Preliminary Acoustic Analysis of Manipuri Vowels

Tulika Basu¹, Arup Saha¹, P Madhubala²

¹Centre for Development of Advanced Computing (CDAC), Kolkata
²Manipur University
(tulika.basu, arup.saha)@cdac.in, pmadhubala@gmail.com

Abstract
Manipuri language is one of the low resourced languages of north eastern part of the India. The need of development of speech technology in the local language is of urgent demand, so as to improve the livelihood of the people. The pre-requisite of development of such speech technology requires the basic research on the acoustic phonetic of the phoneme together with the development of speech resources. Keeping the main objective in mind the current paper aims at preliminary study on acoustic characteristics of Manipuri vowels. The current study is conducted on the speech corpus of around 500 Phonetically Balanced Words (PBW) embedded in neutral carrier sentences spoken by ten informants (5 male and 5 female) in a reading mode. Moreover vowel segments which are in CVC context of the word are selected. The dialect chosen for the purpose of experiment is Imphal dialect of Manipuri language. After analyzing the data, vowel phoneme inventory of Manipuri language has been presented in this paper.

Index Terms: vowels, formants, acoustic analysis, Manipuri language, perceptual distance.

1. Introduction
Manipuri or Meitei is one of the low resourced languages of north eastern part of India whose exact classification still remains unclear but according to some linguists the language is a member of the Kuki-Chin sub-group of Tibeto-Burman languages, one of the co-ordinate members of the vast Sino-Tibetan language family. Manipuri is the official language of Manipur state since time immemorial and still continues to be the official language of the state even after Manipur joined the Indian Union on 15th Oct, 1949. Meitei is also spoken in the Indian states of Assam and Tripura, and in Bangladesh and Myanmar. According to 2001 census the language is spoken by 1.5 million people. It is currently classified as a vulnerable language by UNESCO [1]. The language has been recognized (under the name Manipuri) by the Indian Union and has been included in the list of scheduled languages of India. Manipuri has its own script called Meitei Mayek but Bengali script is also used for many purposes. Meitei contains various dialects. Among them Imphal dialect is treated as the standard one in the state of Manipur [2]. All the affairs of the state are conducted in this language and almost all the communication and medium of instructions in schools and colleges is made in this language.

Manipur is described as having an inventory of six vowels. The six vowels of Manipuri are /a/, /o/, /i/, /u/, /e/ and /l/. There is no evidence of length variation has been attested in the existing literature of Manipuri [3]. Moreover vowel nasalization is not phonemic in this language [3]. There have been a number of studies on the phonemic inventory of Manipuri language but most of the studies are based on perception. This study adds to the previous research in a way that it provides the first acoustic analysis of the vowels in Manipuri.

Rest of the paper is organized in a following way: After introduction methodology of investigation is described in section 2. Experimental findings are presented in section 3. Normalization of Meitei vowels are summarized in section 4. Finally references are given in section 5.

2. Methodology of Investigation
2.1. Recording material
For the current study the database consists of around five hundred isolated words. The wordset is prepared carefully taking into consideration of the fact that the wordset contains all the possible permitted combinations of the phonemes of the language. Moreover all the selected words for this study bear only one tone i.e. low tone to avoid the qualitative and quantitative aspects of the vowels which may be affected by tones. The distribution of the six Manipuri vowels in the current speech corpus taken for this study is given in Figure 1.

It is evident from figure 1 that the highest percentage of occurrence of vowel in Manipuri speech corpus is phoneme /a/, followed by phoneme /e/, /l/, /i/, /u/, and /o/.

![Figure 1: Distribution of Manipuri Vowels](image)

The speech corpus is recorded by ten native Manipuri informants (5 male and 5 female) and their age ranges from 20 to 35 years.

