

Experimental Study of Vowels in Nagamese, Ao and Lotha: Languages of Nagaland

**Joyanta Basu, Tulika Basu, Soma Khan,
Madhab Pal, Rajib Roy**

Centre for Development of Advanced Computing
(CDAC), Kolkata
Salt Lake, Sector – V, Kolkata, India
{joyanat.basu,tulika.basu, soma.khan,
madhab.pal, rajib.roy}@cdac.in

Tapan Kumar Basu

Department of Electrical Engineering,
Academy of Technology
Aedconagar, Hooghly, West Bengal, India
basutk06@rediffmail.com

Abstract

This paper describes the vowels characteristics of three languages of Nagaland namely Nagamese, Ao and Lotha. For this study, nucleus vowel duration, formant structure (1st and 2nd formant i.e. F1 and F2) and intensity of vowels are investigated and analyzed for these languages. This paper includes the nasal context for different vowels and tries to examine its importance in different languages. A detailed analysis is carried out for six vowels namely /i/, /e/, /a/, /ə/, /o/, /u/ for readout speech of Nagamese, Ao and Lotha. Result shows that the vowel duration and formants play important roles in differentiating vowels characteristics. On the other hand, intensity of vowels do not play significant role in the characteristics of the vowels across the languages is observed. This initial study unveil the importance of vowels characteristics and may help to do research and development in the area of language identification, synthesis, speech recognition of three north-eastern languages of Nagaland.

1 Introduction

Culture and language diversity is one of the interesting phenomena in North-Eastern states of India. The seven states (i.e. Arunachal Pradesh, Assam, Meghalaya, Manipur, Nagaland, Mizoram and Tripura) except Sikkim of north-east India cover

an area of 255,511 square kilometers (98,653 sq mi) i.e. about seven percent of India's total area. As of 2011 they had a population of 44.98 million, about 3.7 percent of India's total population. Although there is great ethnic and religious diversity within the seven states, they bear similarities in the political, social and economic spheres (Wikipedia, 2015). According to the 1971 census there are about 220 languages spoken in these states, belonging mainly to three language families, namely Indo Aryan, Sino-Tibetan and Austro-Asiatic. The Indo-Aryan is represented mainly by Asamiya and Bangla, Austro-Asiatic is represented mainly by Khasi and the Sino-Tibetan family of languages is represented by Tani group of languages (Apatani, Galo, Nyishi etc.), Angami, Chakesang, Kuki, Manipuri, Mizo, Kokborak etc.

The entire North-east India is enclosed by major international borders of Bhutan, Nepal and China in the North and North-East, Bangladesh in South and West and Myanmar in East. This region is therefore very sensitive from the point of view of national security and national integrity.

Among the eight states of north-east India, the states that share far east international borders with Myanmar i.e. Manipur and Nagaland are getting much importance now-a-days due to unrest social and political situations. Since the last decade, spread of recent communication mediums like mobiles, telephones and VoIP are supporting spoken communication in regional north-east languages. Speech data in these communications has become

necessary for surveillance purposes. But detailed analysis on any speech data depends largely on pre-defined knowledge on the spoken language and availability of language resources. Unfortunately very little prior works have been done on the languages of Nagaland and Manipur.

In this study, we are mainly concentrating on major languages of the Nagaland state. These are Nagamese, Ao and Lotha. Apart from the basic language grammar, phonetic reader and dictionary very few linguistic resources are available for study and research purposes. Among the previous resources of Ao language, an important and detail study has been reported on the phonetic and phonological description of the Mongsen dialect of Ao (Alexander R. Coupe, 2003). The study is well supported by experimental findings as well as author's personal insights on the studied language. Lotha language is very rarely studied till date. Different aspects of Lotha language has been documented in (Chiang Chen Shan, 2011) which is the only available study on this language. Nagamese, the communication language (lingua franca) of Nagaland is quite well studied since 1921 by J.H. Hutton. The first ever linguistic study of Nagamese has been reported by M. V. Sreedhar (1974) in "Naga Pidgin: A Sociolinguistic Study of Interlingual Communication Pattern in Nagaland". This was followed by "Standardized Grammar of Naga Pidgin" by Sreedhar himself in 1985. B K Boruah's "Nagamiz Kothalaga Niyom - A Primary Grammar on Nagamese" (1985) and "Nagamese: The Language of Nagaland" (1993) are also some relevant studies to understand the basic structure and nature of Nagamese language. The last reported study on Nagamese language is "The structure of Nagamese the contact language of Nagaland" (2003) by Ajii Kumar Baishya of Assam University.

However, the three languages Ao, Lotha and Nagamese were never studied together to frame out similarities or differences in phoneme characteristics. Moreover except Ao language, unfortunately very little work on acoustic analysis has been done so far in two other languages.

