BMC JOURNAL OF SCIENTIFIC RESEARCH
A Multidisciplinary Reviewed Research Journal
ISSN: 2594-3421

Variation of Formant Frequency with Nepalese Vowels

S. K. Adhikari*
E-mail: sheshkant@rediffmail.com

Abstract

The regions of speech spectrum in which the frequency corresponds to relatively large amplitude are known as formants. For any vocalic sounds, number of formants may occur in the frequency range 0 to 4000 Hz. The formant frequencies of speech sounds are directly depending up on the shape and size of vocal tract. The aim of study was to study the variation of formant frequency with Nepalese vowels. Ten Nepalese vowels word in initial position /VC/ as spoken three times by 10 male and 10 female Nepali speakers were recorded in system in the free field of partially acoustically treated room. PRRAT software is used to digitize and analyze the data. Linear predictive coding (LPC) spectra were obtained for each of vowels and formant frequencies were measured. By plotting curve between formant frequencies and vowels, explain their variation.

Keywords: *formant frequency, spectra, front vowels, central vowel, back vowel*

Introduction

Vowels are described in terms of relative position of the tongue the relative position of the lips (spread, round, unrounded), the position of soft plate and the phonemic length of vowel, the tenseness of the articulator and the relative pitch of the vowel[1]. Formant frequencies derived from analysis of natural speech are used to specify the formant pattern of synthetic vowels. Overall success of this approach might be taken to favor formant description, though not necessarily a description based only on static assumptions. A rough rule of thumb in relating the vowel articulation is that F_1 varies mostly with tongue height and F_2 varies mostly with tongue advancement [2]. Cautions should be followed in the use of this rule because there are exceptions.

In speech science a formant is the peak of the spectral envelope that results from an acoustic resonance of the human vocal tract [3]. In acoustic the definition of a formant sometimes differs as it can be defined as a peak, or local maximum, in the spectrum. For harmonic sounds, with this definition, the formant frequency is that of the harmonic partial that is augmented by a resonance.

The formant with the lowest frequency is called first formant frequency (F_1), the second higher than F_1 is called second formant frequency (F_2) and the third which is higher than F_2 is called third formant frequency F_3. Most often the two first formants, F_1 and F_2, are sufficient to identify the vowel [1]. Low vowels have a high F_1 frequency and high vowels have low F_1. Back vowels have a low F_2 and typically a small [1,4]. The first formant F_1 has a higher frequency for an open vowel and a lower frequency for a close

* Dr. Adhikari is Lecturer in Physics at Birendra Multiple Campus (Tribhuvan University)

vowel and the second formant F_2 has a higher frequency for a front vowel and a lower frequency for a back vowel [5,6]. Studies of the frequency spectrum of trained classical singers, especially male singers, indicate a clear formant around 3000 Hz (between 2800 and 3400 Hz) that is absent in speech or in the spectra of untrained singers. It is thought to be associated with one or more of the higher resonances of the vocal tract. This formant is actively developed through Vocal training, for instance through so-called "witch's voice"[7] exercises and is caused by a part of the vocal tract acting as a resonator [8]. It appears, then, that a vowels formant pattern can be used to identify a vowel and even to establish relationships between acoustic and perceptual parameters [8]. In most cases it is necessary to specify only the first three formants to achieve a good result. In addition, the formant patterns of vowels frequently are continuous with a formant pattern of neighboring consonants. Another advantage of formant description is that the formants typically are easily observed in acoustic analysis of speech.

Spectral Analysis of speech

Frequency analyses are important in production and perception and efficient coding. The vocal mechanism is a quasi-stationary source of sound. Its excitation and normal modes change with time. Any spectral measure applicable to the speech signal should therefore reflect temporal features of perceptual significance as well as spectral feature.

