Original Research Article

Variation of Fundamental Frequency of Nepali Vowels with Voiced Consonant in /VC/ Syllables

S. K. Adhikari[1] and Dipak Raj Adhikari[2,*]

[1]Department of Physics, Birendra Multiple Campus, Tribhuvan University, Nepal

[2]Central Department of Physical and Mathematical Sciences, Graduate School of Science & Technology, Mid-West University, Nepal

[*]Corresponding author E-mail: dipak74adhikari@gmail.com

Received: 10th June 2026, Revised: 18th September 2026, Accepted: 25th September 2026

Abstract

Fundamental frequency is an acoustic property of sound which measures the number of complete vibrations. This paper investigates the measurement of variation of fundamental frequency of Nepali vowels word in /VC/ syllables which is followed by voiced consonants sound as spoken three times by ten males and ten female Nepali speakers at age between 15 to 20 years old. Praat software was used to digitize and analyze the recorded data. The fundamental frequency for aspirated and unaspirated voiced sounds was calculated and it is found that variation of fundamental frequency in male speaker in aspirated of voiced (Avg. 212.65 Hz, SD = 8.68) is lower than in un-aspirated voiced (Avg. 226.80 Hz, SD = 7.61) with ratio (Avg. 0.94). Similarly, the average variation of fundamental frequency of female speaker in aspirated of voiced (Avg. 251.73 Hz, SD = 7.32) is lower than in unaspirated of voiced (Avg. 267.72 Hz, SD = 7.87) with ratio (Avg. 0.94). From these data, it was also found that the average fundamental frequency of male speaker is lower than female speakers in aspirated and unaspirated of voiced consonants. The measurement of fundamental frequency of vowels with voiced consonants is very important as it helps to identify the vowels of different languages as it gives the phonetic identity of vowel software made by the speech identification techniques are used in different fields of sound.

Keywords: Fundamental frequency, Nepali language, Vowel, Consonant

Introduction

The fundamental frequency (F_0) of a periodic signal is known as the inverse of its period, can be defined as the smallest positive member of the infinite set of time shifts that leave the signal invariant. This definition applies strictly only to a perfectly periodic signal. The subjective pitch of a sound usually depends on its fundamental frequency, but there are some exceptions. Sounds may be periodic yet "outside the existence region" of pitch [1,2]. Conversely, a sound may not be periodic [3,4]. However, over a wide range pitch and period are in a one-to-one relation, to the degree that the word "pitch" is often used in the place of F_0 and F_0 estimation methods are often referred to as "pitch detection algorithms," or PDA [5]. Modern pitch perception models assume that pitch is made either from the periodicity of neural patterns in the time domain [6,7], or else from the harmonic pattern of partials resolved by the cochlea in the frequency domain [8,9]. Both processes result the fundamental frequency.

There are some applications which give the definition of F_0 according to their uses. For voiced speech, F_0 is defined as the rate of vibration of the vocal folds. Periodic vibration at the glottis may produce speech that is less perfectly periodic because of movements of the vocal tract that filters the glottal source waveform.

F_0 estimation is a topic that continues to attract much effort and ingenuity, despite the many methods that have been proposed. The most comprehensive review is that updated by Hess [5]. Speech F_0 variations contribute to prosody, and in tonal languages they help distinguish lexical categories. Attempts to use F_0 in speech recognition systems have met with mitigated success, in part because of the limited reliability of estimation algorithms. Several musical applications need F_0 estimation, such as automatic score transcription or real-time interactive systems, but here again the imperfect reliability of available methods is an obstacle. F_0 is a useful ingredient for a variety of signal processing methods, for example, F_0-dependent spectral envelope estimation [10]. Finally, a fairly recent application of F_0 is as metadata for multimedia content indexing. Some acoustic cues may be correlated and provide similar information for perceiving speech, it is likely that many cues provide different information, and therefore serve different roles for understanding speech. "Acoustic Theory of Speech Production", which systematized the field of acoustics phonetics [11]. Fant's book presented the theory buttressed by his own empirical work as well as by modeling efforts by scientists working in the field of phonetics [12]. His book remained major theoretical foundation for speech acoustic research in both academy and industry. Clearly defining the roles of different auditory speech properties is essential for the design of more advanced signal-processing technology so that processing can focus on preserving and enhancing the particular auditory speech cues that are most informative in a given auditory environment. There is evidence that a rise versus a fall in F_0 can serve as a perceptual cue to stop consonant voicing [13,14,15]. However, contrary to earlier work recent acoustic analyses show that F_0 tends to decrease in both voiceless and voiced

environments, though more so in the former than in the latter [15,16]. Moreover, F_0 tends to be higher following voiceless consonants than voiced consonants.