2 Purpose of the Study

Main purpose of the study is to find out vowels characteristics of the major languages of Nagaland i.e. Nagamese, Ao and Lotha. Vowels possess one of the defining structures of any language in the

world. Their numbers, acoustic characteristics, particularly timbral ones need to be well defined for technology development (Tulika Basu and Arup Saha, 2011). For this work we have considered three important parameters of vowels like Formants, Nucleus Vowel Duration and Intensity. Nucleus vowel is defined as the steady state of the vowel along with the two transitions (Rajib Roy, Tulika Basu, Arup Saha, Joyanta Basu, Shyamal Kr Das Mandal, 2008) as in figure 1.

One way to objectively differentiate vowels is to examine the first two formant frequencies, namely 1st formant (F1) and 2nd formant (F2), which roughly correlate with tongue height and tongue position respectively (G. E. Peterson and H. L. Barney. 1952). In general high first formant is associated with narrow tongue constriction near the glottis. Similarly second formant frequency is increased as the constriction moves forward (K. N. Stevens and A. S. House. 1961). Using F1 and F2 it is possible to properly place them in a vowel diagram.

This study will help further for different applications like speech synthesis, language identification, speech recognition etc. in the target languages.

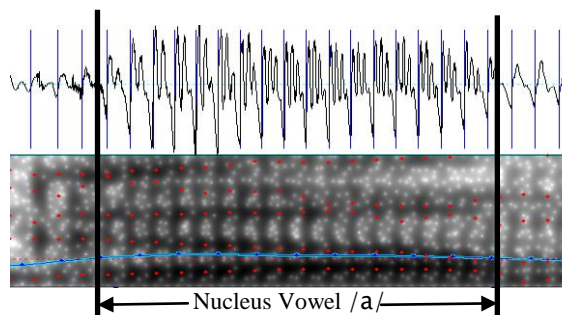


Figure 1. Nucleus Vowel Duration

3 Languages of Nagaland

Nagaland is a state in Northeast India. It borders the state of Assam to the west, Arunachal Pradesh and part of Assam to the north, Myanmar to the east and Manipur to the south. The state capital is Kohima, and the largest city is Dimapur. It has an area of 16,579 square kilometres (6,401 sq mi) with a population of 1,980,602 as per the 2011 Census of India. Nagaland is the home to 16 indigenous tribes namely Ao, Angami, Chang, Konyak, Lotha, Sumi, Chakhesang, Khiamniungan, Dimasa Kachari, Phom, Rengma, Sangtam,

Yimchunger, Kuki, Zeme-Liangmai (Zeliang) and Pochury as well as a number of sub-tribes. Each tribe is unique in character with its own distinct customs, language and dress. Nagaland is one of three states in India where most of the population is Christian (Wikipedia Nagaland).

As per Grierson's classification system, Naga languages can be grouped into three groups- Western, Central and Eastern Naga Groups. The Western Group includes Angami, Chokri and Kheza. The Central Naga group consists of Ao, Lotha and Sangtam, whereas Eastern Group comprises of Konyak and Chang. In addition, there are Naga-Bodo group illustrated by Mikir language, and Kuki group of languages illustrated by Sopvama (also called Mao Naga) and Luppa languages. These languages belong mostly to the Sino-Tibetan language family. Since most of these languages are mutually unintelligible, people depend on a pidgin language called Nagamese for communication. English has been used as the official language of the Nagaland state and it is quite popular among the educated mass of Nagaland. But Nagamese is used as the lingua franca among the various ethnic groups in the state. The languages of Nagaland state are not included in the scheduled list of twenty two languages in India, many of them spoken by dwindling number of speakers. This section presents a brief profile of major languages of the state. Figure 2 shows the languages map of Nagaland in different districts. Though, other dialects also exist in those districts, but from majority perspective they are not shown in the map.

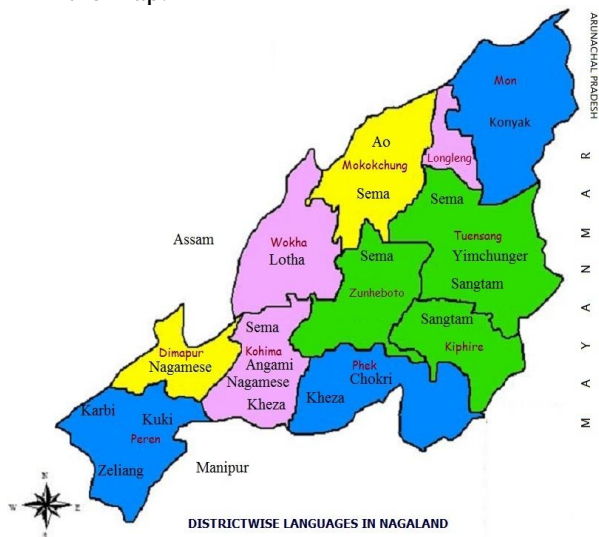


Figure 2. Languages of Nagaland in different districts

For this study we have selected three important languages of Nagaland i.e. Nagamese, Ao and Lotha.