Short –Time frequency Analysis

The relation between a periodic time function f(t) and its complex amplitude-density spectrum F(ω) is the Fourier transform-pair [8].

$$F(\omega)=\int_{-\infty}^{+\infty} f(t)\, e^{-j\omega t}\, dt. \qquad (1)$$

The Fourier transform can be modified by transformation that part of the signal seen through the window at a given instant of time. The desired operation is

$$F(\omega, t) = \int_{-\infty}^{t} f(\lambda)\, h(t-\lambda)\, e^{-j\omega\lambda}\, d\lambda. \qquad (2)$$

If the weighting function h(t) is considered to have the dimension sec^{-1} then F (ω, t) is a short – time amplitude spectrum with the same dimension as the signal.

Like the conventional Fourier transform, F (ω, t) is generally complex with a magnitude and phase spectrum. By the definition, the inverse relation also holds.

$$f(\lambda)\, h(t-\lambda) = (1/2Л) \int_{-\infty}^{+\infty} F(\omega, t)\, e^{j\omega\lambda}\, d\omega. \qquad (3)$$

At any time, $t = t_1$ the product [f (λ) h (t- λ)] is determine for all $\lambda \leq t_1$. If the window function $h(t_1 - \lambda)$ is known, then the original function over the interval -∞ to $-\infty \leq \lambda \leq t_1$ can be retrieved from the product.

for $\lambda = t_1$.

$$f(t)\, h(0) = (1/2Л) \int_{-\infty}^{+\infty} F(\omega, t)\, e^{j\omega\lambda}\, d\omega. \qquad (4)$$

The short-time transformation is therefore uniquely invertible if one non zero value of the window function is known. Typically, h(t) can be chosen so that

H (0) = 1 and

$$f(t) = (1/2Л) \int_{-\infty}^{+\infty} F(\omega,t)\, e^{j\omega\lambda}\, d\omega\ . \qquad (5)$$

This inversion implies that f(t) can be determined for the same points in time that F (ω, t) is known, provided F (ω, t) is known as a continuous function of frequency.

Formant Analysis of Speech

Formant analysis of speech can be considered a special case of spectral analysis [8]. The objective is to determine the complex natural frequencies of the vocal mechanism as they change temporally.

If the band width of the resonance is relatively small then the first moment of amplitude (A) spectrum which estimate the imaginary part of the pole frequency is

$$f = \frac{\int f A(f)df}{\int A(f)df} \qquad (6)$$

As the resonances of the vocal tract are multiple, the output time wave form is therefore a superposition of damped sinusoids and the amplitude spectrum generally exhibits multiple peaks which are notified as formant frequencies.

The n^{th} moment of an amplitude spectrum A (ω) is

$$M_n = \int \omega^n \, A(\omega) \, d\omega \qquad (7)$$

where ω is the radian frequency. If a suitable pre-filtering or partitioning of the spectrum can be made, then a formant frequency can be approximated by

$$\omega = \frac{M_1}{M_0} = \frac{\sum \omega_i A(\omega_i)}{\sum A(\omega_i)} \qquad (8)$$

A number of formant measures based up on this principle have been examined. The spectral partitioning problem remains of considerable importance in the accuracy of these methods.

Subjects: Ten subjects (10 males and 10 females) of 20-25 (adult) were considered. These 20 subjects were selected on the criteria that they had Nepali dialect as their mother tongue, able to read Nepali and normal speech, language and hearing function. Subjects are college students of different level of University which are belongs to different districts (Chitwan, Nawalprasi, Lamjung and Gorkha) of Nepal.

Test materials: In the present study, the ten vowels of Nepali language i.e. /Λ/, /a/, /I/, /i /, /U/, /u /, /e/, / є/, /O/ and / Ɔ /were analyzed. The vowels were in the environment of 15 consonants / p, p^h, b, b^h; t, t^h, d, d^h; ṭ, $ṭ^h$, ḍ; h , ; k, k^h, g, g^h / the test material consist of a list of 150 meaningful monosyllabic word with a /VC/formant. Each test word consisted of one of the 10 vowels as V and one of the 15 consonants as C, and each of these sentences was written on flash cards.