Researchers have found that the laryngeal gesture is qualitatively the same for voiceless aspirated stops and voiceless stops in fricative-stop clusters [17]. Some limited data from a single subject showed that F_0 was higher after voiceless unaspirated stops than after voiced stops [18]. Thus, these findings would suggest that during the production of vowels preceded by voiceless unaspirated stops, there will be relatively large changes in F_0 after voicing onset and high absolute values of F_0, just as in vowels preceded by aspirated stops. After examine the effect of fundamental frequency (F_0) on the voice onset time (VOT) from the voice consonant in word-initial position across high-, medium-, and low-F_0 levels [19]. Natural variations in fundamental frequency were either flattened or removed. The speech envelope and temporal fine structure were also investigated by limiting the availability of these cues via noisy signal extraction. The contribution of these acoustic properties, present during either consonants or vowels, to overall word and sentence intelligibility [20].

One property explored by the current study is the contribution of the F_0 contour to sentence and word intelligibility. The natural kinematics of F_0 have been demonstrated to facilitate speech recognition [21]. as this intonation information facilitates the prediction of syntactic units [22]. Recently proposed that F_0 may be a potential cue that vowels are more appropriate to carry than consonants, and would likely facilitate sentence recognition more so than isolated words [23]. In addition to enhancing syntactic predictability, F_0 cues are believed to aid in stream segregation [24] and facilitate the integration of speech glimpses during interrupted contexts [25]; two properties that may be important for understanding sentences under a noise replacement paradigm where alternating segments (i.e., consonants or vowels) are replaced by noise. As the primary periodic intervals of speech, vowel units include the majority of this dynamic F_0 information. This evidence suggests that F_0 contour cues may play a vital role in recognizing words from vowel segments in sentence contexts.

Method

Nepali language is a major official language in Nepal (population around 30 million) as their mother tongue. A total of twenty speakers aged between 15 to 20 were examined for this study. Ten vowels of Nepali language combine with seven consonants in initial position form VC type of isolated words. These 70 words were recorded three times for each speaker and analyzed using a PC software of Praat at a sampling rate of 44.1 kHz.

Language and Speakers

Altogether 10 boys and 10 girls were selected for the research on the criteria that they had Nepali language as their mother tongue, able to read Nepali language and normal speech, language and hearing function. The speakers having age less than 15 and more than 20 are not included.

Stimuli

The main theme of the study was to find out the variation of fundamental frequency in initial position of vowels. We have taken into consideration of 10 vowels /ʌ/, /a/, /I/, /i/, /U/, /u/, /e/, /ɛ/, /O/ and /Ɔ/ and 7 stop voice consonants of Nepali language, i.e., /p^h, t^h, $ʈ^h$, k^h, b^h, d^h, g^h/ where the vowels are in initial position. These 10 vowels combine with seven consonants in initial position form VC type of isolated words. These 10 vowels form 70 words in this case. So, we analyze 70 words and three times for each speaker.

Recordings

The pupils were instructed to read, the tokens written on the flash card as naturally as possible. The recording was done in partially acoustically treated room for individual subjects by presenting one flash card at a time using voice recording system. This was connected to the computer having software of Praat. The speech signal was digitized at the sampling frequency of 44.1 KHz.

Results and Discussion

To measure the fundamental frequency, the recorded VC (vowel consonant) format of word from different male and female speaker are analyze by Praat software. Here the fundamental frequency of different vowel with different aspirated and unaspirated voiced consonant are tabulated below for male and female speakers.

Table 1 and Table 2 show the variation of fundamental frequency of voiced consonants for male aspirated and unaspirated sound. The ranges of frequency for aspirated sound for male are from 165.60 to 232.12 Hz. The order of the average of variation of frequency for voiced consonant of male aspirated is large for Dental (Avg. 216.38 Hz, SD = 7.56) followed by velar (Avg. 212.60 Hz, SD = 17.73) and labial (Avg. 197.15 Hz, SD = 10.65) sounds. Similarly, the ranges of frequency for unaspirated sound for male is from 193.93 to 254.97 Hz. The order of the average of variation of frequency for voiced consonant is large for labial (Avg. 245.43 Hz, SD = 9.68) followed by Retroflex (Avg. 235.79 Hz, SD = 15.82) and Dental (Avg. 225.53 Hz, SD = 7.63) and Velar (Avg. 220.07 Hz = 13.74) sounds.