3.1 About Ao Language

Ao is one of the important languages in Nagaland. Ao is spoken by a large number of people in the state. Mongsen, Chungli, Chanki etc. are prominent among the Ao dialects. Among all the dialects, Chungli is the widely spoken one and people of other Ao dialects can speak Chungli Ao but not vice versa. The inhabitants of the Mokokchung district mainly converse in this language. The vowel inventory of Chungli Ao is like this: /i/, /e/, /ɨ/, /a/, /ə/, /u/ (Bruhn Daniel, 2009). Nasality is not phonemic in Ao. It is a tonal language with 3 contrasting lexical tones: high, mid and low. All are register tones.

3.2 About Lotha Language

The Lotha language is part of the Sino-Tibetan language family, spoken by approximately 166,000 people in Wokha district, west-central Nagaland, India. It is centered in the small district of Wokha. This district has more than 114 villages such as Pangti, Maraju (Merapani), Englan, Baghty (Pakti) and others, where the language is widely spoken and studied. It is a medium of education up to the post-graduate level in the state of Nagaland. It is also the language in which the church sermons are preached. Lotha has seven dialects. They are Live, Tsontsu, Ndreng, Kyong, Kyo, Kyon and Kyou (Chiang Chen Shan. 2011). Lotha language has six vowels namely /i/, /e/, /a/, /ə/, /o/, /u/. Nasality is not phonemic in Lotha. Like other Tibeto-Burman languages it is a tonal language with three register tones (Low, Mid, and High).

Sl. No.	Nagamese Vowels	Ao Vowels	Lotha Vowels
1	/i/	/i/	/i/
2	/e/	/e/	/e/
3	/a:/	/a/	/a/
4	/ə/	/ə/	/ə/
5	/o/	NA	/o/
6	/u/	/u/	/u/
7	NA	/i/	NA

Table 1. List of Vowels of Nagamese, Ao and Lotha

Above Table 1 shows the list of vowels in three languages.

4 Experimental Data Set

The present study aims at finding out acoustic characteristics of vowels of different languages of Nagaland from readout text for different applications of speech processing in respective languages. For this purpose, text material in each of the three languages has been prepared including digits, numbers, units and paragraphs on different topics. The text material contains around 120 words and 60 sentences of different length. The text is read out by 15 native speakers from Nagaland in the age group between 20 to 40 years with 2 repetitions. All speakers are male speakers and English as their medium of primary education. Table 2 shows the detail Meta data of informants participated in this study.

Sl. No	Informant	Native-Language	Edu- Qualification	Age (in yr)
1	Speaker 1	Nagamese	Secondary	30
2	Speaker 2	Nagamese	Secondary	24
3	Speaker 3	Nagamese	Secondary	35
4	Speaker 4	Nagamese	Secondary	32
5	Speaker 5	Nagamese	Secondary	31
6	Speaker 6	Ao	Secondary	32
7	Speaker 7	Ao	Secondary	30
8	Speaker 8	Ao	Secondary	32
9	Speaker 9	Ao	Secondary	30
10	Speaker 10	Ao	Higher-Secondary	35
11	Speaker 11	Lotha	Secondary	34
12	Speaker 12	Lotha	Higher-Secondary	32
13	Speaker 13	Lotha	Graduate	26
14	Speaker 14	Lotha	Primary	37
15	Speaker 15	Lotha	Secondary	33

Table 2. Speakers' Meta data Information

5 Experimental Procedure

Steps for experimental procedure are as follows:

5.1 Data Collection

For the experiment purpose, speech data has been collected from native speakers of Nagamese, Ao and Lotha. To avoid disfluencies in reading, informants are instructed to read out the text material several times before final recording. Near about 3 hrs of speech data has been collected using Praat (Praat Website, 2016) software. Speech data is being

recorded in a less noisy studio environment with 16 bit 22050 Hz digitization format.

5.2 Data Transcription:

Above collected speech data has been transcribed into phone level using Praat tool. It is worth mentioning here that tone is not considered in the present study. For the present study, only Vowel (V) phonemes (like /u/, /o/, /ə/, /a/, /e/, /i/) have been marked by the transcribers. For simplification and ease of understanding, following symbols are used by transcribers during transcription. Those are 'u', 'o', 'ac', 'a', 'e', 'i' respectively. Transcribers are also instructed to mark the nasal contexts (N) of vowel occurrences. If a vowel is preceded by nasal consonants like /m/, /n/, /ŋ/ etc. then it is marked as N_V and if vowel is followed by nasal consonants then it is marked as (V_N). Then all phone level transcription files are saved as .TextGrid file format. Figure 3 shows the sample transcribed speech data using Praat tool. From the figure three panes can be observed. First one shows the time domain signal, second one shows the spectrographic view of time domain signal and final tier shows the phone level (only vowels) transcription boundary marked manually by transcribers. Transcribers need to zoom in and zoom out the signal and play it repeatedly to perceptually identify the vowels.

