himalayan linguistics

A free refereed web journal and archive devoted to the study of the languages of the Himalayas

Himalayan Linguistics

Acoustics of vowels in Angami

Viyazonuo Terhiija & Priyankoo Sarmah
Indian Institute of Technology Guwahati

ABSTRACT

This work investigates the acoustic properties of Angami vowels with the aim of definitively establishing the phonologically contrasting vowels in the language. Contrary to some previous studies that report seven monophthongs and multiple diphthongs, supported by acoustic-phonetic evidence, this study concludes that there are six monophthongs and two diphthongs in the language. For monophthongs, acoustic characteristics such as the first three formants (F1, F2, F3), and duration were explored. For diphthongs, the first two formants (F1, F2), their slope characteristics, and duration were explored. The salience of the vowels in terms of their acoustic properties was substantiated by statistical analyses.

KEYWORDS

Tenyidie, Angami, vowels, acoustic analysis

This is a contribution from *Himalayan Linguistics, Vol. 24(3): 39–70.*
ISSN 1544-7502

Tables of contents, abstracts, and submission guidelines are available at
escholarship.org/uc/himalayanlinguistics

Himalayan Linguistics, Vol. 24(3). © The authors 2025
ISSN 1544-7502

Acoustics of vowels in Angami

Viyazonuo Terhiija and Priyankoo Sarmah
Indian Institute of Technology Guwahati

1 Introduction

This paper aims to investigate the acoustic properties of monophthongs and diphthongs in Angami (also known as Tenyidie, ISO 639-3: njm). Besides a preliminary study reported in a dialect of Angami (Khonoma) by Blankenship et al. (1993) and a recent study on the acoustics of two southern Angami dialects (Terhiija et al. 2018), no other acoustic studies have been conducted on the Angami vowels. This paper also aims to strengthen the descriptive study of Angami vowels by incorporating objective acoustic analyses.

The descriptions of Angami vowels are conflicting in terms of the numbers and types of vowels reported. Several earlier studies on Angami vowels report six to seven vowels in the language. This discrepancy in Angami vowel description arises due to the status of the schwa vowel in the language which, sometimes is described as a mid-central vowel [ə], and as a high-central vowel [ɨ] at other times. Some studies have also assigned phonemic status to both [ɨ] and [ə]. Apart from the monophthongs, diphthongs in the language have not received much attention in previous works. Hence, this paper also aims to provide an acoustic-phonetic description of the diphthongs in Angami.

Angami is a lesser-studied language of the Tibeto-Burman language family spoken in India's northeast by 152,796 people in Nagaland (Burling 2003; Marrison 1967; Office of the Registrar General & Census Commissioner of India 2011). While the term Angami is considered an exonym, the language is referred to by its endonym, Tenyidie. The term Tenyidie subsumes several linguistic groups related to the Angami language. In this study, we explore the spoken form used by the Angami group, which is broadly divided into four sub-groups, namely, the Northern, Southern, Western, and Chakro,[1] based on their geographic location. Each group consists of several villages, each of which is said to have its own variety of the language. The current study investigates the Northern Angami sub-group, as spoken in the Kohima village. This variety, spoken in and around the Kohima village, is considered the standard form of Angami (Mahapatra et al. 1989; Ravindran 1974).

While linguistic analysis of Angami is scarce, there are works since the beginning of the 20th century that primarily describe the grammar of the language. Angami was first reported in McCabe's field report where he had given a brief description of the grammar and provided a list of words translated from English to Angami (McCabe 1887). General descriptions of the grammar of standard Angami are reported in the works of Burling (1960); Giridhar (1980); Kuolie (2006); Ravindran (1974) etc. One of the earliest preliminary acoustic studies of the language was on the voiceless nasals, as reported by Bhaskararao and Ladefoged (1991). Another early study provides a phonetic description of Khonoma Angami vowels, tones, and voiceless nasals (Blankenship et al. 1993). Of late, some acoustic phonetic works have also focused on Angami segmental and

[1] The term Chakro refers the native residents by the highway or the residents of lower altitude area.

suprasegmental features. For example, Terhiija et al. (2018) reports the characteristics of vowels produced in the Angami Kigwema and Viswema villages. Terhiija and Sarmah (2020; 2024(a); 2024(b)) report the phonetic properties of Angami nasals and stop consonants, respectively. In Lalhminghlui et al. (2019), how Angami vowels and tones phonetically affect each other is investigated. Meyase has studied Angami (Tenyidie) extensively in terms of its morphology and phonology (2016; 2021; 2024). He has examined the tonal phonology of the language. However, Meyase's work restricted to descriptive analyses of the Angami language. Besides the works in standard Angami, there are works done in dialects of the language as well. Dialectal variations and internal variations based on kinship ties and geography have also been reported (Suokhrie 2015; 2016). An attempt to develop speech corpus and automatic speech recognition in Angami is also one of the latest works on the language (Terhiija et al. 2019).

However, as seen above, there are no dedicated studies that reported the acoustic characteristics of the vowels in Angami. Hence, in the current study, we provide a description of the Angami vowels as spoken in the Kohima village variety. We also attempt to address the inconsistencies in terms of the number and type of vowels in the standard Kohima village variety reported in the earlier studies. Considering that, we provide literature survey on the monophthongs and diphthongs in Angami as reported in the previous studies. In the following Section 1.1, we provide a literature review of the monophthongs in Angami, followed by a literature review on the diphthongs in Section 1.2. Following this, the methodology of the current work is discussed in Section 2. Section 3 reports the results of the study and Section 4 concludes the work with a discussion.