Table 1 Fundamental frequency (in Hz) of different vowel with voiced consonants for male speakers in aspirated sound. In the Nepali language, VC-type isolated words for retroflex consonants are not found with aspirated sound.

Vowel	Voiced (Aspirated male)				
	Labial	Dental	Retroflex	Velar	Average value
/ʌ/	203.32	203.34	--	214.69	207.11
/a/	**165.60**	216.37	--	203.28	**195.09**
/I/	197.28	220.04	--	211.93	209.75
/i/	**232.12**	212.35	--	211.62	218.70
/U/	218.08	227.40	--	230.74	**225.41**
/u/	222.20	215.72	--	205.93	214.62
/e/	198.11	228.24	--	226.09	217.48
/ɛ/	202.62	206.92	--	191.26	200.27
/O/	210.00	219.88	--	212.42	214.10
/Ɔ/	222.12	213.56	--	218.07	223.92
Avg.	**197.15**	**216.38**	**--**	**212.60**	**212.64**
SD	**17.73**	**7.56**	**--**	**10.65**	**8.68**

Table 2 Fundamental frequency (in Hz) of different vowel with voiced consonants for male speakers in unaspirated sound

Vowel	Voiced unaspirated male				
	Labial	Dental	Retroflex	Velar	Average value
/ʌ/	238.89	225.98	228.64	203.59	224.28
/a/	221.67	210.18	204.38	**193.93**	**207.54**
/I/	248.23	228.23	250.34	213.18	226.00
/i/	227.85	232.98	240.42	213.98	232.81
/U/	235.65	225.01	**254.97**	225.01	**235.16**
/u/	241.19	221.72	212.44	239.19	228.63
/e/	226.91	238.82	247.29	219.39	233.10
/ɛ/	222.46	229.87	233.44	223.54	221.33
/O/	241.95	225.57	236.78	238.51	227.70
/Ɔ/	249.47	216.94	249.18	230.33	231.48
Avg.	**245.43**	**225.53**	**235.79**	**220.07**	**226.80**
SD	**9.68**	**7.63**	**15.82**	**13.74**	**7.61**

The variation of fundamental frequency of voiced consonants for female aspirated and unaspirated sound can be observed from Table 3 and Table 4. It has been found that the ranges of frequency for aspirated sound for female are from 232.60 to 282.17 Hz. The order of the average of variation of frequency for voiced consonant of male aspirated is large for Labial (Avg. 261.88 Hz, SD = 13.63) followed by Dental (Avg. 247.57 Hz, SD = 6.70) and Velar (Avg. 245.73 Hz, SD = 5.64) sounds. Similarly, the ranges of frequency for unaspirated sound for female is from 204.08 to 301.27 Hz. The order of the average of variation of frequency for voiced consonant is large for Velar (Avg. 274.90 Hz, SD = 11.95) followed by Dental (Avg. 274.28 Hz, SD =11.79) and Retroflex (Avg. 260.95 Hz, SD = 8.77) and Labial (Avg. 260.73 Hz, SD = 21.59) sounds.

Table 3 Fundamental frequency (in Hz) of different vowel with voiced consonants for female speakers in aspirated sound. In the Nepali language, VC-type isolated words for retroflex consonants are not found with aspirated sound.

Vowel	Voiced aspirated female				
	Labial	Dental	Retroflex	Velar	Average value
/ʌ/	252.48	246.46	--	245.16	248.37
/a/	238.02	240.89	--	237.26	**238.72**
/I/	263.77	246.58	--	243.66	251.34
/i/	272.84	244.66	--	252.68	256.73
/U/	281.27	257.02	--	243.18	**260.49**
/u/	**282.17**	253.60	--	253.19	252.99
/e/	268.16	250.38	--	247.66	255.38
/ɛ/	248.76	251.14	--	240.94	246.65
/O/	255.59	**232.60**	--	239.60	252.60
/Ɔ/	255.70	252.37	--	253.92	254.00
Avg.	**261.88**	**247.57**	--	**245.73**	**251.73**
SD	13.63	6.70	--	5.64	7.32