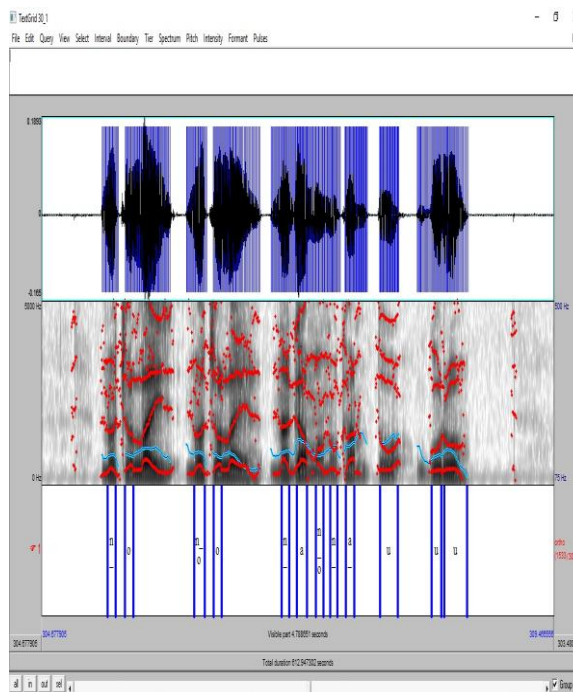


Figure 3. Transcription using Praat Tool

5.3 Extraction of duration, formants and intensity of vowels

Nucleus vowel duration, 1st formant (F1), 2nd formant (F2) and intensity are calculated using Praat scripts for further analysis. All vowels are segmented automatically using transcription output file i.e. from TextGrid file. These segmented files are required to test the perceptual appropriateness of different vowels by listeners.

6 Result and Discussion

Table 3 presents the number of vowel segments collected for analysis after transcriptions have been done. It has been observed that within our experiment data, occurrence of vowel /a/ is highest in Nagamese. Similarly vowel /u/ in Ao and vowel /o/ in Lotha language has the highest occurrence.

Sl. No.	Vowels	Nagamese	Ao	Lotha
1	/u/	448	694	405
2	/o/	569	NA	782
3	/ə/	578	670	605
4	/a/	809	527	525
5	/e/	576	579	576
6	/i/	735	547	623

Table 3. Vowel count in three languages under analysis

6.1 Analysis of Nucleus Vowel Duration

For present study on vowel duration, six vowels are considered including all vowel phonemes of Nagamese, Ao and Lotha languages. Those vowels (V) are /u/, /o/, /ə/, /a/, /e/, /i/ i.e. ‘u’, ‘o’, ‘ac’, ‘a’, ‘e’ and ‘i’ respectively. Using Praat scripts, nucleus vowel durations are extracted from transcription files. Vowels in Nasal context V_N and N_V are also analyzed in this study.

6.1.1 Vowel duration in Nagamese

Figure 4 shows the mean and +/- standard deviation of duration for each vowel in Nagamese including all contexts. It has been observed that nucleus vowel duration of Nagamese vowel /e/ i.e. ‘e’ is higher than other vowels and on the other hand duration of vowel /ə/ i.e. ‘ac’ is lesser than the others.

In Figure 5, nucleus vowel durations in preceding and succeeding nasal contexts are shown separately. It has been found that duration of vowels followed by nasal consonants i.e. V_N is always lesser than that of N_V and V. In all cases, duration of /ə/ i.e. ‘ac’ is less irrespective of nasal context. Duration of /e/ i.e. ‘e’ is highest in V and N_V cases. But duration of vowel /e/ i.e. ‘e’ is smaller than duration of /a/ i.e. ‘a’ in V_N.

