Journal of Acoustical Society of India : Vol. 37, No. 3, 2010 (pp. 121-129)

A Preliminary Acoustic Study of Mizo Vowels and Tones

Priyankoo Sarmah[1] and Caroline R. Wiltshire[2]
[1]*Department of Linguistics and Cognitive Science Hankuk University of Foreign Studies*
202-1, 89 Wangsan-ri, Choein-gu, Mohyeon-myeon
Yongin, South Korea 449-791
e-mail: priyankoo@gmail.com
[2]*Department of Linguistics University of Florida 115454, Turlington Hall*
Gainesville, Florida, USA 32611
e-mail: wiltshir@ufl.edu

[Received: 12.04.2010; Revised: 27.05.2010; Accepted: 11.06.2010]

ABSTRACT

This paper provides quantitative measures of Mizo vowels and tones. The first two formants and durations of the five Mizo vowels were extracted, and their perceptual distances were measured. As F0 is the primary cue for the identification of Mizo tones, a quantitative parameter is built for instrumentally identifying the four Mizo tones namely, high, low, rising and falling tones using F0 as the major cue. While doing that, we also examined (a) if consonantal segments affect the pitch of Mizo tones in any way. In the presence of positive evidence for consonantal influence, this study will try to determine (b) how far into the F0 can the consonantal effects perturb. The results indicated that (c) in Mizo; consonantal segments in the onset do affect the pitch of the following tone in a significant manner. With the methodology adopted in this study it is concluded that the effects of consonants into the F0 may be highly predictable.

1. INTRODUCTION

Mizo is a Tibeto-Burman language spoken by over half a million people: approximately 539,000 in India, 1,000 in Bangladesh, and 12,500 in Myanmar. In India, it is spoken primarily in the state of Mizoram (Fig. 1). There has been descriptive work on its sound system in general [see 1-4], and its vowel inventory and tone system has been described and analyzed by native speakers [5-7]. However, none of the studies conducted on Mizo has based their conclusions on acoustic analysis.

Mizo is described as having five vowels with long and short distinction for each [5-7]. According to the same studies, the ten total vowels of Mizo are [i:, i, u:, u, ɛ:, ɛ, ⊃:, ⊃, a:,a]. The low vowels [a:, a] are described as central vowels. No other information about the Mizo vowel system is available from these studies.

The Mizo tone system is described as having an inventory of four tones: High (H), Low (L), Rising (R) and Falling (F) [5-7]. These studies do provide a few rules for the interactions of tones in suffixed or compounded words; however, both these researchers stress the desirability of acoustic analyses to support their descriptions.

The tones described in previous literature are more or less consistent with slight variations. For example, Bright is of the view that Mizo has H, F, L and an allophonic mid level tone sometimes realized as a mid to low