Table 4 Fundamental frequency (in Hz) of different vowel with voiced consonants for female speakers in unaspirated sound

Vowel	Voiced				
	Labial	Dental	Retroflex	Velar	Average value
/ʌ/	280.35	263.37	256.53	258.11	264.59
/a/	**204.08**	254.14	264.08	263.40	**246.43**
/I/	279.46	277.53	251.55	272.72	268.32
/i/	261.17	278.40	266.89	284.73	272.80
/U/	279.16	275.96	266.95	280.14	275.55
/u/	249.38	299.06	267.44	**301.27**	269.28
/e/	265.69	285.19	270.86	278.97	**275.18**
/ɛ/	254.87	272.90	260.88	272.94	265.40
/O/	274.38	264.56	264.13	267..51	267.65
/Ɔ/	258.76	271.71	240.20	269.22	269.97
Avg.	**260.73**	**274.28**	**260.95**	**274.90**	**267.72**
SD	**21.59**	**11.79**	**8.77**	**11.95**	**7.87**

From Table 5, the average value of fundamental frequency of vowel with voice aspirated consonants varies from 195.09 to 225.41 Hz having average value 212.64 Hz (SD = 8.68) but in case of voice unaspirated consonants that varies from 207.54 to 235.16 Hz having average value 226.80 Hz (SD = 7.61) for male speakers. The average ratio of aspirated to unaspirated is measured as 0.94 which is less than one that means the fundamental frequency of unaspirated voiced consonant is greater than aspirated voiced consonant. The value of correlation is 0.93 that means the fundamental frequency of vowel with aspirated and unaspirated voiced consonants are highly correlated.

From Table 6, average value of fundamental frequency of vowel with voiced aspirated consonants varies from 238.72 to 260.49 Hz having average value 251.73 Hz (SD = 7.32) but in case of voiced unaspirated consonants that varies from 246.43 to 275.18 Hz having average value 267.72 Hz (SD = 7.87) for female speakers. The average ratio of aspirated to unaspirated is measured as 0.94 which is less than one that means the fundamental frequency of unaspirated voiced consonant is greater than aspirated voiced consonant. The value of correlation is 0.95 that means the fundamental frequency of vowel with aspirated and unaspirated voiced consonants are highly correlated.

Table 5 Average fundamental frequency (in Hz) and statistical value of vowel with aspirated and unaspirated voiced consonant for male speakers

Vowels	Aspirated	Unaspirated	Ratio	Correlation
/ʌ/	207.11	224.28	0.92	0.93
/a/	195.09	207.54	0.94	
/I/	209.75	226.00	0.93	
/i/	218.70	232.81	0.94	
/U/	225.41	235.16	0.96	
/u/	214.62	228.63	0.94	
/e/	217.48	233.10	0.93	
/ɛ/	200.27	221.33	0.91	
/O/	214.10	227.70	0.94	
/ɔ/	223.92	231.48	0.97	
Avg.	212.65	226.80	0.94	
SD	8.68	7.61		

Table 6 Average fundamental frequency (in Hz) and statistical value of vowel with aspirated and unaspirated voiced consonant for female speakers

Vowels	Aspirated	Unaspirated	Ratio	Correlation
/ʌ/	248.37	264.59	0.94	0.95
/a/	238.72	246.43	0.97	
/I/	251.34	268.32	0.94	
/i/	256.73	272.80	0.94	
/U/	260.49	275.55	0.95	
/u/	252.99	269.28	0.94	
/e/	255.38	275.18	0.93	
/ɛ/	246.65	265.40	0.93	
/O/	252.60	267.65	0.94	
/ɔ/	254.00	269.97	0.94	
Avg.	251.73	267.72	0.94	
SD	7.32	7.87		

The fundamental frequency (F_0) of a vowel provides the information about the rate of vibration of the vocal folds. When the fundamental frequency is high, then rate of vibration of vocal fold is higher, and vice versa. Differences in fundamental frequency appears due to variations in glottal airflow during the production of vowel sounds. In case of aspirated consonants, their

articulation affects vocal cord vibration during the production of vowel sounds. As a result, average fundamental frequency of vowels for voiced aspirated consonants is lower than that of voiced unaspirated consonants for both genders.