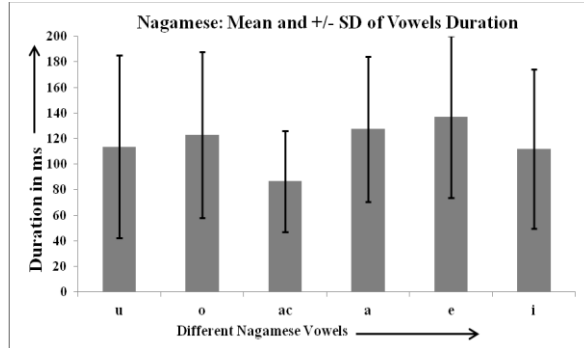


Figure 4. Nucleus vowel duration of Nagamese

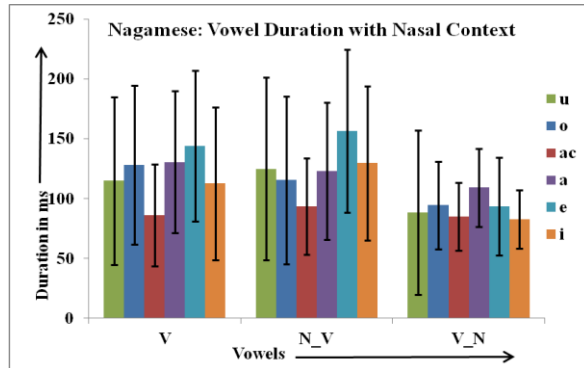


Figure 5. Nucleus Vowel Duration of Nagamese with nasal context

6.1.2 Vowel duration in Ao

Vowel inventory of Ao consists of /i/, /e/, /i/, /a/, /ə/, /u/. But the speech data which have used in this study does not contain any /i/ vowel. It is also interesting to note that though /o/ vowel is not included in the vowel inventory of Ao (be it in Chungli or Mongsen), in course of transcription /o/ vowel is found corresponding the grapheme ‘u’ like in words “tuko” which is pronounced sometimes as /tuko/ and sometimes as /toko/ by the native speakers of Chungli AO. This phenomenon is further supported by the previous

study on AO language where it has been mentioned that vowel /u/ and /o/ are in free variation in AO language (Alexander R. Coupe, 2003).

From figure 6 it has been observed that nucleus vowel duration of Ao vowel /a/ i.e. ‘a’ is higher than other vowels and on the other hand duration of vowel /ə/ i.e. ‘ac’ is smaller than others.

Vowel duration of Ao with nasal context has been shown in figure 7. All vowels followed by nasal consonants i.e. V_N are lesser in duration than that of N_V and V.

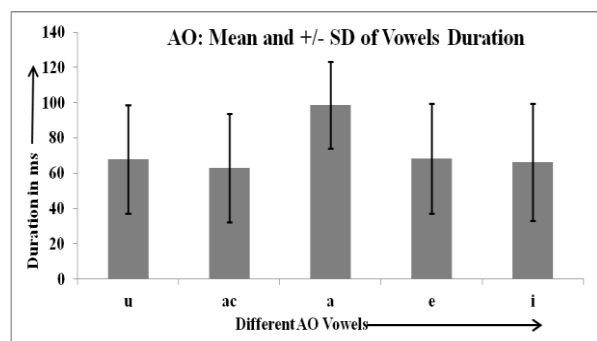


Figure 6. Nucleus Vowels Duration of Ao

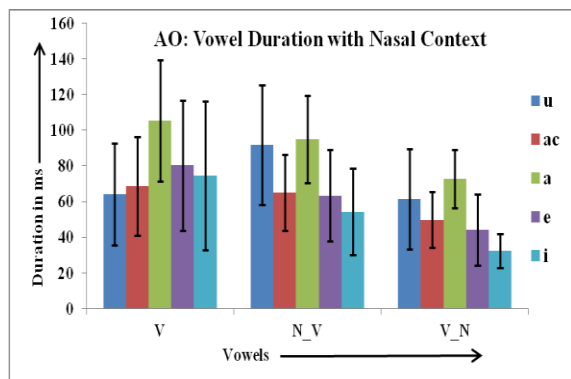


Figure 7. Nucleus Vowel Duration of Ao with nasal context

6.1.3 Vowel duration in Lotha

From figure 8 it has been observed that nucleus vowel duration of Lotha vowel /a/ i.e. ‘a’ is higher than other vowels like Ao language and on the other hand duration of vowel /ə/ i.e. ‘ac’ is smaller than others like Nagamese language.

From figure 9, it has been found that duration of vowel followed by nasal consonants i.e. V_N are smaller than duration of N_V and V except /e/ i.e. ‘e’. And in all the cases duration of /u/ i.e. ‘u’ is smaller irrespective of nasal context.