Procedure: The subjects were instructed to read, the token written on the flash card as naturally as possible. The recording was done in partial acoustically treated room for individual subjects by presenting one flash card at a time using a SANYO voice activated recording system (TRC- 860C). This was connected to the computer (Pentium IV) having SOFT-WARE OF PRRAT. The speech signal was digitized at the sampling frequency of 16 KHz.

Observation

Table: - 1: Average value of first formant frequencies (F_1) in Hz of Nepali vowels for male and female.

Gender	Front vowel				Central vowel		Back vowel			
	/i/	/I/	/ɛ/	/e/	/a/	/Λ/	/Ɔ/	/O/	/U/	/u/
Male	410	422	498	517	765	682	530	520	463	455
Female	451	485	588	611	868	792	599	583	498	475

Table: - 2: Average value of second formant frequencies (F_2) in Hz of Nepali vowels for male and female.

Gender	Front vowel				Central vowel		Back vowel			
	/i/	/I/	/e/	/ɛ/	/a/	/Λ/	/Ɔ/	/O/	/u/	/U/
Male	2532	2413	2084	1999	1517	1489	1345	1256	1193	1182
Female	2701	2624	2256	2187	1817	1776	1545	1506	1393	1382

Table: - 3: Average value of third formant frequencies (F_3) in Hz of Nepali vowels for male and female.

Gender	Front vowel				Central vowel		Back vowel			
	/i/	/I/	/e/	/ɛ/	/Λ/	/a/	/O/	/Ɔ/	/u/	/U/
Male	3393	3166	3055	3033	2643	2631	2467	2378	2311	2207
Female	3593	3354	3221	3103	2834	2801	2666	2518	2491	2406

Result and discussion

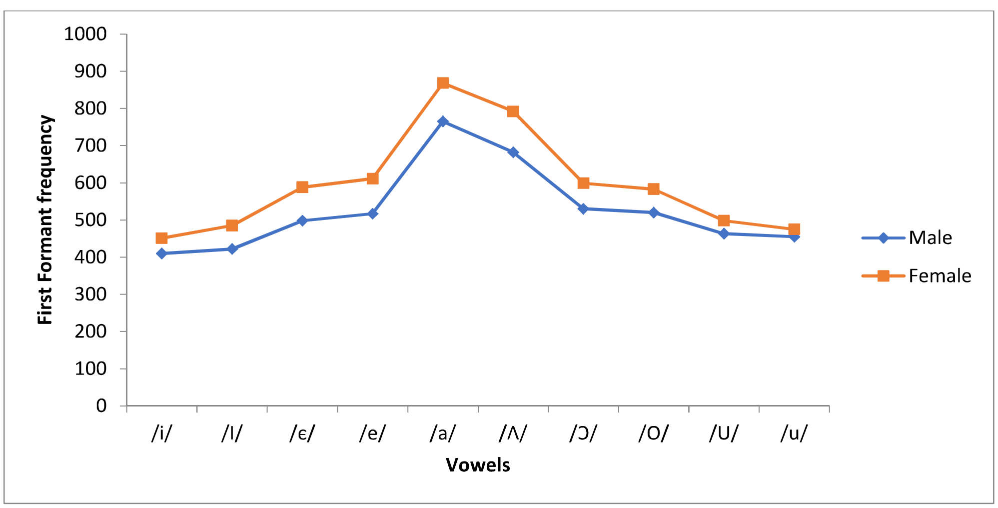

Fig. 1: Variation of first formant frequencies F_1 of different vowels for male and female.