1.1 Previous descriptions of Angami monophthongs

Except for the study on the Khonoma dialect, one of the western varieties of Angami, studies on Angami vowels have been impressionistic and did not use any acoustic description (Giridhar 1980; Kuolie 2006; Ravindran 1974). These studies have also resulted in contradictory accounts on the phonemic inventory of the Angami language. For example, in previous studies, the vowel inventory size proposed for Angami vowels ranges from six to seven phonemic vowels (Burling 1960; Ravindran 1974; Giridhar 1980; Kuolie 2006). While all of them agree on the five peripheral vowels in Angami /i, e, u, o, a/, they diverge on the description of the central vowels. Some report one mid-central vowel or one high central vowel, and others report the existence of both in the language.

The seventh vowel arises when authors differentiate between the mid-central vowel /ə/ and the high-central vowel /ɨ/.[2] In one of the earliest works on the language by McCabe (1887), the language is described to have six vowels /i, e, u, ɨ, o, a/. He described the central vowel as a high vowel. Studies on the Kohima variety by Burling (1960), Marrison (1967), Ravindran (1974), and Giridhar (1980) agree that there are six vowels in Angami /i, e, u, ə, o, a/, and the central vowel is a mid-central vowel. In the study of the Khonoma Angami variety, Marrison (1967) and Blankenship et.al (1993) also maintain that there are six vowels in the inventory. As opposed to these studies, Kuolie (2006) has described that there are seven vowels in the inventory. The author states that there are two central vowels; one is the high central vowel /ɨ/, and the other is the mid-central vowel /ə/. These descriptions of Angami vowels in the literature are summarized in Table 1.

[2] The high central vowel /ɨ/ is often represented as /ü/ in the studies reported and in the orthography.

Author (year)	Variety	No. of vowels	Vowel inventory
McCabe (1887)[3]	NA[4]	7	/i, e, u, ɨ, o, ö, a/
Burling (1960)	Kohima	6	/i, e, u, ə, o, a/
Marrison (1967)	Khonoma	6	/i, e, u, ə, o, a/
Marrison (1967)	Kohima	6	/i, e, u, ə, o, a/
Ravindran (1974)	Kohima	6	/i, e, u, ə, o, a/
Giridhar (1980)	Kohima	6	/i, e, u, ə, o, a/
Chase (1992)	Khonoma	6	/i, e, u, ə, o, a/
Blankenship et.al (1993)	Khonoma	6	/i, e, u, ə, o, a/
Kuolie (2006)	Standard	7	/i, e, u, ə, ɨ, o, a/
Meyase (2016)	Standard	6	/i, e, u, ə, o, a/
Suokhrie (2015)	Kohima	7	/i, e, u, ə, ɨ, o, a/

Table 1. Number and types of monophthongs in Angami as described in previous studies

Ravindran (1974) and Giridhar (1980) have described the central vowel for the word <ü> /ə/ 'poetry' as [ə]. However, Kuolie (2006) has made a distinction between a mid-central and a high central vowel claiming that they are both distinct phonemes. He assigned /ɨ/ for <ü> 'song' and /ə/ for <cü> /tʃə/ 'small'. Note that in the standard Angami orthography, also used by Kuolie (2006), both [ɨ] and [ə] are ascribed a single grapheme <ü>.[5]

The mid-central vowel, or schwa, denoted by the <ə> symbol, is one of the common vowels reported in Tibeto-Burman languages, including Angami. The schwa vowel is the most common non-peripheral vowel which occurs in languages with more than four vowels (Schwartz et al. 1997). Recasens (2022) observed in a recent cross-linguistic study of 110 languages that schwa most often occurs in languages with an inventory of seven vowels (20.9%), followed by languages with inventories of six, eight, and nine vowels (18.7%), and lastly in languages with inventories of four and ten vowels (12.1%). The same study showed that when the number of mid-vowels in the language system increases, the first formant (F1) is more variable among languages than the second formant (F2). The male speaker's formant frequency values for the schwa vowels are estimated to be 500 Hz (F1) and 1500 Hz (F2). These values correspond to the resonances of a vocal tract of 17.5 cm long (Fant 1970).

According to cross-linguistic studies, the acoustic properties of the schwa vowel vary depending on the context. In languages such as English and Dutch, the schwa vowel arises due to vowel reduction and hence, is restricted to unstressed syllables. According to Flemming (2007; 2009), there are two kinds of schwas in English: schwa that occurs word-finally and schwa that occurs word-

[3] McCabe (1887: 5) used /ö/ for the word /krö/ (*to wash, to buy*).

[4] The examples given in the texts seem to be from the Southern variety.

[5] The <ü> is also written as <ii> in Angami, possibly due to typographic convenience.

medially. The authors observed that the F2 of word-medial schwa is highly variable due to the co-articulatory effects of the contextual phonemes. A similar case of F2 variability has also been reported in Dutch (Beinum 1994). The duration of schwa in English when produced word-finally is relatively longer than when it is produced in other contexts. Flemming (2009) reports that non-final schwa vowels averaged 64 milliseconds (ms) in duration, while the word-final schwa vowels had a mean duration of 153 ms. In Dutch, overall, the duration of the schwa vowel is reported to be shorter than the other vowels (Beinum 1994). Hence, while schwa is a commonly occurring vowel in systems larger than four vowels, it is a phonetically vulnerable phoneme, varying extensively due to contextual and prosodic effects.

1.2 Previous descriptions of Angami diphthongs

Previous literature describes Angami diphthongs with an inventory size that ranges from one to six. Marrison (1967) listed a single diphthong, [ie], in the Khonoma variety of Angami. While Kuolie (2006) and Suokhrie (2015) have listed two diphthongs namely, [ie, uo], for the Kohima variety, Burling (1960), Marrison (1967), and Ravindran (1974) have reported four diphthongs for the same variety. Giridhar (1980) has listed six diphthongs. Table 2 summarizes the number and types of diphthongs described in previous works.