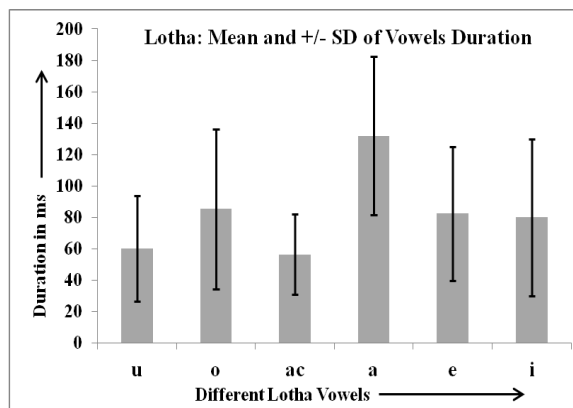


Figure 8. Nucleus Vowels Duration of Lotha

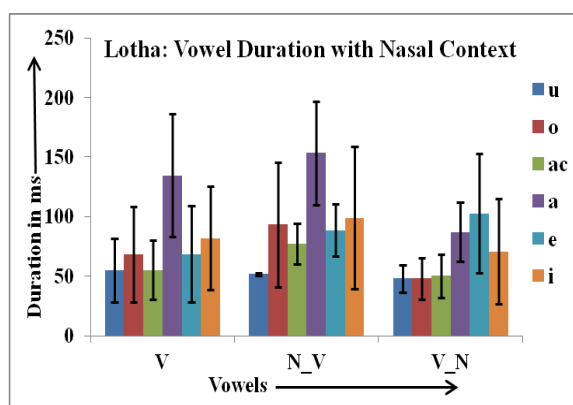


Figure 9. Nucleus Vowel Duration of Lotha with nasal context

6.2 Analysis of Vowel Formants

Formant analysis of the vowels in the three languages has been done including all contexts of their occurrence.

6.2.1 Vowel Formants in Nagamese

Figure 10 shows the F1 vs. F2 plot for Nagamese vowels. Six vowels have been observed and they are clustered in different zones. Zones of vowels /ə/ and /a/ are overlapped. But value of F1 for vowel /a/ is higher than the F1 value of /ə/. Some portion of /u/ and /o/ are also overlapped. But from then diagram it has been clearly identified the vowels like /i/ from /u/ or /e/ from /ə/ etc.

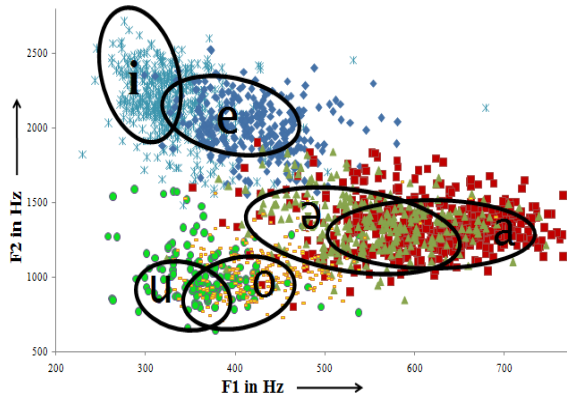


Figure 10. F1 vs. F2 of Nagamese vowels

6.2.2 Vowel Formants in Ao

Figure 11 shows the general F1 vs. F2 plot of different vowels in Ao language. It has been observed that there is a great amount of overlap in both the formant frequencies of vowel /u/ and /o/. So in this study these two vowels are merged together and analyzed as a single vowel /u/.

From the F1 vs. F2 plot it has been clearly identified the vowels like /i/, /ə/, /a/ and /u/. Values of /e/ and /ə/ are overlapped to some extent.

6.2.3 Vowel Formants in Lotha

Figure 12 shows the F1 vs. F2 plot of different vowels in Lotha language. In Lotha, six vowels have formed well separable clusters and thus they can be identified by F1 and F2 values. Only some cases it creates confusion for identifying the vowels /e/ and /ə/. In Lotha /u/ and /o/ vowels can be clearly identified by their values.

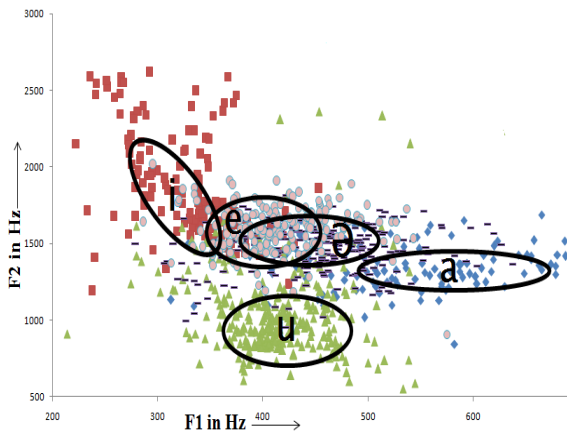


Figure 11. F1 vs. F2 of Ao vowels

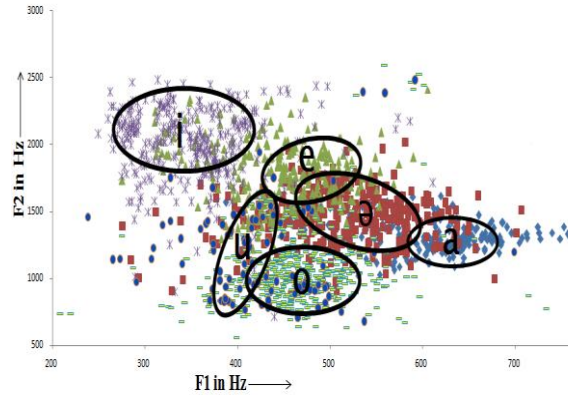


Figure 12. F1 vs F2 of Lotha vowels

Figure 13 shows overall comparison of F1 vs. F2 values of different vowels in Nagamese, Ao and Lotha languages. In the three languages (Nagamese, Ao and Lotha) F2 values have varied significantly for vowels /i/ and /e/ and mean F2 value of Nagamese is higher than that of Ao and Lotha.