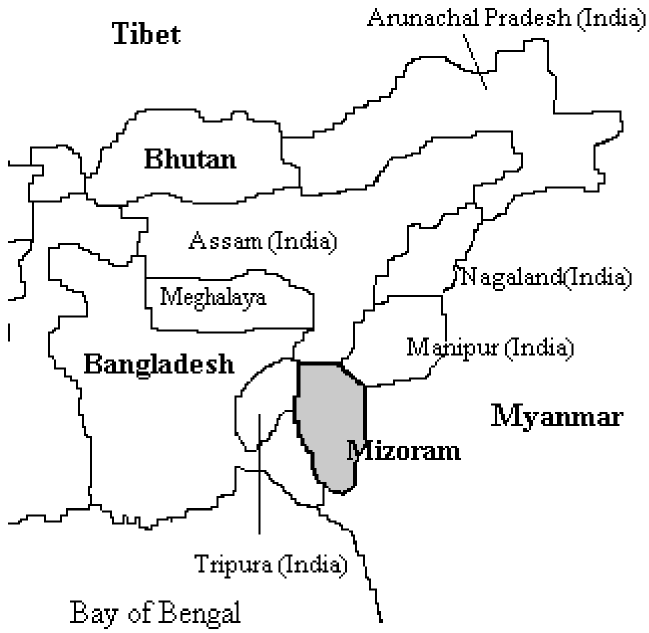

Fig. 1. Location of Mizoram in North-East India

falling tone [1]. Weidert describes the Mizo tones as high-level, high-falling, low-rising and low-level tones [4]. Chhangte describes the Mizo tone inventory as including H, R, F, and an unmarked tone, where the unmarked is phonetically mid or low [5]. Fanai also describes the four tones of Mizo as H, L, R and F where the L tone can also have an allophonic variation realized as an extra low tone that she marks as L [7]. As the L tones in Mizo are realized with a mid to low fundamental frequency (F0) movement, H is the only tone with a level F0 track [7]. The H and the L tones can occur on syllables with both long and short vowels, however; the R and the F tones can only occur with syllables with more than one sonorant in the rime (vowels and following consonants in a syllable). Hence, the four tones in Mizo can surface in monosyllabic, disyllabic and trisyllabic words as shown in Table 1.

Table 1. Tone distribution in Mizo syllables as in [7]

Syllable Types	Distributions								
Monosyllable	H	F	L	R					
Disyllable	HH	HL	HR	FL	LL	LH	LR	LF	RR
Trisyllable	LLL	LHH							

This work adds to the previous research on Mizo in various ways. First, this study provides the first acoustic analysis of the vowels in Mizo. Second, it provides the measures of perceptual difference (D_{ij}) between vowels in Mizo following Lindblom [8]. Third, this study provides an acoustic measure of the tones in words whose tones have been provided by previous researchers. Corresponding to the previous research findings it was confirmed that Mizo has four tones distinct tones namely, L, H, R and F. Fourth, the extent to which a syllable initial consonant affects the pitch of the part that hosts the tone or the tone bearing unit (TBU) in Mizo is examined, and some preliminary results about the differential effect of initial consonants are reported.

2. METHOD

2. 1 Data collection

One female native speaker of Mizo was recorded reading a list of target words in a frame sentence. The speaker was 25 years old, had been born and raised in Aizawl, Mizoram, and spoke Mizo, English, French, and a little Hindi. The speaker reported no developmental or acquired conditions that could lead to impaired speech or hearing. The frame sentence was "ka X ati", which translates to 'I said X' in English; the frame was used to keep the segmental and suprasegmental context consistent.

The word list was constructed by drawing on previous studies and included monomorphemic words representing all four aforementioned tones [4, 7]. For tone analysis 59 words appearing in minimal pairs were recorded. For vowel quality and temporal measurements 49 additional words were recorded resulting in a total of 108 words. To avoid any confusion for the speaker about which word was intended, English glosses were provided, as Mizo orthography contains no representation for tones [5]. The speaker was asked to produce the words four times each, each time embedded in the frame, and the fourth repetition was discarded to avoid list intonation effects. High quality recordings were made using a Sony TCD-D8 DAT recorder with a 44.1 KHz sampling ceiling, and a head-mounted Shure SM10A microphone, then digitized into a CSL model 4400. The data was acoustically analyzed using Praat 4.3.09 [9].

3. RESULTS

3.1 Vowels

3.1.1. Vowel quality and duration

The first and the second formant frequencies of vowels are argued to be the primary cue for identification of vowels [10]. Hence, the F1 and F2 for the Mizo vowels produced between two consonants were calculated. The formant frequencies were measured at the mid point of the vowels, following the methodology described in Abramson [11]. The F1 and F2 values for the Mizo vowels are presented in Table 2.

Table 2. Average formant frequencies (in Hz) and durations of Mizo vowels

Vowel	Long			Short		
	F1	F2	Duration (ms)	F1	F2	Duration (ms)
a	839	1418	178	781	1548	87
ε	617	2313	205	507	2218	99
⊃	779	1956	213	621	1416	96
i	NA	NA	NA	402	2490	85
u	548	1407	186	486	1208	82

As noticed from Table 2, the five vowels in Mizo are considerably distinct from each other at the mid point of the first two formants. In terms of vowel length, it is noticed that the long vowels are at least twice as long as the short vowels.