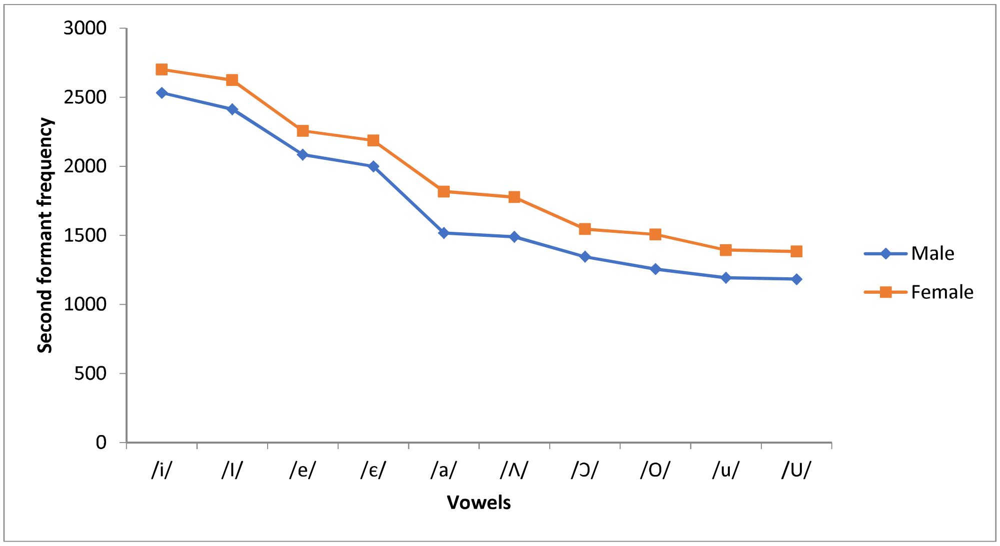

Fig. 2: Variation of second formant frequencies F_2 of different vowels for male and female.

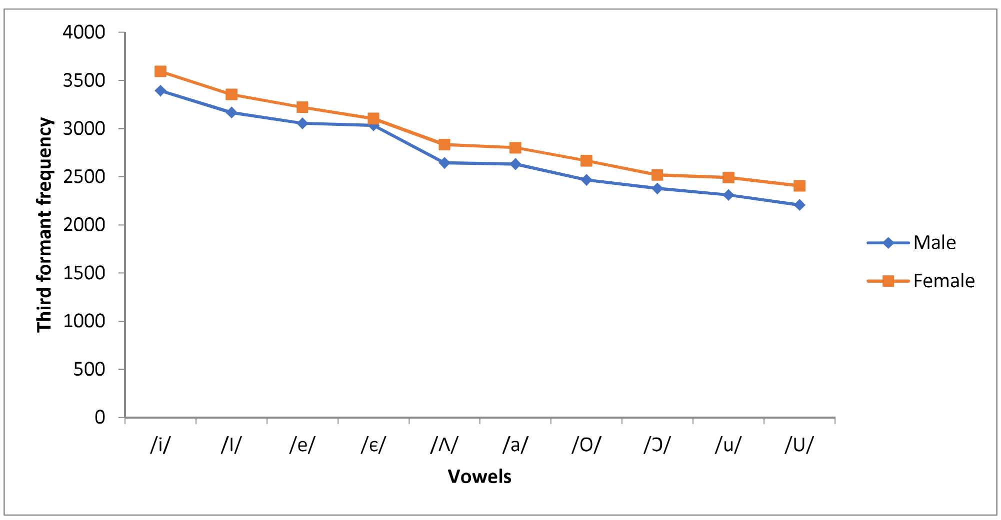

Fig. 3: Variation of third formant frequencies F_3 of different vowels for male and female

First formant frequency (F_1)

From table1, it found that central vowel /a/ and /Λ/ have high F_1 having value 765Hz and 682Hz for male and 868 Hz and 792 Hz for female. Front vowel /i/, /I/ have low F_1 having value 410 Hz and 422 Hz for male and 451 Hz and 485 Hz for female respectively.

