonant Context in Telugu

Krishna Y., Ph.D. (Sp. & Hg.), CCC-A
Dr. B Rajashekhar, Ph.D.

==
Language in India www.languageinindia.com **ISSN 1930-2940 Vol. 13:3 March 2013**
==

Introduction

Communication in general can be considered as the process involved in generation, transmission, or reception of messages to oneself or another which is usually mutually understood set of signs. To understand human speech communication, one needs to have knowledge of the speech code and its' characteristics. The speech code differs based on linguistic rules of the language and organs involved in speech production. The speech code also differs based on the anatomical variation that exists between gender and age groups.

Human speech sounds are produced with the vocal cords vibration in the glottis and then transmitted into the vocal tract. The vibrations determine the fundamental frequency of the sound and the resonances in the vocal tract are known as formants (Pickett, 1996). Each format frequency thus produced has a varying bandwidth depending upon the spread of energy which is dependent on the oral structures, dampening of energy and place of articulation.

Vowels

Acoustically, vowels are characterized by formant pattern, spectrum, duration, bandwidth, amplitude and fundamental frequency. Among these, it is believed that, formant pattern, duration and fundamental frequency play a major role in the vowel perception (Pickett, 1980). Formant bandwidth is the difference in frequency between the points on either side of the peak which have amplitude that corresponds to 3 dB down from the peak.

Damping of the formants was first reported by Fletcher in 1929 and gave a bandwidth of 500 nepers per second, corresponding to 159 cps, for the first formant of natural vowel /a/. Dunn (1961) in his review reported that, there were considerable variations in the band widths obtained by various authors, and this could be due to the method they obtained and the context they used.

Resonant Frequency Bandwidth

Resonant Frequency Bandwidth estimation is very essential in understanding the quality of spoken vowels and vocal tract acoustics (Yasojima, Takahashi & Tohyama, 2006). Bandwidth can be calculated using different methods. Initially with the help of Fourier analysis using Oscillograms (Dunn, 1961) was calculated which questioned on its reliability and the accuracy. Later, Bogert (Dunn, 1961) introduced Bandwidth analysis using the sound spectrograph with sectioner. Another method used was curve fitting method, where true resonance curves of different widths were calculated and plotted in amplitude and frequency scales of spectrogram. However, its accuracy was questioned based on the calculations used for obtained curves (Dunn, 1961).

Bandwidth of the resonant frequency can also be estimated using Clustered Line-Spectrum Modeling (CLSM) (Yosida, Kazama & Toyama, 2001). In this model, the formant frequencies are subjected to analysis where in the bandwidth is calculated by decaying the signal. According to Dunn (1961) the Bandwidths varied from 39 cps to 130 cps for the first formant, 50 to 190 cps for the second formant and 70 to 260 cps for the third formant based on the technique used.

Male and Female Speakers

The bandwidths increased for both male and female speakers as the formant frequencies increased. The bandwidths for females also had greater variation and were wider compared to males (Yasojima et al., 2006). Formant bandwidth has little effect on the quality or intelligibility of isolated vowels (Klatt, 1982; Rosner & Pickering, 1994); however, it has an effect on the identification of vowels in competition with other vowels (Cheveigne, 1999). The author inferred that at constant root mean square amplitude, identification of a vowel is enhanced by sharpening its formants, or widening those of its competitor.

Telugu

Telugu, one of the four major Dravidian languages (the others being Kannada, Tamil and Malayalam), is the second most widely spoken language in India. Telugu has ten basic vowels, among which, five are short and five long (Duggirala, 2005). In modern Telugu, two short and long vowels (/{/, /{:/) have been introduced (Krishnamurti, 1961). In the published literature in Telugu language, studies have not studied the association between the bandwidth of the vowels and age, gender and consonant context they appear.

The aim of the present study, which is part of larger study (Krishna, 2009) is to understand the association between the bandwidth of the formant frequencies and age, gender consonant context they appear in Telugu language.