Author (year)	Variety	No.	Vowel description
Burling (1960)	Kohima	4	/ie, uo, əi, uu/
Marrison (1967)	Khonoma	1	/ie/
Marrison (1967)	Kohima	4	/ie, uo, ei, ou/
Ravindran (1974)	Kohima	4	/ie, uo, ou, əi/
Giridhar (1980)	Kohima	6	/ie, uo, əi, io, eo, ou/
Chase (1992)[6]	Khonoma	7	/ie, ei, ai, ia, io, oi, ui/
Kuolie (2006)	Standard	2	/ie, uo/
Suokhrie (2015)	Kohima	2	/ie, uo/

Table 2. Description of Diphthong vowels in Angami by various authors

As seen in Table 2, [ei, ou] are two of the most commonly occurring diphthongs in the previous studies. Besides the two common diphthongs, there are other diphthongs which are described. The inconsistencies in the description of Angami diphthongs arise partially due to the orthographic convention used for the language. The Angami orthography uses [ei] for the /i/ sound, while [ou] is used for /u/ sound. This has prompted Marrison (1967), Ravindran (1974), and Giridhar (1980) to describe the orthographic <ei> and <ou> as diphthongs.

The other issue in the previous descriptions of the diphthongs is the phonemic representation of the [əi] sound. Burling (1960), Ravindran (1974), and Giridhar (1980) provided examples of the sound [əi] occurring in words such as [zəi] 'dark'; [təi] 'sky' and [pfhenəi] 'clothes', however the [əi]

[6] Chase (1992) notes that diphthong /ie/ is the most frequent in occurrence, whereas the others are marginal.

in these words is pronounced as a /i/. The phonetic representation of these words will be [zi], [ti], and [pf̠eni], respectively, in the standard variety of the language. Burling (1960) also describes [uu] as ending in an offglide, resembling a diphthong in words such as, [suu] /su/ 'deep' and [rəluu] /rəlu/ 'war dance'. Again, the pronunciation of these words does not seem to contain any dynamic movement of the vowels, prompting us to consider them as pure monophthongs. At the same time, Angami does not have phonemic vowel duration, hence, the [u] and the [uu] are not distinct in terms of duration.

The identification of vowels in Angami is further complicated by the existence of suffixes that consist of only vowel nuclei. For example, the definitive suffix -u and the nominative suffix -e. While in derived representations they may seem to be diphthongs, they are considered vowel sequences as described by Kuolie (2006). The examples in (1) and (2) show the derivation of the definitive and nominative lexemes. In these examples, the resulting sequence of vowels are pronounced separately as two morae and not as a glide, which may be true for many other vowel sequences, mistakenly considered as diphthongs in the language in previous descriptions.

(1) ha-u
this-DEF
'This one'

(2) a-e
1SG-NOM
'I'

As mentioned in the earlier sections, some of the misinterpretation of the Angami phonemes emanates from the writing system of the language. Angami is written using the Roman orthography and it has used the orthography innovatively to meet the requirement of the language. Table 3[7] summarizes some of the vowel representations in the Angami orthography and provides their corresponding International Phonetic Alphabet (IPA) transcriptions. As seen in the table, the vowel /i/ is represented both as <ei> and <i>. Similarly for /u/ both <ou> and <u> are used.[8] It should be noted that the authors did not find any vowel duration differences or tone differences among the various ways of writing.

[7] Angami has 4-5 level tones. Each of the 5 tones in Angami, namely, T1-T5 as reported in the Angami dictionary, MKS Dieda (Liezietsu et al. 2019). T1 being the highest and T5 being the lowest; T2, T3 and T4 are the intermediate tones.

[8] The authors examined words from one of the earliest written documents in Angami, such as the Hymns by Rev. S.W. Rivenburg (1903), where vowels are represented in two ways to signify a single sound.

Orthography	IPA	Meaning	Tone
mei	/mi/	tail	T1
mi	/mi/	fire	T1
nuo	/ɲo/	child	T4
nou	/nu/	heart	T4
nu	/nu/	inside	T4
pfü	/pf̪ə/	carry	T5

Table 3. Orthography vs IPA of the vowels in Angami

Taking into consideration the existing literature on diphthongs, it is evident that the large number of diphthongs reported in previous literature is due to multiple reasons, such as misinterpretation of the Angami orthography, and considering vowel sequences of distinct morae as vowel glides. However, phonemically there are only two diphthongs in Angami, namely, [ie] and [uo]. Hence, in the current study, we consider only [ie] and [uo] for acoustic analysis. The methodology adopted for analyses of vowels is reported in the following section.

2 Methodology

This study investigates the six monophthong and two diphthong vowels in open syllables as shown in Table 4. As for the monophthong vowels, the supposed high central vowel [ɨ], written as <ü>, was examined and was compared with the mid-central vowel /ə/ and other peripheral vowels. For this, a wordlist comprising words found in MKS Dieda (Liezietsu et al. 2019) as well as in Kuolie (2006) was considered. Kuolie (2006) mentions that [ɨ] can occur in V, CV and VC type of syllables, while /ə/ is restricted to monosyllables and to the initial V in disyllables. The distinctiveness of the two central vowels is discussed in detail in Section 3.2 of this work.

Considering the discussion above, we will investigate the acoustic-phonetic characterization of the monophthongs and diphthongs in Angami. The monophthongs in the language are analyzed using their first three formant frequencies and the duration, whereas the diphthongs are analyzed using the first two formants (F1, F2) and their slopes. The detailed methodology is discussed in the subsections to follow.