One of the earliest experimental investigations into variations in speaking fundamental frequency (F_0) was conducted through a series of studies initiated at the University of California, Los Angeles (UCLA) during the 1960s [26-28]. Numerous studies in recent decades have indicated that speakers of various languages or dialects [29], as well as different social groups within the same language [30], exhibit distinct ranges and typical values [31] of speaking fundamental frequency (F_0).

By setting the minimum pitch 75 Hz to maximum pitch 800 Hz, analyses the fundamental frequency in Praat [32] by the method of autocorrelation. The fundamental frequency is one acoustical cue and it has great role in identification of isolated words and sentence [23,33]. This frequency also varies with vocal effort [34], loudness of vowels [35] of different speakers. When compare the fundamental frequency between two villages of Standard Mandarin speakers [36] then finding some significant difference. In case of Chinese-English speakers there is some difference in fundamental frequency between languages [37]. The fundamental frequency of English male native speaker has lowest fundamental frequency [38] in comparison to other speakers Spanish, Japanese, [39] and Tagalog but lower than Mandarin [40]. Although there are different research work on the fundamental frequency of speaker over different language [41,42], community [30], and gender [43], this research work absolute over the Nepali speaker. Here compares the fundamental frequency of Nepali vowels with aspirated and unaspirated consonants, and gender.

The results of fundamental frequency variations in male voiced consonants shows significant differences between aspirated and unaspirated sounds. After study over the above result, it observed that, aspirated consonants have a wider frequency range. Among them, dental sounds have the highest average frequency (216.38 Hz). On the other hand, unaspirated consonants have a narrower frequency range, with labial sounds having the highest average frequency (254.43 Hz). For female voiced consonants, the pattern observed in male speaker is reverse in which unaspirated consonants show a wider frequency range, with velar sounds leading in average frequency (274.90 Hz). In contrast, aspirated consonants have a narrower frequency range, with labial sounds having the highest average frequency (261.88 Hz).

The analysis of the average fundamental frequency of vowels with aspirated and unaspirated consonants for both male and female speakers are also more informative. For male speakers, the average frequency for vowels with aspirated consonants is 212.64 Hz, while unaspirated consonants have a higher average frequency of 226.80 Hz, resulting in an average ratio of 0.94. This indicates that unaspirated consonants produce a greater vocal frequency.

Similarly, for female speakers, the average frequency for vowels with aspirated consonants is 251.73 Hz, compared to 267.72 Hz for unaspirated consonants, yielding an average ratio of 0.94. The high correlation values of 0.93 for males and 0.95 for females demonstrate a strong

relationship between consonant type and their fundamental frequencies. In both cases, unaspirated consonants produce a greater vocal frequency than aspirated. These findings suggest that unaspirated consonants may enhance pitch perception in both male and female speech. Overall, this research enhances our understanding of how consonant aspiration affects vocal characteristics, underscoring the importance of considering these differences in phonetic and linguistic studies.

Figure 1 is a box plot of the comparing the average fundamental frequency of vowel with aspirated and unaspirated voiced consonant between male and female speakers. From the box plot (Figure 1) we study the distribution of data and which provides a visual information of central tendency, dispersion, outliers and skewness of the data.

The central line of the box plot lies outside the other boxes, indicating that the fundamental frequency of vowels with aspirated and unaspirated consonants for male and female speakers. The central line of unaspirated consonant is higher than aspirated consonants which signifies that the fundamental frequency of vowel with unaspirated consonants is greater than aspirated consonant for both genders. When analyzed across the gender then we found the fundamental frequency of female is higher than male speakers for both aspirated and unaspirated consonants.