Method

Material

A list of 60 meaningful words (Krishna, 2009) consisting of all ten short and long vowels present in Telugu, in all possible preceding consonant and semivowel (CVCCV/CVCV) context was used. The target word was embedded in the final position of a carrier sentence "/i: padamu (target word) /" (This word is ______), so as to obtain reasonable uniform stress and intonation patterns (Bennett, 1981; Most, Amir & Tobin, 2000). The words were grouped based on manner and place of articulation and voicing feature of the preceding consonant. The influence of the first consonant on the fundamental frequency of the first vowel in the CVCV/CVCCV context was studied.

Participants

A total of 72 Telugu speaking normal individuals from three different regions (Coastal, Rayalaseema and Telengana) in three different age groups (Group I: 06 to 09 years; Group II: 13 – 15 years; Group III: 20 – 30 years) with equal gender ratio participated in the study. The mean age of Group I was 8 years, Group II - 13 years and Group III-23 years. All the participants were born in Andhra Pradesh and were native Telugu speakers. A qualified Audiologist and Speech-Language Pathologist evaluated and certified their speech, language, and hearing, as being normal at the time of data collection. The investigator has selected participants from three age groups, two gender (male and female) and three regions (Coastal, Rayalaseema and Telengana) so as to study the influence of age, gender and region.

Method

After an informed consent, the randomly selected participants from the respective groups were comfortably seated in a quiet room and speech sample was recorded using condenser microphone and Wave Surfer recording software. The participants were asked to read three times the sentence presented to them visually. Researcher selected perceptually best produced word, and was extracted using Adobe Audition 3 for the further analysis. The speech sample was recorded at a sampling rate of 22,050 kHz and bit rate of 256 kbps. A total of 4320 samples were analyzed using Computerized Speech Lab (CSL) 4500 for the analysis of fundamental frequency of the target vowel present in the target word. An anti-aliasing filter with a 10 kHz cutoff frequency was used before A/D conversion and a pre-emphasis factor of 0.8 was applied.

The target vowel was selected from its' onset to the offset. The onset and the offset of a vowel are determined by the presence and absence of clearly visible first two formants on the spectrogram respectively. (Krause, 1982; Gopal, 1987). Formant bandwidth was calculated from the spectrogram by identifying each format frequencies. Formant bandwidth values were recorded from the CSL.

The researcher re-measured 10% of the tokens (random selection) after 6 months of the first measure for intra-judge reliability. Results from the paired *t*-test suggest that the two measures are highly reliable ($t_{(431)} = 1.026$, $p = 0.309$). An experienced speech pathologist, unaware of the purpose of the study, measured temporal and spectral characteristics of 10% of the tokens (random selection) for inter-judge reliability. Results from the paired *t*-test suggest that the two measures are highly reliable ($t_{(431)} = 0.10$, $p = 0.920$).

Mean and standard deviation was used to summarize the variations in Bandwidths of the vowels. To evaluate the effect covariates and their association (age and gender) on bandwidth and to understand the influence of preceding consonant and individual variations on the fundamental frequency a multilevel approach (Quene & Bergh, 2004) was used. All analysis was carried out using SPSS 16 and MlWin 1.1.

Results and Discussion

The bandwidths of the formant frequencies F1, F2 and F3 (B1, B2 and B3 respectively) of all the short and long vowels are given in Table 1.

Table 1: Mean B1, B2 and B3 (Hz) and SD for all short and long vowels

N=4320

Vowel	B1		B2		B3	
	Mean	SD	Mean	SD	Mean	SD
/i/	58.55	18.59	133.97	17.53	229.9	18.54
/e/	56.39	17.04	137.73	16.04	230.26	17.62
/a/	57.71	20.19	137.11	15.04	228.3	17.84

/o/	57.78	19.05	134.75	16.48	230.37	16.63
/u/	56.69	19.25	136.9	17.1	229.98	19.34
/i:/	58.07	19.44	134.07	17	229.93	14.58
/e:/	55.51	16.95	137.63	17	231.93	17.44
/a:/	58.07	23.69	137.57	16.04	228.91	17.95
/o:/	56.95	19.18	136.12	18.42	227.36	20.06
/u:/	56.16	21.47	134.67	15.27	229.81	17.9