IPA	Word	Meaning	Tone
a	t͜sa	less	T1
e	t͜se	break	T1
ə	t͜sə	sprout	T3
i	t͜si	over	T2
o	t͜so	reach	T3
u	t͜su	went	T3
ie	tie	lips	T3
uo	tuo	future marker	T4

Table 4. A minimal set in Angami showing contrastive monophthongs and diphthongs

2.1 Speakers

Thirty-three Angami native speakers (16 female and 17 male) were recorded for conducting the current study of Angami monophthongs and diphthongs. Of the 33 speakers, 28 were born and raised in the Kohima village, Nagaland. The speakers were chosen from Kohima village as the variety spoken in this area is regarded as the standard variety of the Angami language. The remaining 5 speakers were language teachers who have a minimum of master's degree in Tenyidie (Angami). From the subset of the data, 21 speakers' speech data was used for the diphthong study. The average speaker's age is 34.9 years (sd = 8.6) at the recording time. All the speakers were native speakers of Angami (Tenyidie). All the participants could also speak English (official language of the Nagaland state) and Nagamese (lingua franca) besides Angami. All the speakers had a minimum education of a bachelor's degree, and some were working professionals. They were all proficient readers of Angami.

2.2 Recording of data

The recordings were conducted in a noise-controlled environment either at the speaker's home or in the village community building. Data was collected at different times by the first author from 2018-2022. The recordings were made using a Shure unidirectional head-worn microphone (model: Shure SM10-CN) connected to a Tascam linear PCM recorder (model: TASCAM DR-100MKII) through an XLR jack. The sampling frequency at the recording was 44.1 kHz, 24 bits in WAV format.

2.3 Data type

The speech data collected was designed to incorporate all 6 monophthongs and 2 diphthongs in Angami. Data for the vowels was derived from monosyllabic and disyllabic words in Angami.[9] Each word recorded in three contexts, namely, *sentence*, *carrier* and *isolation*. In the *sentence* context, the target word appeared in a natural sentence, while in the *carrier* context the target word was embedded in a carrier sentence which was constant for all words and iterations. In the *isolation*

[9] Angami restricts syllable types to V, VC and CV type of syllables and hence, a majority of the data collected was in the CV syllable structure.

context the speakers were asked to produce the target word in isolation. An example of the target word in three contexts is provided in Table 5. In case of a small subset of the data, words containing the vowels /i/ and /ə/ were produced only in isolation in CV syllables where the onset was always a nasal. As this additional data was also included for analyses, the number of /i/ and /ə/ is relatively higher as seen in Table 6. As seen in the table, we analyzed 12,066 monophthongs and 734 diphthongs in total for the current study.

Context	IPA	Meaning
Sentence	puo a ne ʃə	s/he pushes me
Carrier	a ne puba	I said push
Isolation	ne	push

Table 5. Example of vowel target sound /ne/ produced in the three environments

Vowels	Sentence	Carrier	Isolation	Total
a	545	552	680	1,777
e	545	569	723	1,837
ə	640	655	1290	2,585
i	725	750	1032	2,507
o	531	537	655	1,723
u	484	491	662	1,637
ie	131	132	183	446
uo	85	89	114	288
Grand Total				**12,800**

Table 6. Distribution of monophthong and diphthong tokens in the current analysis

2.4 Annotation

After recording the speech data, they were transferred to a computer for analysis. The sound files were manually segmented, with vowel boundaries using Praat 6.0.43 (Boersma & Weenink 2020). Beginning and end of steady state formants in the vowel segments were considered the vowel boundaries. After that the vowels were visually inspected using the waveforms and spectrograms. Further, the segmented files were cross-checked three times to avoid any errors in annotation and segmentation. As seen in Figure 1, the vowel boundaries, along with the consonant boundaries, were annotated.