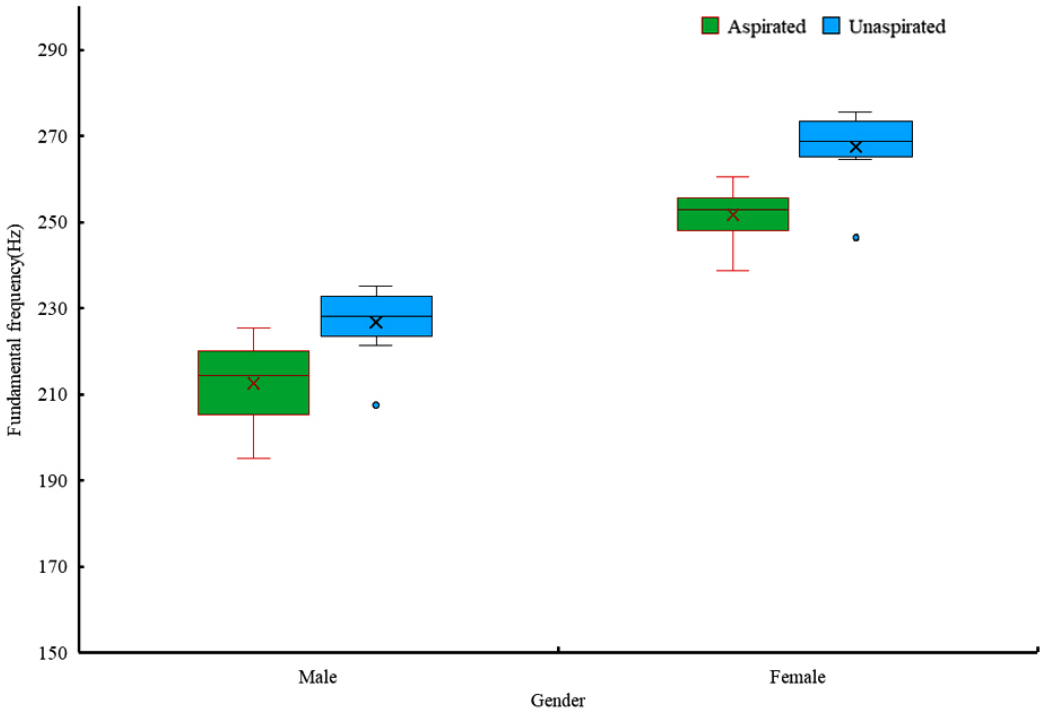


Figure 1 *The box plots of comparing the average fundamental frequency (in Hz) of vowel with aspirated and unaspirated voiced consonant between male and female speakers*

The inter-quartile range (IQR) of aspirated and unaspirated consonant for male speaker is higher than female speaker. This result indicate that the fundamental frequency of male speakers is more variable than female speakers. There is a significant difference in the IQR of aspirated and unaspirated consonant for male speaker but in case of female less significant difference are observed. The fundamental frequency of male frequencies is more variable than female speakers. Outliers appear in case of unaspirated consonant for both genders. That means fundamental frequencies of vowels with unaspirated consonant are inconsistent for both genders. The data of fundamental frequency of vowels with aspirated consonant is left-skewed but right skewed is observed in case of unaspirated for male and female speakers.

Conclusion

The fundamental frequency for aspirated and unaspirated sounds have been calculated and it is found that variation of fundamental frequency in male speaker in aspirated of voiced (Avg. 212.65 Hz, SD = 8.68) is lower than in un-aspirated voiced (Avg. 226.80 Hz, SD = 7.61) with ratio (Avg. 0.94). Similarly, the average variation of fundamental frequency of female speaker in aspirated of voiced (Avg. 251.73 Hz, SD = 7.32) is lower than in unaspirated of voiced (Avg. 267.72 Hz, SD = 7.87) with ratio (Avg. 0.94). From these data it also found that the average fundamental frequency of male speaker is lower than female speakers in aspirated and unaspirated of voiced consonants.

From above observation Table and Graph, it found that the fundamental frequency of unaspirated voiced consonant (Avg. 226.80 Hz, SD = 7.61 for male, Avg. 267.72 Hz, SD = 7.87 for female) is higher than aspirated voiced consonant (Avg. 212.65 Hz, SD = 8.68 for male and Avg. 251.73 Hz, SD = 7.32). When study across the gender, the average fundamental frequency of male speaker is lower than female speakers in aspirated and un-aspirated of voiced consonants.

Measuring the fundamental frequency of voiced sounds along with voiced consonants is very important for identifying vowels in various languages. It will serve as the best tool for speech identification techniques in artificial intelligence for speech analysis and synthesis. Various communication machines are also going on the research process which could give more advantages for better life.

Acknowledgement

The authors gratefully acknowledge academic support from the Department of Physics, Birendra Multiple Campus, Bharatpur and Central Department of Physical and Mathematical Sciences, Mid-West University Nepal. This research received no specific financial support or funding from any public, private, or commercial funding agency.