2.2. Recording
The recording of speech data for all ten speakers is done in a studio environment and digitized at a sampling rate of 48,000 Hz with an accuracy of 16 bits/sample. The data is recorded using the Shure SM58 vocal microphone with the
help of Cool Edit Pro Software. During the recording a constant distance from the microphone element and the
speaker’s mouth is maintained. After each recording, the moderator checks for any wrong pronunciation due to slip
of tongue or any hesitation during the recording, and if the moderator finds so, the utterances are recorded again. For
this study the vowel segments which are in CVC context of the word are selected. During the recording a movie of the
informant’s lips is also captured.

2.3. Analysis Procedure

A total of 19985 vowel segments (5 male and 5 female Manipuri informants) are considered for the analysis of
vowels.

From the literature survey it is evident that the articulatory position of the vowel can be determined from
the acoustic formant measurement [4]. In general high first formants are associated with narrow tongue constriction
near the glottis. Similarly the second formant frequency is increased as the constriction moves forward [5]. So for the
determination of the articulatory position of Manipuri vowels, the two formant frequencies F1 and F2 are measured.

It is worth mentioning here that for the analysis of vowel duration, steady part along with two transitions is
considered whereas measurement of formants frequency is done only from the steady part of the vowel segments itself.
Initially the vowel phonemes within each word are labeled along with time stamp by the native Manipuri transcribers.
In the next step the labeled vowel segments are cut automatically using the labeled files. The extraction of the
steady state of the vowel is done by using simple amplitude tracking algorithm. The minimum length of the steady state
vowel segment considered in the study for the extraction of formants is 40 ms. A Praat script [6] is used to
automatically extract first and second formant corresponding to vowel segments. The parameter settings
for the extraction of the formants using Praat are as follows:

- Analysis window length 0.025s
- Pre-emphasis factor 0.85
- Frame Interval 0.001s

The formant values of those vowel segments which are outside mean ± standard deviation are once again checked
and if found erroneous these are corrected manually using spectrum section. The extracted formant frequency is then
converted into Mel scale (as in equation 1) for further analysis.

\[ m = 1127 \log_10(1 + f / 700) \]  

(1)

where m is frequency in Mel scale and f is the frequency in Hertz.

To investigate the rounded/unrounded feature of the vowel, photographic evidence of the lip rounding during the
articulation is used. Audio-visual data has been captured in course of the articulation of the vowel. A frame
at the steady state of utterance of the vowel is selected for the analysis purpose. The degree of lip rounding is defined
as a ratio of height and width of the area between the opening of the lips as shown in figure 2.

3. Experimental Findings

3.1. Analysis on vowel duration

Using Praat script, vowel duration has been derived from labeled files. For the quantitative analysis of Manipuri
vowels frequency distribution of duration has been done. After getting the duration from the labeled file the
frequency distribution of duration of each of the 6 Manipuri vowels has been derived. Finally the distribution of
duration is being normalized by dividing the maximum number of occurrence in a duration interval of that
individual vowel. It is evident from figure 3 that in Manipuri language there is no long short variation of
vowels in Manipuri language as the frequency distribution of duration of each of the nucleus vowel exhibits single
peak.

![Figure 3: Frequency distribution of vowel duration](image)

Figure 4 depicts the vowel duration of the six Manipuri vowels. It is evident from figure 4 that the duration of
vowel /a/ is maximum and that of the vowel /u/ is lowest. In fact all the low vowels exhibit higher duration. This may
be related to the jaw movement which has low position for these vowels.

![Figure 4: Intrinsic Vowel Duration](image)

The low standard deviation indicates the average intrinsic vowel duration is meaningful.

3.2. Analysis on vowel formant

To obtain the acoustic features of vowels, formants were measured. Table 1 gives average (\( \mu \)) and standard deviation
(\( \sigma \)) of the values of first (F1) and second (F2) formant of all the Manipuri vowels with respect to male, female. It is
observed from Table 1 that there is a difference in mean of the formant frequency for both male and female.

Table 1: Formant Frequency (in Mel) of male, female.