3.1.2 Perceptual distance

According to Lindblom [8] perceptual distinctiveness between vowels can be numerically expressed by calculating the Euclidean distance between two vowels. Hence, the perceptual spaces among the Mizo vowels were calculated using Eq. (1), where D_{ij} is the Euclidean distance between two vowel points and M1 and M2 are the frequency of F1 and F2 expressed in mels [8]. A higher D_{ij} value confirms that the two vowels i and j are highly distinct, for example in case of English, it has been noted that the average D_{ij} value is about 70.2 mels [12].

(1)

Table 3. Perceptual distance (D_{ij}) in mels among Mizo vowels

Vowels	i	u	ɛ	⊃
a	308	140	233	77
i		280	84	231
u			233	125
ɛ				156

Table 3 shows that Mizo vowels maintain an average perceptual difference of 187 mels. While, the vowels [i] and [a] maintain the highest difference. However, vowels pairs like [ɛ]-[i] and [a]-[⊃] have much smaller perceptual differences namely, 84 mels and 77 mels, respectively. However, comparing with the perceptual difference scores obtained for English, these pairs are highly distinct.

3.2 Mizo tones

Mizo tone bearing units (TBUs) constitute the vowel of the syllable and the following voiced consonant, if the latter is available [7]. With the help of a script designed for extracting pitch values in Praat, the F0 values of each tone was measured. A Praat script calculated the pitch at every 2% of the total duration of the TBU. Praat was also used to calculate the average pitch of the TBU based on the whole span from initiation to termination of the pitch on the TBU.

3.2.1 Pitch of the four tones in Mizo

The results of the analyses of tones in this study closely concur with the findings of previous researchers. It is confirmed by means of acoustic analyses that Mizo has four distinct tones in the language system, and show that these can be differentiated by our two measures of pitch. In Figure 3 the average pitch contours of the four tones in Mizo are presented. In Figure 3 the pitch tracks of the four tones in Mizo are normalized for duration and plotted on the basis of six equidistant points on the pitch contour. The whole pitch track from the point of initiation to the point of termination of the TBU is extracted and presented in the figure. As claimed in Fanai [7], the high tone in Mizo remains fairly level through its realization. The point of initiation of the falling tone is

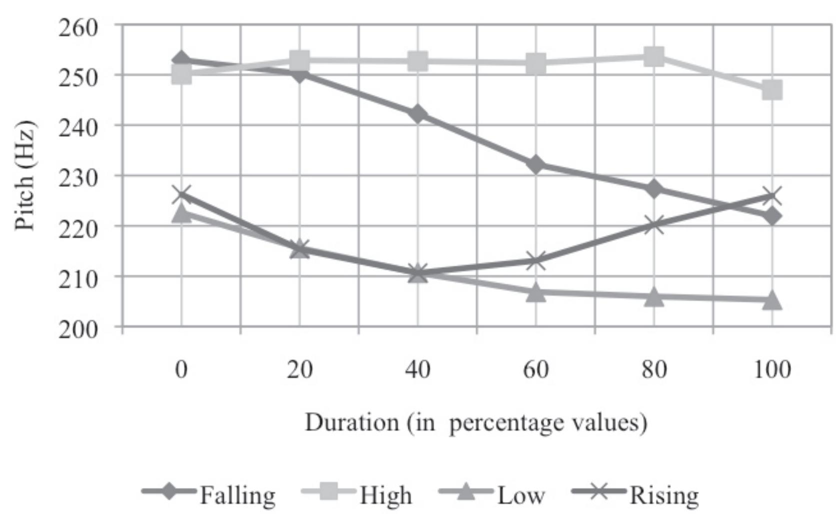

Fig. 3. Average pitch tracks of the Mizo tones

similar to the point of initiation of the high tone. The rising tone in Mizo shows an initial downward dip however from about 40% of the total duration it rises up to the point of termination of the falling tone. Similarly, the low tone in Mizo demonstrates a falling contour. Considering the phonetic property of the Mizo low tone, it is more apt to label it as a low falling tone than a low tone.