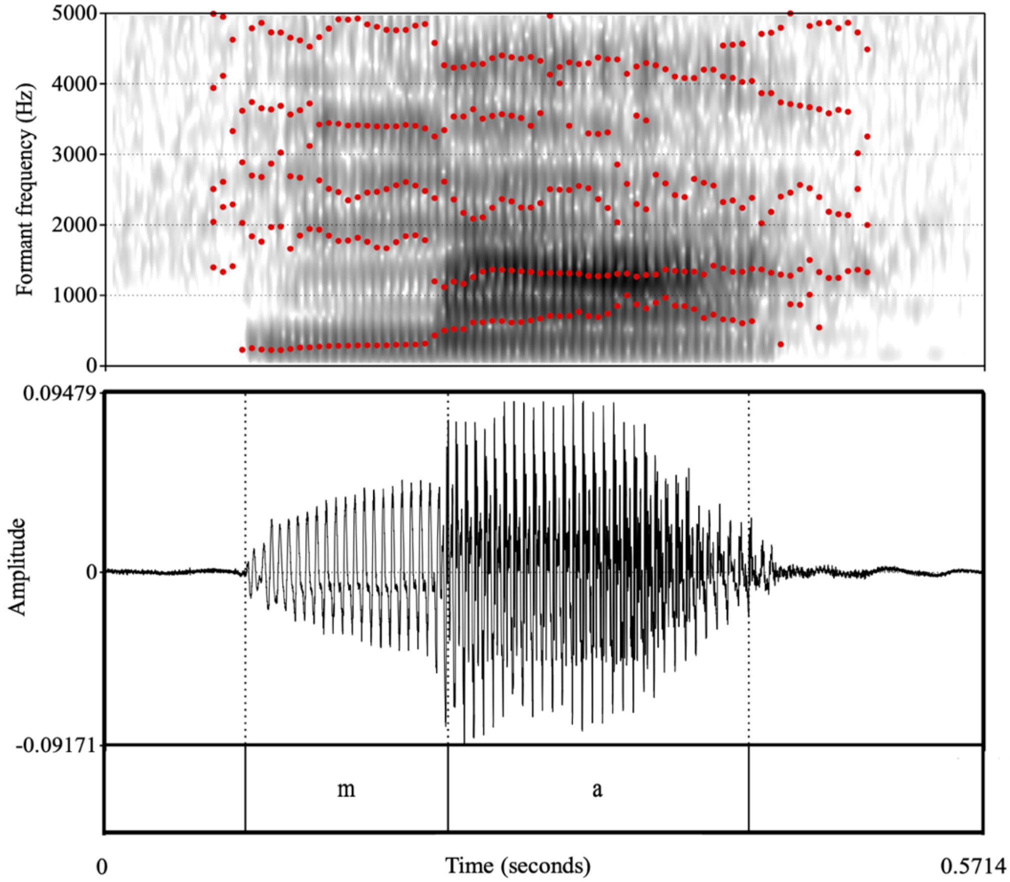

Figure 1. Annotation example showing the vowel boundary of the word /ma/ 'fire' as produced by a female speaker

2.5 Acoustic measurements

Vowels articulatory features such as height, backness and roundness can be ascertained with acoustic features, namely the first three formant frequencies. The first formant (F1) is inversely proportionate with the height of the vowel, the second formant (F2) is directly proportionate to the frontness of the vowel and the third formant (F3) is inversely proportionate to the roundness of the vowel. Hence, to investigate the acoustics of Angami vowels, we have extracted the F1, F2, and F3 values of the vowels from the midpoint of the vowels. Apart from the vowel quality features (F1, F2, and F3), we also extracted vowel quantity information as expressed by the duration of the vowels. A Praat script was used to extract the F1, F2 and F3 values in the middle of the vowel. The same script also measured the total duration of the vowels. Formant ceilings for females were at 5500, while for males they were at 5000 following Williams and Escudero (2012).

Considering the dynamic properties of diphthongs, F1 and F2 values were extracted at every 2% of the total duration of the vowels. To estimate the slope of the F1 and F2 trajectories, the values of the formant contours were converted to discrete cosine transform (DCT) coefficients. The first three coefficients, namely, C0, C1 and C2 were used to characterize the diphthongs. In case of the DCT coefficients, C0 provides the average formant values calculated from the entire formant trajectory, C1 provides information about the slope of the formant trajectory by comparing it to a falling half cycle of a cosine curve and C2 provides values corresponding to the formant trajectory when it is compared to a full cycle of a cosine curve. Hence, a falling trajectory has high positive C1 values, and a raising trajectory has high negative values. Similarly, a convex contour results in high positive C2 values whereas a concave contour results in high negative C2 values.

While raw Hertz values were extracted for vowel formants, considering speaker variability, the raw formants were normalized using the z-score normalization method (Rose, 1987). Speech data collected from multiple speakers usually vary across speakers due to anatomical, physiological, and talker-specific factors. These factors systematically affect the fundamental frequency and the formant frequencies of vowels. This variation arising due to speaker-specific features is considered an artifact and speech data is usually processed with speaker normalization (Peterson 1952). In a study by Adank et al. (2004), the z-score method of frequency normalization used by Lobanov (1971) for vowel normalization, has been found to be the best vowel normalization metric, which is a vowel extrinsic, talker intrinsic normalization procedure. Hence, this method reduces the anatomical variation of the formant frequencies but maintains phonological and sociolinguistic variation. The equation for z-score speaker normalization is provided in Equation (1) where N_i is the normalized value of the i^{th} formant and F_i is the raw value of the i^{th} formant produced by a specific speaker. $\bar{x}$ is the mean and σ is the standard deviation all the i^{th} formant produced by the speaker. In the current study, formant normalization was conducted with by using the *normLobanov* function from the *phonR* package on the open-source R platform (R Core Team 2022).

(1) $$N_i = \frac{F_i - \bar{x}}{\sigma}$$

2.6 Statistical analyses

Utilizing R's statistical packages (R Core Team 2022), descriptive and exploratory statistics were calculated. While exploratory statistical modeling, such as the Linear Mixed Effects (LME) models, was carried out using the *lme4* package on R (Bates et al. 2014), descriptive statistics like mean, standard deviation, and standard errors were calculated using R's core functions. Using the Anova function in the *car* package, Type II Wald chi-square tests were performed to determine the significance of the fixed effects (Fox and Weisberg 2019). Finally, using the *emmeans* package, Bonferroni post-hoc tests were carried out to estimate the contrasts between the levels in the fixed factors (Lenth 2017). The plots were produced with R using the *ggplot2* package (Wickham 2016).

3 Results

3.1 Monophthongs in Angami

The data used in this study is drawn from several different datasets. Hence, it is important to evaluate the normality of the data. To assess this, we examine both skewness and kurtosis. Skewness indicates the symmetry of the data, whereas kurtosis indicates the magnitude of the tails. A symmetrically distributed data has skewness between -0.5 and 0.5. While kurtosis values around 3 indicate a normal distribution, values more than 3 and less than 3 are short-tailed and heavy-tailed distribution indicators, respectively.