References

[1] R. J. Ritsma, "Existence region of the tonal residue. I", Journal of the Acoustical Society of America, 34, 1224–1229 (1962). https://doi.org/10.1121/1.1918307.

[2] D. Pressnitzer, R. D. Patterson and K. Krumbholz, "The lower limit of melodic pitch", Journal of the Acoustical Society of America, 109, 2074–2084 (2001). https://doi.org/10.1121/1.1359797.

[3] K. N. Reddy, "The duration of Telugu speech sounds: an acoustic study", IETE Journal of Research, 34, 57–63 (1988).

[4] W. A. Yost, "Pitch strength of iterated rippled noise", Journal of the Acoustical Society of America, 100, 3329–3335 (1996). https://doi.org/10.1121/1.416973.

[5] W. Hess, "Pitch determination of speech signals: Algorithms and devices", Springer Science & Business Media (2012).

[6] B. C. J. Moore, "An introduction to the psychology of hearing", (4th Edition), Academic press (1997).

[7] P. A. Cariani and B. Delgutte, "Neural correlates of the pitch of complex tones. I. Pitch and pitch salience," Journal of Neurophysiology, 76, 1698–1716 (1996).

[8] P. L. Wightman, "The pattern-transformation model of pitch", Journal of the Acoustical Society of America, 54, 407–416 (1973). https://doi.org/10.1121/1.1913592.

[9] E. Terhardt, "Pitch, consonance, and harmony", Journal of the Acoustical Society of America, 55, 1061–1069 (1974). https://doi.org/10.1121/1.1914648.

[10] A. De Cheveigné and H. Kawahara, "YIN, a fundamental frequency estimator for speech and music", Journal of the Acoustical Society of America, 111, 1917–1930 (2002). https://doi.org/10.1121/1.1458024.

[11] R. Jakobson, G. Fant and M. Halle, "Preliminaries to speech analysis: The distinctive features and their correlates", Cambridge, Massachusetts: MIT Press (1963).

[12] G. Fant, "Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations", (2nd Edition). Berlin, Boston: De Gruyter Mouton (1971).

[13] L. A. Chistovich, "Variations of the fundamental voice pitch as a discriminatory cue for consonants", Soviet Physics: Acoustics, 14, 372–378 (1969).

[14] M. Haggard, S. Ambler and M. Callow, "Pitch as a voicing cue", Journal of the Acoustical Society of America, 47, 613–617 (1970). https://doi.org/10.1121/1.1911936.

[15] M. Haggard, Q. Summerfield and M. Roberts, "Psychoacoustical and cultural determinants of phoneme boundaries: evidence from trading F_0 cues in the voiced–voiceless distinction", Journal of Phonetics, 9, 49–62 (1981).

[16] W. A. Lea, "Segmental and suprasegmental influences on fundamental frequency contours", (pp. 15–70), In: L. M. Hyman (Ed.), Consonant Types and Tone, Los Angeles: University of Southern California Press (1973).

[17] H. Yoshioka, A. Löfqvist and H. Hirose, "Laryngeal adjustments in the production of consonant clusters and geminates in American English", Journal of the Acoustical Society of America, 70, 1615–1623 (1981). https://doi.org/10.1121/1.387227.