As far as the rising tone in Mizo is concerned, it is seen that the rise in F0 occurs in the latter half of the pitch track. It has also been noticed that the rising tone in Mizo occurs only on TBUs that have at least two constituents (such as a long vowel, a diphthong or a vowel followed by a voiced consonant), called rimes [7]. Hence, it also raises the question of whether the rising tone is realized actually only on the final constituent of the rime. To investigate this issue, the rime lengths of different Mizo rimes based on their moracity are examined in the following subsection.

3.2.2 Mizo rime durations

Rime durations of different Mizo syllables were calculated and statistically compared (see Figure 4 and Figure 5). The rime lengths by syllable types in Mizo are significantly different from each other [$F (6, 301) = 39.8$, $p < 0.001$]. The syllable types were further categorized by the sonority properties of the codas. Bimoraic and trimoraic rimes, where all the moras were sonorants, appeared to be longer in duration than rimes that have a non-sonorant as the final mora (Fig. 5). This difference is found to be statistically significant [$F (3, 304) = 75.38$, $p < 0.001$]. The average duration of the rime in a CVC syllable in Mizo is 88.0 milliseconds, whereas a CV-Son rime on average is 206.63 ms long. Similarly, a bimoraic, CVV rime is 199.37 ms long; however, the average duration of a CVVSon is 235.34 milliseconds.

Mizo R and F tones occur only in bimoraic or trimoraic rimes whereas the H tone and the L tone can be associated with both bimoraic and monomoraic rimes [7]. Hence, when the duration of the rimes is compared by tone types, it is possible that they will all turn out to be significantly different from each other in terms of their average durations. On the other hand, if the duration of only the bimoraic rimes are compared across all tone types, it is assumed that they will not show any significant interaction among each other. Hence, the duration of the rimes in Mizo in the bimoraic context were compared across all four tone types. Figure 6 shows the average duration of Mizo tones in bimoraic rimes. Figure 6 also shows that the duration of the F tone is shorter than the duration of the rimes bearing the rest of the four tones. A statistical analysis of variance (ANOVA) test conducted on average duration for tone types effect showed overall significance, indicating [$F (3, 172) = 17.74$, $p < 0.001$] significant interaction between tone types and average duration of rimes. However, a Bonferroni post-hoc test confirmed that among the four tones in Mizo; only the F tone differs from the other three tones and that the H, R and L tones do not show any significant difference in terms of rime duration among them.

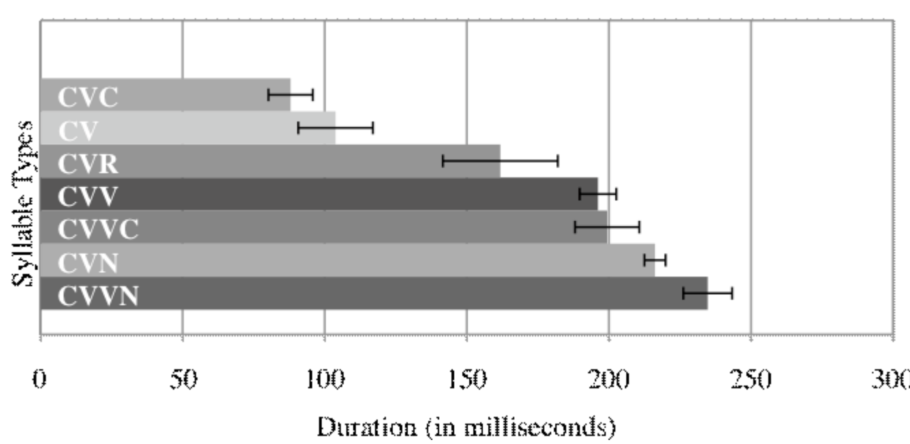

Fig. 4. Average duration of various rime types in Mizo (with standard error bars)