To visualize the three formants' differences in the six Angami monophthong vowels, density plots for normalized values of the three formants are presented in Figure 2. The figures also accompany the kurtosis and skewness values for each vowel representing the normalcy of the distribution of the normalized F1 and F2 values. As seen in the figure, vowels /a/, /e/, and /ə/ have distributions peaking at 2, 0, and 0.5 in normalized F1, respectively, which are symmetrical considering their skewness values. As indicated by the kurtosis values, /e/ and /ə/ have normal normalized F1 distribution while /a/ is slightly short tailed. Similarly, the vowel /i/, /o/, and /u/, have F1 distributions peaking at -1, 0, and -1 hertz, respectively. The F1 values for all three vowels are highly skewed. The high kurtosis values indicate that the F1 of these three vowels has a short-tailed distribution.

In Figure 2, we also investigated the distribution of the normalized F2. Vowels /a/, /e/, and /ə/ have distribution peaking at 0.5, 1, and 0 hertz, respectively, and they have symmetrical distribution based on their skewness values (between -0.5 and 0.5). Except for vowel /e/, the other vowels in F2 values have short-tailed distribution. Overall, we see that the vowels /a/, /e/ and /ə/ have symmetrical distribution showing normality in the distribution of the data based on the normalized F1 and F2 values, whereas vowels /i/, /o/, and /u/ their distributions are highly skewed and short-tailed, as shown by their high kurtosis values. The high kurtosis and skewness in vowels /i/, /o/, and /u/ shows that there is a potential variation in articulation of these vowels across speakers. Also, the back vowels /o/ and /u/ show significant variability. However, this may not be of concern as it is a cross-linguistic trend for back vowels to be more variable (Becker-Kristal 2010).

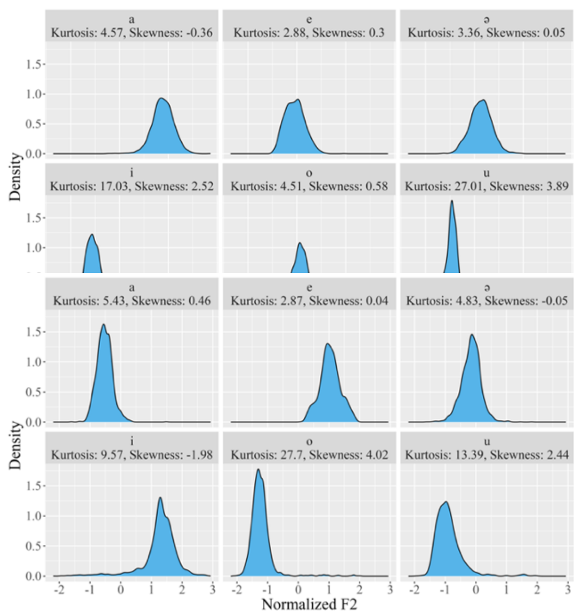

Figure 2. Skewness and Kurtosis graph of the six vowels in Angami representing the normalcy of the distribution of the normalized F1 (top) and F2 (bottom) values

3.2 Acoustic features of monophthongs in Angami

To examine the vowel characteristics, the first two formant frequencies for each vowel were measured. While the first formant (F1) frequency gives information about vowel height, the second formant (F2) frequency provides information on vowel backness (Ladefoged 1996). Both F1 and F2 frequencies of the vowels were calculated at the vowel midpoint in Hertz (Hz) values so that vowel-specific formants are captured without the co-articulatory features arising due to the preceding and the following consonants. After extracting the vowel formants, vowels were normalized, and formant means were plotted with one standard deviation ellipses. The following subsections show the distribution of the vowel formant values and the mean vowel values.

3.2.1 Vowel density

Vowel Density helps us to see how uniformly the F1 and F2 values of each vowel are distributed. It shows whether the vowels are produced in one distinct way or in more than one way. When the vowels are distributed uniformly, we get one core, otherwise, there may be more than one core visible. Vowel density plots further help in understanding the pattern of vowel dispersion. The vowel density plots of the six vowels were plotted for visualization in Figure 3. As seen in Figure 3 (left), the values without normalization have two cores which are prominent in vowels such as /i/, /e/, and /ə/. However, when formant values were normalized using the Labonov normalization method (equation shown in Section 2.6) and were plotted as shown in Figure 3 (right), multiple cores disappeared. As seen in the figure, after normalization of the values, the vowels are distributed uniformly and there is only one core for all the vowels. This makes us conclude that the bimodal distribution of the formants in Figure 3 (left) arises because of gender on formant values. This also justifies the use of vowel normalization methods in acoustic analyses of vowels.

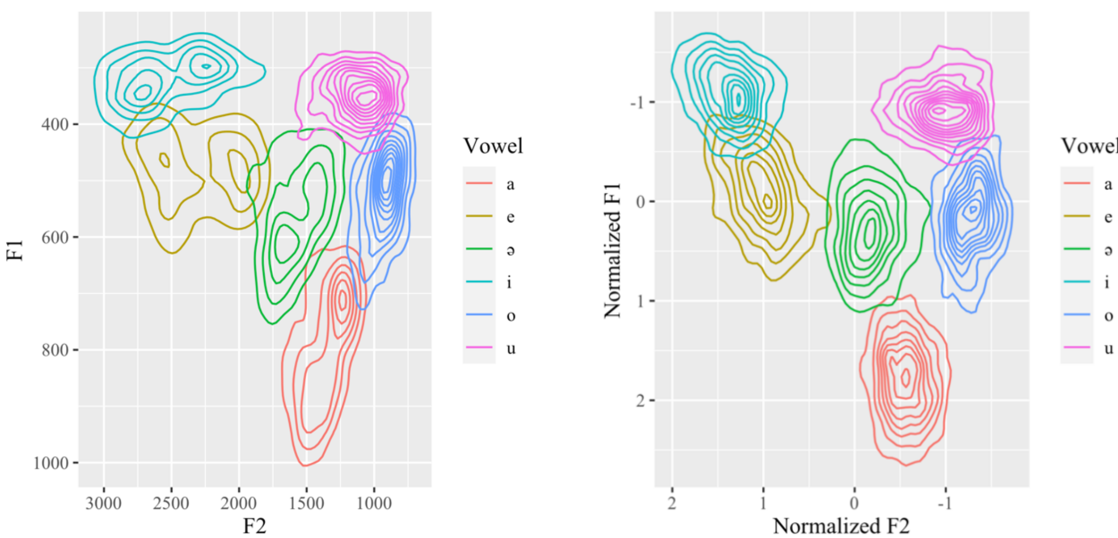

Figure 3. Vowel density plot showing the distribution and expansion of vowels in Angami in terms of F1 and F2 values, the left plot shows the raw formant values and the plot on the right normalized formant values; number of vowels =12066