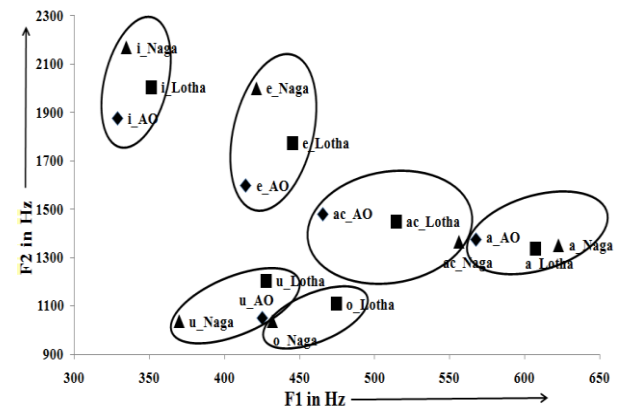


Figure 13. Comparison of F1 vs. F2 for Nagamese, Ao and Lotha vowels

On the other hand, F1 values have varied significantly for vowels 'ac' and 'a'. From this F1 vs. F2 comparison figure, all major vowels can be clearly identified by respective F1 and F2 values.

Changes in vowel formant characteristics for nasal context are also studied. Figure 14 and Figure 15 is showing the occurrence frequency distribution of F1 and F2 values respectively for different vowels in Nagamese. It has been observed from figure 14 that, there is no significant change of vowel 1st formant F1 with V_N and N_V context. But some changes of vowel 2nd formant F2 can be observed in figure 15 for V_N and N_V context. F2 values for vowels /ə/, /e/ and /i/ are showing

different frequency distribution with two or three major peaks. After careful observations from data it has been found that these peaks are coming due to nasal contexts of vowels. Frequency distribution of F1 and F2 for the other two languages Ao and Lotha has also been calculated and similar pattern has been found in those cases. No effect of nasal context on F1 has been observed for Ao and Lotha also. But F2 plays important role in nasal context. F2 of Ao vowels /u/ and /i/ is found to be affected for nasal context. Similarly F2 of /e/ is affected in Lotha.

6.3 Analysis of Vowel Intensity

During this study intensity of different vowels of Nagamese, Ao and Lotha are analyzed. Figure 16 shows the intensity wise graph of different vowels.

It has been observed that there is no significant change in intensity of different vowels of the three languages. Intensity of some Ao vowels like *u*, *o*, *ax*, *ac*, *a*, *i* is smaller than that of Nagamese and Lotha.