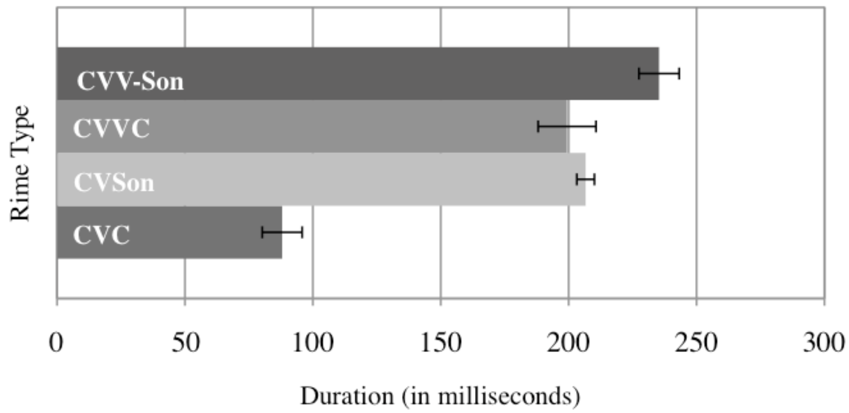

Fig. 5. Average duration of rime in sonorant and non-sonorant codas (with standard error bars)

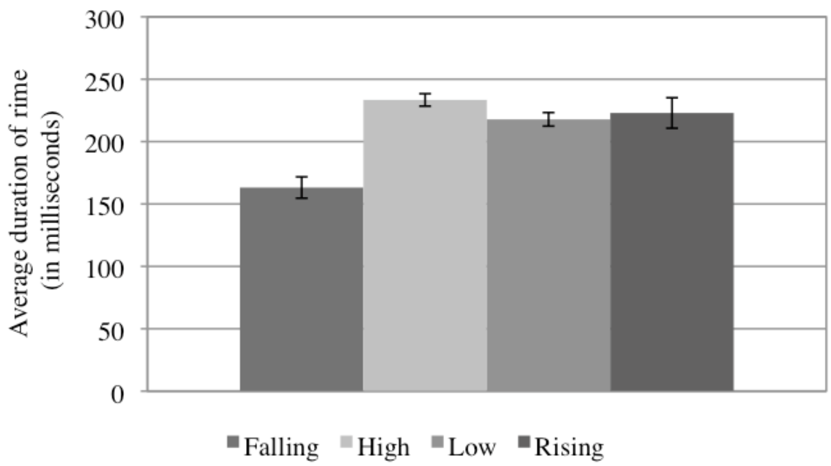

Fig. 6. Duration of Mizo tones in bimoraic rimes (with standard error bars)

3.2.3 Consonantal Interference on pitch

The effect of the consonants on the following F0 has been well attested in the literature [13, 14]. The F0 at the onset of a vowel is significantly higher after a voiceless stop than after a voiced stop [15]. Similarly, an aspirated stop induces a higher F0 in the following vowel than an unaspirated stop [13]. In case of Tibeto-Burman languages, it was demonstrated that in Dimasa consonant types do affect the following fundamental frequency in a statistically significant way [16]. More significantly it has been shown in the same study that the consonantal effects permeate up to the initial 20% of the F0 contour. Hence, for Mizo, it was decided to see if the consonant types would have any effect on the following F0 contour. The onset of the F0 contour was compared by different consonant types. The results of the comparisons are shown in Table 5.

Table 5. The F and p values of statistical comparisons of onset F0

Factor	F	p value
Voicing	(1, 321) = 40.75	<0.001
Place	(2, 252) = 4.73	<0.05
Manner	(3, 319) = 7.67	<0.05

Table 5 shows that in terms of voicing, the voiced and the voiceless consonants affect the following F0 contour in different ways. Concurring with the previous conclusions on voicing effect of F0, this study also shows that the F0 onset following a voiceless consonant is higher than after a voiced consonant [15]. While the F0 contour following a voiceless consonant has an average onset of 239.09 Hz, the onset following a voiced consonant has an average F0 of 220.2 Hz.

Similarly, a significant effect of place feature of the consonants on the onset F0 was also noticed. A Bonferroni post-hoc test further showed that among the three places of articulation that were compared, alveolar consonants differed significantly from the other two places of articulation namely velar and bilabial (Bonferroni adjusted $p = 0.0125$). While alveolar consonants exerted the maximum effect on the following F0 onset, the effect of the other two types of consonants did not differ significantly among themselves. The degree of effect of consonants on the onset F0, based on their place of articulation can be expressed as *alveolar > velar = bilabial.