Scrutiny of the collected data revealed the front high short and long vowels /i/ and /i:/ had larger mean bandwidth B1 compared to front mid vowel /e/. The front mid short and long vowels /e/ and /e:/ had larger mean bandwidth B2 as compared to front high vowels /i/ and /i:/. The front high short and long vowels /o/ and /i:/ had larger mean bandwidth B3 as compared to vowels /a/ and /o/. Central vowels had larger mean bandwidth for F1 and F2 while front vowels had larger mean bandwidth for F3. It was also observed that, mean B1 and B2 varied with increase in the phonetic length of the vowel.

Further to study on the association between age, gender and region and it's kind with B1, B2 and B3 of the vowels studied were addressed by Random intercept model. The results for B1, B2 and B3 are given in Tables 2, 3 and 4 respectively.

Table 2: Statistical analysis using random intercept model for B1

N=4320

Covariates	**Estimate**	**Std. Error**	**Wald ratio**	**P value***
Constant (β_{0ij})	68.94	1.07	64.43	< 0.01
Age (β_{1ijk})	-4.89	1.01	-4.84	< 0.01
Gender (β_{2ijk})	0.52	0.13	4	< 0.01
Region (β_{3ijk})	-4.86	0.36	-13.5	< 0.01

Variance components	
Random Error:	374.95
Consonant Level:	0.00
Individual level:	0.00
Total variation:	374.95
*-2*log likelihood(IGLS) = 37416.36*	

**significant at 0.05 level*

Table 3: Statistical analysis using random intercept model for B2

N=4320

Covariates	**Estimate**	**Std. Error**	**Wald ratio**	**P value***
Constant (β_{0ij})	137.59	0.93	147.95	< 0.01
Age (β_{1ijk})	2.75	0.86	3.19	< 0.01
Gender (β_{2ijk})	-0.34	0.11	-3.09	< 0.01
Region (β_{3ijk})	-0.74	0.31	-2.39	0.01
Variance components				
Random Error:	269.19			
Consonant Level:	2.09			
Individual level:	0.13			
Total variation:	271.41			
*-2*log likelihood(IGLS) = 35594.57*				

**significant at 0.05 level*

Table 4: Statistical analysis using random intercept model for B3

N=4320

Covariates	**Estimate**	**Std. Error**	**Wald ratio**	**P value***
Constant (β_{0ij})	226.00	1.08	209.26	< 0.01

Age (β_{1ijk})	3.21	0.99	3.24	< 0.01
Gender (β_{2ijk})	-0.29	0.13	-2.23	0.01
Region (β_{3ijk})	0.78	0.36	2.17	0.02
Variance components				
Random Error:	316.91			
Consonant Level:	0.39			
Individual level:	0.02			
Total variation:	317.32			
*-2*log likelihood(IGLS) = 31227.03*				

**significant at 0.05 level*

From the results, it could be inferred that, there is a significant association between age, gender and region and bandwidths (B1, B2 and B3). Individual vowel variations within the region and gender groups are not significant; but tend to be significantly influenced with consonant and individual level variations.

Bandwidth increase with increase in formant frequency for all vowels in Telugu as noted in the current study has been reported (Dunn, 1961; Yosida, Kazama & Toyama, 2001; Yasojima, Takahashi & Tohyama, 2006). The changes noticed in bandwidths of the formant frequencies as the age progressed, could be attributed to the changes in vocal tract acoustics. Although bandwidth has a significant association with region, individual vowels did not show much significant differences. The significant association between bandwidth and region noted from multilevel analysis could be due to the preceding consonant and individual level influences and also measurement technique.

From the current study it can be concluded that minimal variations in bandwidths exist among the vowels. The front high short and long vowels /i/ and /i:/ had larger bandwidth for F1 compared to front mid vowel /e/. The front mid short and long vowels /e/ and /e:/ had larger bandwidth for F2 as compared to front high vowels /i/ and /i:/. The front high short and long vowels /o/ and /i:/ had larger bandwidth for F3 as compared to vowels /a/ and /o/. Central vowels had larger bandwidth for F1 and F2 while front vowels had larger bandwidth for F3. Bandwidths B1 and B2 varied with increase in the phonetic length of the vowel. There is significant association between age, gender and region and consonant has an influence on the bandwidth of the formants. This study emphasis the need for studying the bandwidth and establish normative values for each age, gender and region groups.