3.2.2 Vowel quality

It is claimed that in Angami, there are two central vowels (see Section 1.2); one that occurs without any onset or coda as a V type syllable, and the other that occurs in CV type of syllables. We marked the former as <ü> and the latter as a standard /ə/. These two vowels were plotted separately in Figure 4 (left) to see whether they show any distinctiveness or not. When plotted it was seen that the two vowels were completely overlapped. This makes us conclude that acoustically there is no distinction between the two vowels and hence, they should be treated as belonging to the same phonemic category.

As we conclude that there are only six vowels in Angami, the six-vowel inventory is plotted in Figure 4 (right). However, the question remains if the vowels have durational differences in V type syllables, and this is discussed in detail in Section 3.2.3. The front close vowel /i/ and front close-mid vowel /e/ show proximity of the ellipses, whereas the back close vowel /u/, central vowel /ə/, back close-mid vowel /o/ and front open vowel /a/ are distinct from the rest of the vowels and are independent vowels without any overlapping ellipses. The mean values and standard deviation of the F1, F2, and F3 of the six vowels in Angami are displayed in Table 7. As seen in the table, the F1 value is inversely related to the height of the vowel where the low vowel /a/ has a high F1 and the two high vowels /i/ and /u/ have a low F1. As for F2 values, the front vowels /i, e/ have higher F2 values whereas the back vowels have lower F2 values.

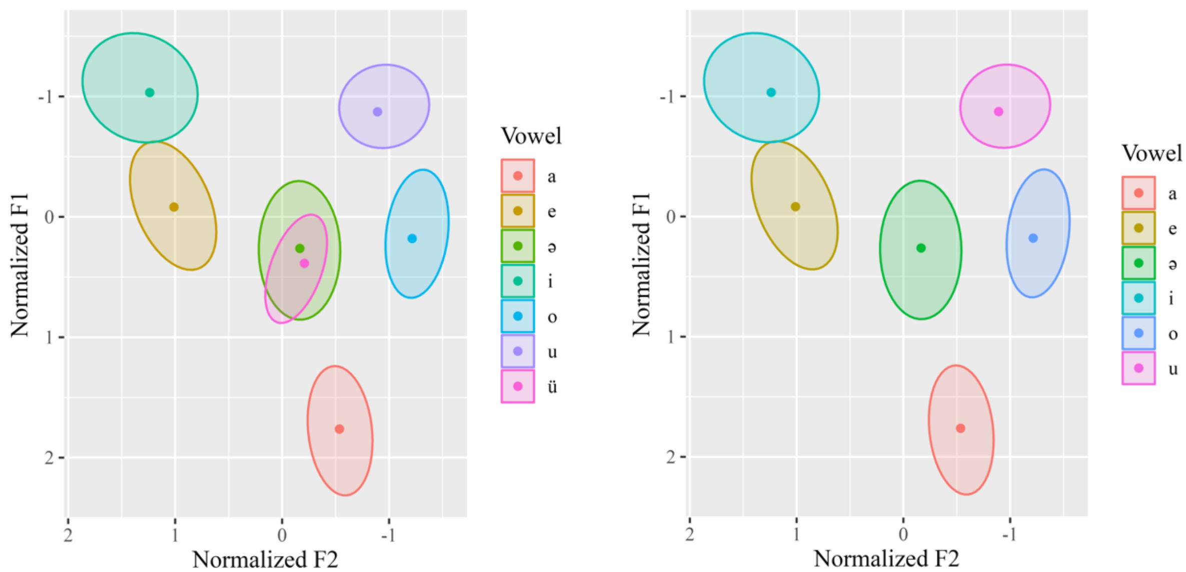

Figure 4. Normalized F1 and F2 with one standard deviation ellipses for 7 vowels (left) and 6 vowels (right) in Angami; number of vowels =12066

Vowels	F1	F2	F3
i	334 (67)	2441 (417)	3071 (303)
e	488 (75)	2304 (315)	2908 (325)
ə	555 (97)	1605 (222)	2748 (394)
a	793 (104)	1371 (164)	2667 (395)
o	535 (87)	966 (220)	2794 (358)
u	362 (81)	1167 (261)	2670 (316)

Table 7. The mean values and standard deviation (SD) of F1, F2, and F3 of the six vowels in Angami

We look further into the variation of the vowel spectral dispersion based on gender as plotted in Figure 5. It is clear from the figure that the vowels produced by male speakers do not have any overlapping ellipses and are distinct from each other. However, in the case of the female speakers, the front close vowel /i/ and front close-mid vowel /e/, the ellipses overlap minimally. Table 8 provides the mean values and standard deviation of the three formants across the six vowels by gender. We see that female speakers have higher formant Frequencies (F1, F2 & F3).