Manner of articulation of consonants also significantly affected the onset F0 of the following F0 contour. A Bonferroni post-hoc test confirmed that among the four manners of articulation, the stop consonants had the highest effect on the following F0 onset. However, the effect of the fricative, liquid and the nasal consonants did not differ among the three manners of articulation. Hence, the degree of effect of the consonants based on their manner of articulation can be expressed as *stop > fricative = liquid = nasal.

From the discussion above, it becomes evident that voicing, place and manner of articulation of the consonants affect the following F0 onset in a significant way. It was also noticed that the effect on the following F0 contour permeates until the initial 20% of the pitch contour of the vowel.

3.2.4 F0d values and average F0 of Mizo tones

As Mizo tones are dynamic tones, i.e. tones that change their F0 contour by the factor of time, it is not possible to quantify the acoustic properties of the tones only with average F0 of the tone. Hence, apart from calculating the average F0 of each tone the slope of the four Mizo tones was calculated using Eq. (2). In Eq. (2), y2 is the F0 value at the 78% into the F0 contour and y1 is the F0 value at the 22% into the F0 contour, normalized for duration. The rational behind calculating the F0d between 22% and 78% comes from the findings in the previous section where, it was shown that the consonantal effects can permeate up to the initial 20% of the F0 contour. Similarly, it is also noticed that the consonant following an F0 contour can also affect 20% into the preceding F0 contour [15]. These two claims were also attested by a visual examination of the average F0 contours of the four tones in Mizo (see the H tone F0 contour of Mizo in Fig. 3). Hence, it was decided to exclude the initial and final 20% of the F0 contour in calculating the slope (F0d) of the Mizo tones.

$$F0d = y_2 - y_1 \quad (2)$$

Table 4. Average F0 and F0d values in Hz by tone types

Tone	Average F0	Mean F0d
Falling	236.80	-22.51
High	251.71	0.72
Low	210.81	-9.03
Rising	214.54	11.57

The average F0 values and the F0d values for each tone in Mizo are shown in Table 4. A univariate ANOVA test showed that the four tones are distinct from each other in terms of their F0d [$F(3, 172) = 48.24$, $p < 0.001$]. A Bonferroni Post-hoc test for multiple comparisons of the tones confirmed that each of the four tones in Mizo differ from the other three tones in a significant manner in terms of their F0d (Bonferroni adjusted $p = 0.0125$).

A univariate ANOVA test conducted on the tone types as factor and average F0 as the dependent variable showed significant effect of tone types [$F(3, 172) = 55.23$, $p < 0.001$]. However, a Bonferroni post-hoc test also showed that among the four tones in Mizo, the R and the L tones and the F and the H tones do not differ from each other in terms of the average F0 (Bonferroni adjusted $p = 0.0125$). Hence, it is apparent that the average of F0 contour alone is not sufficient for identifying tone types in Mizo. As indicated before, owing to the dynamic property of the F0 contours associated with the Mizo tones, average F0 may not be the reliable measure to distinguish one tone from the other. The dynamic tones in Mizo differ from each other only in terms of the slope of the F0 contour. Therefore, the F0d measure suggested in this section is a more reliable way of measuring and identifying Mizo tones.

4. CONCLUSION

4.1 Vowels

The spectral cues that distinguish the Mizo vowels from one another were identified. It was shown that, in terms of their F1 and F2 values, Mizo vowels are quite distinct from each other. This claim is further strengthened by the perceptual difference measures that were calculated for the Mizo vowels. The minimum perceptual difference in our data is for the vowel pair a and ⊃ (da⊃ = 77 mels). This value is higher than the perceptual difference calculated for vowel pairs in other languages where the difference is perceptually well attested. For example, the average d_{ij} for English speakers across all vowel pairs is 70.2 mels [12]. In terms of temporal qualities of Mizo vowels it was noticed that the long vowels in Mizo are at least twice as long in duration as the short vowels.