<table>
<thead>
<tr>
<th>Manipuri Vowel</th>
<th>F2 (Male)</th>
<th>F1 (Male)</th>
<th>F2 (Female)</th>
<th>F1 (Female)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/u/</td>
<td>950</td>
<td>438</td>
<td>971</td>
<td>453</td>
</tr>
<tr>
<td></td>
<td>±164</td>
<td>±58</td>
<td>±199</td>
<td>±62</td>
</tr>
<tr>
<td>/o/</td>
<td>934</td>
<td>590</td>
<td>965</td>
<td>601</td>
</tr>
<tr>
<td></td>
<td>±100</td>
<td>±79</td>
<td>±91</td>
<td>±90</td>
</tr>
<tr>
<td>/ə/</td>
<td>1267</td>
<td>782</td>
<td>1308</td>
<td>858</td>
</tr>
<tr>
<td></td>
<td>±110</td>
<td>±103</td>
<td>±96</td>
<td>±85</td>
</tr>
<tr>
<td>/a/</td>
<td>1250</td>
<td>914</td>
<td>1299</td>
<td>990</td>
</tr>
<tr>
<td></td>
<td>±74</td>
<td>±85</td>
<td>±68</td>
<td>±59</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>1615</td>
<td>582</td>
<td>1668</td>
<td>606</td>
</tr>
<tr>
<td></td>
<td>±79</td>
<td>±107</td>
<td>±83</td>
<td>±35</td>
</tr>
<tr>
<td>/e/</td>
<td>1691</td>
<td>412</td>
<td>1762</td>
<td>415</td>
</tr>
<tr>
<td></td>
<td>±91</td>
<td>±51</td>
<td>±73</td>
<td>±57</td>
</tr>
</tbody>
</table>

It is observed from Table 1 that there is a difference in mean of the formant frequency for both male and female. Student t-test was performed to find whether the difference between the mean of the formant frequencies between male and female is significantly different or not. Result of the t-test is tabulated in Table 2.

Table 2: Significance of difference of formants with respect to sex of the informants

<table>
<thead>
<tr>
<th>Manipuri Vowel</th>
<th>F2</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>/u/</td>
<td>Significant at 0.05 level</td>
<td>Significant at 0.05 level</td>
</tr>
<tr>
<td>/o/</td>
<td>Significant at 0.05 level</td>
<td>Significant at 0.05 level</td>
</tr>
<tr>
<td>/ə/</td>
<td>Significant at 0.05 level</td>
<td>Significant at 0.05 level</td>
</tr>
<tr>
<td>/a/</td>
<td>Significant at 0.05 level</td>
<td>Significant at 0.05 level</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>Significant at 0.05 level</td>
<td>Significant at 0.05 level</td>
</tr>
<tr>
<td>/e/</td>
<td>Significant at 0.05 level</td>
<td>Not significant at any level</td>
</tr>
</tbody>
</table>

It is evident from Table 2 that there exists a significant difference in the formant value with respect to sex of the informants. It is interesting to note that there is no significant difference between the mean of the male and female informant’s formant frequency F1 at any level for the vowel /i/.

4. Normalization of Manipuri Vowels

In order to objectively verify that whether the overlapped vowels are perceptually distinct from one another, a perceptual distance between the overlapped vowels has been measured using Eq. (2).

$$D_{ij} = \sqrt{(M_{1,i} - M_{1,j})^2 + (M_{2,i} - M_{2,j})^2}$$

where $D_{ij}$ is the Euclidean distance between two vowel (i, j) points and $M_1$ and $M_2$ are the frequency of F1 and F2 expressed in Mel [5]. 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 Mel [7].

Table 3 presents the perceptual distance between overlapping vowel pairs.

Table 3: Perceptual Distance between overlapping vowel pairs

<table>
<thead>
<tr>
<th>Vowel Pair</th>
<th>u</th>
<th>o</th>
<th>ə</th>
<th>a</th>
<th>e</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>/u/</td>
<td>0</td>
<td>151</td>
<td>482</td>
<td>580</td>
<td>688</td>
<td>753</td>
</tr>
<tr>
<td>/o/</td>
<td>394</td>
<td>468</td>
<td>685</td>
<td>786</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ə/</td>
<td>0</td>
<td>133</td>
<td>412</td>
<td>581</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td>0</td>
<td>503</td>
<td>685</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ɛ/</td>
<td>0</td>
<td>192</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/i/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From Table 3 it can be observed that vowel /u/ and /o/ maintains the highest perceptual distance. Though perceptual distance between vowels /u/ and /o/, /ə/ and /a/, /ɛ/ and /i/ are less in comparison with other vowels, they are much distinct in comparison with perceptual difference scores obtained for English [7].