[18] M. Caisse, “Cross-linguistic differences in fundamental frequency perturbation induced by voiceless unaspirated stops”, Journal of the Acoustical Society of America, 70, S76–S77 (1981). https://doi.org/10.1121/1.2019036.
[19] C. R. McCrea and R. J. Morris, “The effects of fundamental frequency level on voice onset time in normal adult male speakers”, Journal of Speech, Language, and Hearing Research, 48, 1013-1024 (2005). https://doi.org/10.1044/1092-4388(2005/069).
[20] D. Fogerty and L. E. Humes, “The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the intelligibility of words and sentences”, Journal of the Acoustical Society of America, 131, 1490–1501 (2012). https://doi.org/10.1121/1.3676696.
[21] J. S. Laures and G. Weismer, “The effects of a flattened fundamental frequency on intelligibility at the sentence level”, Journal of Speech, Language, and Hearing Research, 42, 1148–1156 (1999). https://doi.org/10.1044/jslhr.4205.1148.
[22] A. Wingfield, C. J. Lahar and E. A. L. Stine, “Age and decision strategies in running memory for speech: Effects of prosody and linguistic structure”, Journal of Gerontology, 44, P106–P113 (1989).
[23] D. Fogerty and L. E. Humes, “The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the intelligibility of words and sentences”, Journal of the Acoustical Society of America, 131, 1490–1501 (2012). https://doi.org/10.1121/1.3676696.
[24] A. S. Bregman, C. Liao and R. Levitan, “Auditory grouping based on fundamental frequency and formant peak frequency”, Canadian Journal of Psychology / Revue canadienne de psychologie, 44, 400-413 (1990). https://doi.org/10.1037/h0084255.
[25] C. J. Darwin, D. S. Brungart and B. D. Simpson, “Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers”, Journal of the Acoustical Society of America, 114, 2913–2922 (2003). https://doi.org/10.1121/1.1616924.
[26] D. Graddol and J. Swann, “Speaking fundamental frequency: Some physical and social correlates”, Language and Speech, 26, 351–366 (1983). https://doi.org/10.1177/002383098302600403.
[27] T. D. Hanley and J. C. Snidecor, “Some acoustic similarities among languages”, Phonetica, 17, 141–148 (1967). https://doi.org/10.1159/000258584.
[28] T. D. Hanley, J. C. Snidecor and R. L. Ringel, “Some acoustic differences among languages”, Phonetica, 14, 97–107 (1966). https://doi.org/10.1159/000258520.
[29] M. Dolson, “The pitch of speech as a function of linguistic community”, Music Percept: An Interdisciplinary Journal, 11, 321–331 (1994). https://doi.org/10.2307/40285626.
[30] C. G. Henton, “Fact and fiction in the description of female and male pitch”, Language & Communication, 9, 299–311 (1989). https://doi.org/10.1016/0271-5309(89)90026-8.
[31] R. J. Podesva, “Phonation type as a stylistic variable: The use of fal setto in constructing a persona”, Journal of Sociolinguistics, 11, 478–504 (2007). https://doi.org/10.1111/j.1467-9841.2007.00334.x.
[32] P. Boersma, “Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound”, Institute of Phonetic Sciences, University of Amsterdam, Proceedings, 17, 97-110 (1993).

[33] D. Fogerty and L. E. Humes, "The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the intelligibility of words and sentences", Journal of the Acoustical Society of America, 131, 1490 –1501 (2012). https://doi.org/10.1121/1.3676696.
[34] H. Traunmüller and A. Eriksson, "Acoustic effects of variation in vocal effort by men, women, and children", Journal of the Acoustical Society of America, 107, 3438-3451 (2000). https://doi.org/10.1121/1.429414.
[35] A. Geumann, "Invariance and variability in articulation and acoustics of natural perturbed speech", Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM), 38, 265-393 (2001).
[36] D. Deutsch, J. Le, J. Sheng and T. Henthorn, "The pitch levels of female speech in two Chinese villages", Journal of the Acoustical Society of America, 125, 208–213 (2009). https://doi.org/10.1121/1.3113892.
[37] S. A. Xue, F. Hagstrom and J.-P. Hao, "Speaking fundamental frequency characteristics of young and elderly bilingual Chinese'English speakers: A functional system approach", Asian Pacific Journal of Speech, Language and Hearing, 7, 55–62 (2002). https://doi.org/10.1179/136132802805576544.
[38] R. J. Baken and R. F. Orlikoff, "Clinical measurement of speech and voice", (pp. 145–223), (2nd Edition), San Diego: Singular Publishing (2000).
[39] G. Fairbanks, "Voice and articulation: Drillbook", (pp. 168), New York: Harper & Brothers (1940).
[40] P. Keating and G. Kuo, "Comparison of speaking fundamental frequency in English and Mandarin", Journal of the Acoustical Society of America, 132, 1050-1060 (2012). https://doi.org/10.1121/1.4730893.
[41] Y. Todaka, "A cross-language study of voice quality: Bilingual Japanese and American speakers", (pp. 145–147), (Doctoral dissertation), University of California, Los Angeles (1993).
[42] I. Mennen, F. Schaeffler and G. Docherty, "Cross-language differences in fundamental frequency range: A comparison of English and German", Journal of the Acoustical Society of America, 131, 2249–2260 (2012). https://doi.org/10.1121/1.3681950.
[43] H. Yamazawa and H. Hollien, "Speaking fundamental frequency patterns of Japanese women", Phonetica, 49, 128–140 (1992). https://doi.org/10.1159/000261907.