4.2 Rime Durations in Mizo

In this study Mizo rime durations were compared for different type of consonants that constitute the rime. The results, confirmed by statistical tests, indicated that the duration of rimes of CV, CVV, CVC, CVR (R = [r], [l]), CVN (N = [m], [n], [ŋ]), CVVC and CVVN are significantly different from each other. While the average rime length of a CVC syllable is about 88 ms, a CVVN syllable has an average rime duration of 235 ms. When the rime duration of CVVN syllables were compared by tone types in Mizo, it was noticed that the F tone is significantly shorter than the H, L and F tones. However, the latter three tones do not show any statistically significant difference in rime duration. Hence, it is concluded that rime duration in Mizo cannot effectively identify the tone type in the language.

4.3 Consonantal Effects

It was also found that consonant quality affects the F0 of the following tone significantly, at least up to the 20% into the F0 contour. While the voiceless consonants raise the initial F0 of the following tone, voiced consonants lower the same. Similarly, stop consonants and alveolar consonants have significant raising effect on the initial part of the following F0 contour. This observation corresponds to the previous observations on consonantal interference on fundamental frequency [13, 16].

4.4 Tones in Mizo

The results obtained in this study strongly correspond to the results obtained in the previous non-acoustic studies as far as the number and nature of tones are concerned [7]. However, it was shown in this study that it is not possible to distinguish Mizo tones from each other merely by their average F0. This conclusion was statistically verified where it was seen that the four tones are not uniformly different from each other. However, in terms of the slope of the F0 contour (F0d), the four tones in Mizo are significantly different from each other. Hence, it can be concluded that distinctive identification of the Mizo tones is possible only by means of the F0d measures proposed in this study.

5. REFERENCES

[1] E. J. A. HENDERSON, 1948. Notes on the syllable structure of Lushai, *Bulletin of the School of Oriental and African Studies*, **12 (3/4)**, 713-725.

[2] W. BRIGHT, 1957. Singing in Lushai, *Indian Linguistics,* **17**, 24-28.
[3] R. BURLING, 1957. Lushai Phonemics, *Indian Linguistics*, **17**, 148-155.
[4] A. WEIDERT, 1975. Componential Analysis of Lushai Phonology, John Benjamins Publishing Company, Amsterdam.
[5] L. CHHANGTE, 1986. A Preliminary Grammar of the Mizo Language, Master's thesis, University of Texas, Arlington.
[6] L. FANAI, 1989. Some Aspects of the Autosegmental Phonology of English and Mizo. M.Litt. dissertation, CIEFL, Hyderabad, India.
[7] L. FANAI, 1992. Some Aspects of the Lexical Phonology of Mizo and English: An Autosegmental Approach, PhD dissertation, CIEFL, Hyderabad, India.
[8] B. LINDBLOM, 1986. Phonetic universals in vowel systems, Experimental Phonology, 13-44, Academic press, San Diego, California.
[9] P. BOERSMA, 1991. Praat, a system for doing phonetics by computer, *GLOT International*, **5 (9/10)**, 341-345.
[10] P. LADEFOGED, 2003. Phonetic Data Analysis, Wiley-Blackwell, USA.
[11] A.S. ABRAMSON, 1962. The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments. Indiana U. Res. Center in Anthropology, Folklore, and Linguistics, Pub. 20, Bloomington, USA.
[12] B. LEE, S.G. GUION and T. HARADA, 2006. Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals, *Studies in Second Language Acquisition*, **28**, 487-513.
[13] Y. LAI, C. HUFF, J. SERENO and A. JONGMAN, 2009. The raising effect of aspirated consonant on F0 in Taiwanese, Proceedings of the 2nd International Conference on East Asian Linguistics.
[14] J-M HOMBERT, J. OHALA and W. EWAN, 1979. Phonetic explanations for the development of tones, *Language*, **55 (1)**, 37-58.
[15] B. MOHR, 1971. Intrinsic Variations in the Speech Signal, *Phonetica*, **23**, 65-93.
[16] P. SARMAH and C. WILTSHIRE, 2010. An acoustic study of Dimasa tones, *North East Indian Linguistics,* **(2)**, 16-44.