Conclusion

In the literature bandwidth of the vowel formant has always been studied not extensively. This could be due to varying in measurement techniques, and limited understanding of its association with individual factors and linguistic factors. This study was an attempt to study the association between the bandwidth and individual factors such as age, gender, region and consonant. A speech sample from 72 individuals was recorded and bandwidth of the formant frequencies F1, F2 and F3 were recorded using computerized speech lab 4500. To study the association and effect random intercept model was applied on the data. From the results it can be concluded that, bandwidth did vary between age, gender, region and consonant. It was also observed that, consonant has a significant effect on the bandwidth as compared to other individual factors.

==

References

Cheveigné, A. D. (1999). Formant Bandwidth affects the identification of competing vowels. from http://recherche.ircam.fr/equipes/pcm/cheveign/pss/1999_ICPhS.pdf

Duggirala, V. (2005). Development of Instructional materials for improving linguistic awareness of Telugu speaking Hearing Impaired children. (pp. 63 - 79). Department of Linguistics, Osmania University, Hyderabad: University Grants Commission, New Delhi.

Dunn, H. K. (1961). Methods of Measuring Vowel Formant Bandwidths. *Journal of Acoustical Society of America, 33*(12), 1737 - 1746.

Fletcher, H. (1929). *Speech and Hearing*. New York: D. Van Nostrand Company, Inc.

Gopal, H. S. (1987). Temporal aspects of tense and lax vowels in American English: Effects of speaking rate and post vocalic consonant voicing. Unpublished Dissertation. University of Texas.

Klatt, D. H. (1982). *Prediction of perceived phonetic distance from critical-bank spectra: a first step.* Paper presented at the IEEE ICASSP.

Krause, S. E. (1982). Developmental use of vowel duration as a cue to postvocalic stop consonant voicing. *Journal of Speech and Hearing Research, 25,* 388 - 393.

Krishna, Y. (2009). Acoustic characteristics of vowels in Telugu. *Languages in India*, 13(1). http://www.languageinindia.com/jan2013/krishnaacousticdissertation.html.

Krishnamurti, Bh. (1961). *Telugu Verbal Bases: A Comparative and Descriptive study* (Vol. 24). Berkeley and Los Angeles: University of California Press.

Pickett, J. M. (1996). *The Sounds of Speech Communication - A Primer of Acoustic Phonetics and Speech Perception*. Boston: Allyn and Bacon.

Pickett, J. M. (1980). *The sounds of speech communication a primer of acoustic phonetics and speech perception*. Baltimore: University Park Press.

Quene, H., & Bergh, H. V. D. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication, 43*(1-2), 103-121.

Rosner, B. S., & Pickering, J. B. (1994). *Vowel perception and production*. Oxford: Oxford University Press.

Yasojima, O., Takahashi, Y., & Tohyama, M. (2006). *Resonant Bandwidth Estimation of Vowels using Clustered-Line Spectrum Modeling for Pressure Speech Waveforms.* Paper presented at the 2006 IEEE International Symposium on Signal Processing and Information Technology, Vancouver, Canada.

Yosida, K., Kazama, M., & Toyama, M. (2001). Signal Analysis by Clustered Line-Spectrum Modeling. *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers), 101*(382(EA2001 62-73)), 31 - 36.

==

Krishna. Y. Ph.D. (Sp. & Hg.), CCC-A
Professor
Department of Speech & Hearing
Manipal College of Allied Health Sciences
Manipal University
Manipal – 576 104
Karnataka
India
krishna.y@manipal.edu

B. Rajasekhar, Ph.D.
Dean and Professor
Manipal College of Allied Health Sciences
Manipal University
Manipal – 576 104
Karnataka
India
b.raja@manipal.edu