In fig. 1, the curve of F_1 increases from front vowel /i/, /I/, /e/, /є/ to central vowel /a/, / Λ/ having maximum F_1 and decrease to back vowel /O/, /Ɔ/, /U/, /u/ in case of both male and female speaker.

Second formant frequency (F_2) : From table 2, it found that the front vowel /i/, /I/ have high F_2 having value 2532 Hz and 2413 Hz for male and 2701 Hz and 2624 Hz for female. The back vowel /U/, /u/, /O/, / Ɔ/ have lowest F_2 having value 1182 Hz,1193 Hz,1256 Hz,1345 Hz for male and 1382 Hz, 1393 Hz,1506 Hz,1545 Hz for female respectively.
As the front vowel /I/, /i/, /e/, /є/ have high F_2 so that curve decrease from front vowel to central vowel /a/, /Λ/ and to back vowel /Ɔ/, /O/, /u/, /U/ respectively for both male and female in fig 2.

Third formant frequency (F_3) : From table 3, it found that the front vowel /i/, /I/ have high F_3 having value3393 Hz and 3166 Hz for male and 3593 Hz and 3354Hz for female. The back vowel /U/, /u/, / Ɔ/, /O/ have lowest F_3 having value 2207 Hz, 2311 Hz, 2378 Hz, 2467 Hz for male and 2406 Hz, 2491 Hz, 2518 Hz, 2666 Hz for female respectively.
As the front vowel /I/, /i/, /e/, /є/ have high F_3 so that curve decrease from front vowel to central vowel /a/, /Λ/ and to back vowel /Ɔ/, /O/, /u/, /U/ respectively for both male and female in fig.3.

Conclusion

From above discussion it concludes that the F_1 increases from front vowel /i/, /I/, /e/, /є/ to central vowel /a/, / Λ/ and decrease to back vowel /O/, /Ɔ/, /U/, /u/ in case of both male and female speaker.
F_2 decrease from front vowel /i/, /I/, /e/, /є/ to central vowel /a/, /Λ/ and back vowel /Ɔ/, /O/, /u/, /U/ respectively for both male and female.
F_3 decrease from front vowel /i/, /I/, /e/, /є/ to central vowel /a/, /Λ/ and back vowel /Ɔ/, /O/, /u/, /U/ respectively for both male and female.

References

1. Gunnar, Fant (1970). “Acoustic Theory of Speech Production”, Mouton The Hague Paris”.
2. Ilse Lehiste (1974). “Acoustic Phonetics”, The M. I. T. Press, ISBN 262 -12025- 9.
3. Titze, I.R., Baken, R.J. Bozeman, K.W., Granqvist, S. Henrich, N., Herbst, C.T., Howard, D.M., Hunter, E.J., Kaelin, D., Kent, R.D., Löfqvist, A., McCoy, S., Miller, D.G., Noé, H., Scherer, R.C., Smith, J.R., Story, B.H., Švec, J.G., Ternström, S. and Wolfe, J. (2015). “Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization”. J. Acoust. Soc. America. 137, 3005–3007.
4. Nuzhakufran, Rajani Sebastian and Savithri S.R. (2004). “Acoustic characteristics of Sinhala Vowels”. J. Acoust. Soc. Ind. 32, 233-236.
5. Ladefoged, Peter (2006). “A Course in Phonetics” (Fifth Edition), Boston, MA: Thomson Wadsworth, ISBN 1-4130-2079-8.
6. Ladefoged, Peter (2001). Vowels and Consonants: An Introduction to the Sounds of Language, Maldern, MA: Blackwell, ISBN 0-631-21412-7.
7. Frisell, Anthony (2007). Baritone Voice. Boston Branden Books. ISBN 0-8283-2181-7
8. James L.Flanagan (1972). “Speech analysis, Synthesis and Perception”. Springer-verlag, Berlin, Heidelberg, New York ISBN 3-540-05561-4.