In order to objectively verify that whether the overlapped vowels are perceptually distinct from one another, a perceptual distance between the overlapped vowels has been measured using Eq. (2).

$$D_{ij} = \sqrt{(M_{1,i} - M_{1,j})^2 + (M_{2,i} - M_{2,j})^2}$$

where $D_{ij}$ is the Euclidean distance between two vowel (i, j) points and $M_1$ and $M_2$ are the frequency of F1 and F2 expressed in Mel [5]. 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 Mel [7].
this procedure individual vowels are well separated by reducing the overlapping between them, on the other hand data of both the sexes are concentrated to form separate cluster for a particular vowel. The above said normalization procedure is effective in the sense that it is, in one hand, useful in acoustic phonetic study of vowels and on the other this procedure of vowel normalization preserves the relative position of vowels in vowel space as the raw formant data does. Fig. 6 represents the intrinsic-cum-extrinsic normalized value of first and second formant (henceforth DF1 and DF2) in mel scale.

![Figure 6: Plot of DF1 and DF2 in mel scale](image)

It can be observed from the above figure that almost all the vowels form well separated clusters by reducing the spread due to talkers except vowel /a/ and /ə/. The spread is more along the DF1 axis which is also reflected in raw formant plot.

5. Lip Rounding of Manipuri Vowels

The degree of lip rounding is measured by the ratio of height and width of the area between the opening of the lips. The measurement has been taken when the opening of the lips was maximum. Table 4 presents the height-width ratio of lip opening.

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Ratio between height and width</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/</td>
<td>0.96</td>
</tr>
<tr>
<td>/e/</td>
<td>0.93</td>
</tr>
<tr>
<td>/æ/</td>
<td>0.73</td>
</tr>
<tr>
<td>/i/</td>
<td>0.53</td>
</tr>
<tr>
<td>/ɪ/</td>
<td>0.45</td>
</tr>
</tbody>
</table>

It is interesting to note that there is a clear distinction in the values of ratio between the rounded and unrounded vowels. The rounded vowels have ratio closer to one whereas the ratio for unrounded vowel is less than 0.6. It is quite evident from Table 4 that value of the ratio for the phoneme /æ/ lies in the range 0.6 to 0.9, as the phoneme /æ/ is a central vowel and hence the spread of the lips is less than the rounded vowel but more than the unrounded vowel.

<table>
<thead>
<tr>
<th>Table 4: Ratio of height and width of Vowel</th>
</tr>
</thead>
</table>

Figure 7 shows the photographical evidence of lip rounding for six Manipuri vowels.

![Figure 7: Lip rounding of six Manipuri vowels](image)

6. Conclusion

Based on the experimental findings following conclusions for Manipuri vowels can be drawn

- There is no long and short variation of vowel in Manipuri language
- There is significant difference in the value of both F1 and F2 of male and female speakers.
- As per the experimental findings Manipuri has 2 rounded and 3 unrounded vowels.
- Incase of /a/ lip is neither rounded nor spread.
- Though some vowels are overlapped in the acoustic vowel space they are distinct in perceptual domain.

Keeping in mind the findings of Manipuri vowels the following is summarized in Table 5 along with existing description of the Manipuri vowel in the literature.

<table>
<thead>
<tr>
<th>Table 5: Articulatory description of Manipuri Vowel</th>
</tr>
</thead>
</table>

7. Acknowledgement

The Speech Processing Section of C-DAC, Kolkata is very much thankful to the informants and transcribers of Manipuri University. This paper is part of the project done under IPA-PLS consortium, MeitY, Govt. of India. The authors are grateful to MeitY, Govt. of India for supporting their research.
8. Reference