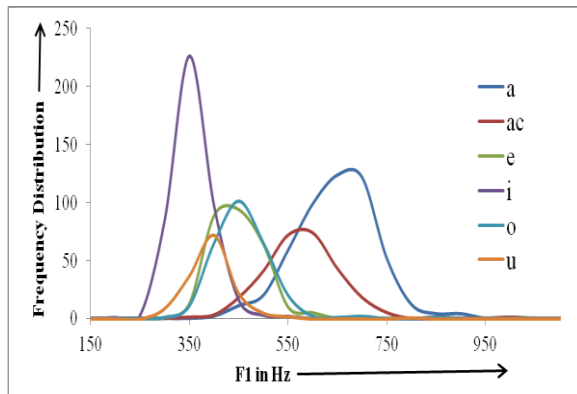


Figure 14. F1 of Nagamese Vowels with Nasal Context

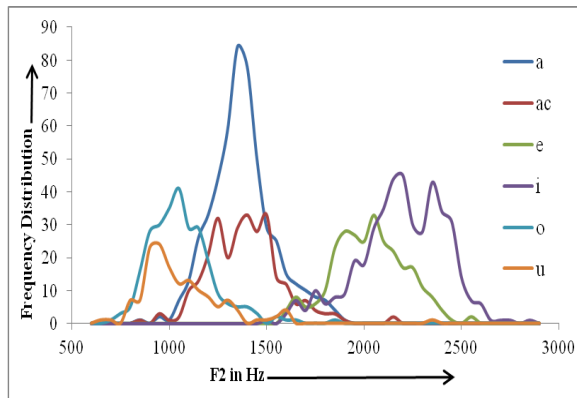


Figure 15. F2 of Nagamese Vowels with Nasal Context

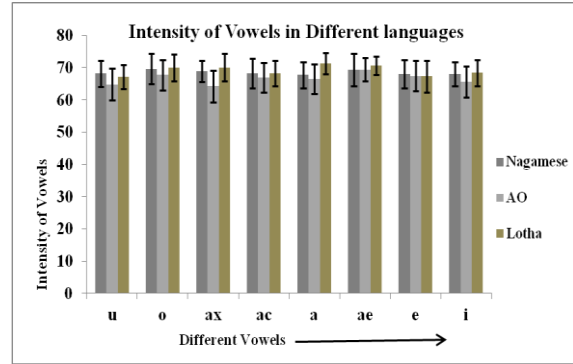


Figure 16. Intensity of Different Vowels of three languages of Nagaland

7 Conclusion

In this paper we have reported characteristics of vowels of three languages of Nagaland namely Nagamese, Ao and Lotha and carried out experimental study to find out language specific features. In this paper nucleus vowel duration, formant (F1 and F2) of vowels and intensity has been observed. The present study tried to find out significant influence on the nucleus vowel in presence of adjacent nasal phoneme i.e. preceding and succeeding nasal phoneme in all three languages.

In conclusion, the following points can be summarized for three languages of Nagaland:

- Nucleus vowel duration of vowel /e/, /a/ and /a/ is higher for Nagamese, Ao and Lotha respectively. The duration of /e/ in Nagamese is highest than the others may be due to the fact that most of the Nagamese verbs end with 'e' vowel and the speakers try to lengthen it to indicate the clause boundary.
- Similarly, duration of vowel /ə/, /u/ and /ə/ is smaller for Nagamese, Ao and Lotha respectively.
- In most of the cases duration of vowels followed by nasal consonants i.e. V_N is lesser than duration of N_V and V.
- In overall comparison of mean F1 vs. F2 all vowels are well separated in Nagamese, AO and Lotha.
- Vowels /u/ and /o/ are in free variation in Ao language because in F1-F2 plane they overlapped. Therefore samples of the two vowels are merged together and analyzed as a single vowel /u/ in Ao language.

- In case of nasal context i.e. V_N and N_V no significant influence of F1 has been observed for all three languages. But F2 plays important roles in nasal context. In frequency distribution of F2 of all vowels multiple peaks have been observed due to nasal context.
- No significant changes in intensity for different vowels are observed in all three languages.

However, there is scope of further study on vowels characteristics with respect to other different context like fricative, sibilants, plosives etc. This study may help the researches in the area of language identification, duration modeling for synthesis system as well as speech recognition on languages of Nagaland.

Acknowledgments

This work is a part of ongoing initiatives on “Deployment of Automatic Speaker Recognition System on Conversational Speech Data for North-Eastern states” under CDAC North-East grant. The authors are thankful to CDAC, Kolkata, India for necessary financial and infrastructural support. Authors like to thank user agency for enabling them to collect speech data on different languages under controlled environment at single place from a number of native speakers. They also like to thank Ms. Sushmita Nandi for her efforts in manual verification of recorded speech data and transcriptions.

References

- Alexander R. Coupe. 2003. *A Phonetic and Phonological Description of Ao: A Tibeto-Burman Language of Nagaland North-East India* (Pacific Linguistics, 543), Publisher: The Australian National University (2003), ISBN-10: 0858835193, ISBN-13: 978-0858835191
- Baishya, Ajit Kumar. 2004. *The structure of Nagamese: The contact language of Nagaland*, Silchar: Assam University (Doctoral dissertation).
- Bhim Kanta Boruah. 1993. *Nagamese: the Language of Nagaland*, Mittal Publications, New Delhi, India
- Boruah B.K.1993. *Nagamese: The Language of Nagaland*, Mittal Publications, New Delhi.
- Bruhn Daniel. 2009. *The Tonal Classification of Chungli AO Verbs*, UC Berkeley Phonology Lab Annual Report

Chiang Chen Shan. 2011. *Language Documentation of Different Aspects of Lotha, a Tibeto-Burman language of Nagaland, north-east India*, Division of Linguistics and Multilingual Studies, Nanyang Technological University

G. E. Peterson and H. L. Barney. 1952. *Control Methods used in Study of Vowels*, Journal of the Acoustical Society of America, vol. 24, no. 2, pp. 175-184

K. N. Stevens and A. S. House. 1961. *An Acoustical theory of vowel production and some of its Implications*, Journal of Speech and Hearing Research, vol.4

Praat Website. 2016. <http://www.fon.hum.uva.nl/praat/>

Rajib Roy, Tulika Basu, Arup Saha, Joyanta Basu, Shyamal Kr Das Mandal. 2008. *Duration Modeling for Bangla Text to Speech Synthesis System*, International Conference on Asian Language Processing 2008, Chiang Mai, Thailand, November 12-14, 2008

Tulika Basu and Arup Saha. 2011. *Qualitative And Quantitative Classification Of Bangla Vowel*, O-COCOSDA 2011

Wikipedia Nagaland.
<https://en.wikipedia.org/wiki/Nagaland>

Wikipedia. 2015.
https://en.wikipedia.org/wiki/Seven_Sister_States