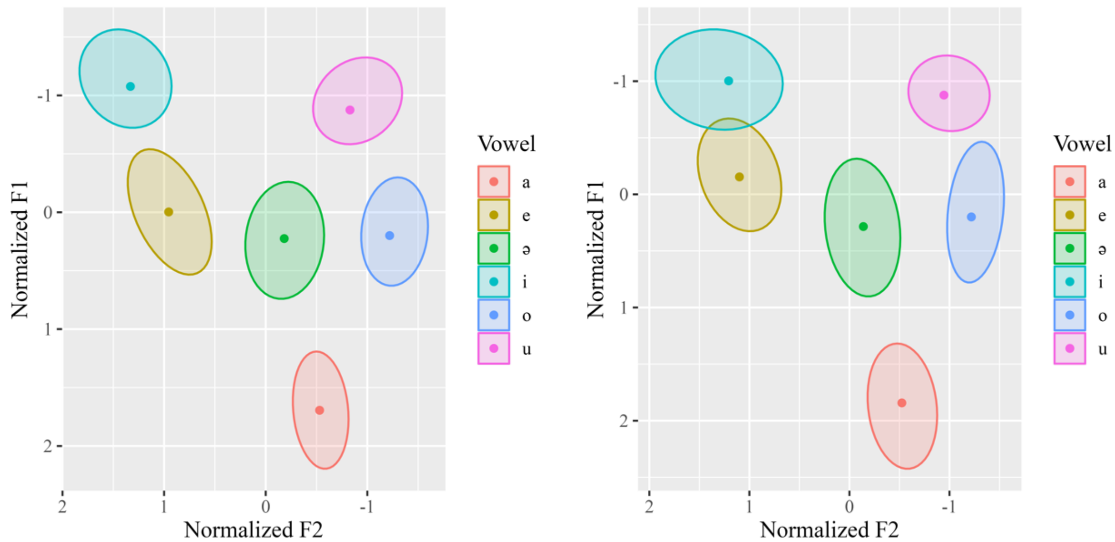

Figure 5. Normalized F1 and F2 with one standard deviation ellipses for Male (left) and Female (right) speakers; number of vowels =12066

	Male			Female		
Vowels	F1	F2	F3	F1	F2	F3
i	320 (67)	2250 (285)	2885 (222)	344 (63)	2609 (435)	3237 (271)
e	477 (62)	2043 (172)	2669 (199)	497 (81)	2523 (217)	3105 (267)
ə	513 (73)	1469 (163)	2574 (219)	588 (95)	1701 (199)	2891(418)
a	720 (68)	1284 (126)	2502 (217)	864 (85)	1452 (154)	2819 (444)
o	507 (64)	918 (218)	2562 (246)	564 (98)	1003 (208)	2999 (307)
u	352 (85)	1131 (262)	2469 (247)	370 (76)	1195 (250)	2884 (228)

Table 8. Mean values and standard deviation (SD) of F1, F2, and F3 of the six vowels in Angami as produced by Male and Female speakers

To confirm the statistical significance of the formant frequencies (normalized F1 and F2) among the six vowels, we constructed an LME model where the dependent variables were the two normalized formant frequencies, F1 and F2. Vowels, context, gender, tones, and their interactions were the fixed effects, while speaker, syllable type, structure type (mono or disyllabic), and individual words were the random effect, as shown in (1). The model was subjected to a Type II Wald χ^2 test for analysis of deviance (ANOVA) using the *car* package.

Formants ~ vowels + context + gender + tones + vowels*context*gender*tones + (1|speaker) + (1|syllable) + (1|structure) + (1| code) (1)

The results of the ANOVA conducted on the LME models are summarized in Table 9 and Table 10. The results show vowel types, tone, and context significantly affect the first formant (F1), however, gender, has no effect on F1. At the same time, all the interactions, including the ones with gender are significant. In case of F2, as shown in Table 10, only vowel and context have significant effects. However, all interactions have significant effects on the F2. In case of F3, as seen in Table 11, vowel and context has significant effects, however, when the χ^2 values are compared, the effect of vowel types is more pronounced in case of F1 and F2, indicating these two formants are better predictors of the vowels. Figure 6 displays the interaction plot of the LME models, providing a visual representation of how the variables of interest, such as gender, tone, and vowel, interact and impact the formant frequencies. In terms of F1, we see very minimal effects of tone and gender. For example, the vowel /e/ is nominally higher in terms of F1 in Tone 1 (T1) than in Tone 4 (T4) or Tone 5 (T5). Similarly, in most of the cases, the F1 of female speakers appear to be slightly higher than the male speakers. In case of F2, we can see sporadic effects of gender however, no consistent effects of tones. The trends in F1 and F2 are not uniform and hence, we can conclude that the normalized data for the Angami vowels are consistent. However, in case of F3, we notice clear effects of gender and tones, making us conclude that F3 may not be a reliable measure of vowel quality in Angami.

	χ^2	df	*p*-value
vowel	41674.0	5	<0.001
context	574.2	2	<0.001
gender	0.17	1	0.67
tone	70.4	4	<0.001
vowel × tone	245.13	20	<0.001
vowel × context	315.1	10	<0.001
vowel × gender	102.3	5	<0.001
vowel × context × tone	92.4	40	<0.001
vowel × gender × tone	**68.9**	**20**	**<0.001**
Conditional R^2 = 0.85, Marginal R^2 =0.79			

Table 9. Results of an ANOVA conducted on an LME model with normalized F1 as the dependent variable and vowel, context, gender, tone, and their interactions as fixed effects. Only statistically significant interactions are shown in this table

	χ^2	df	*p*-value
vowel	48316.5	5	<0.001
context	17.1	2	<0.001
gender	0.2	1	0.60
tone	8.7	4	0.06
vowel × tone	159.2	20	<0.001
vowel × context	82.2	10	<0.001
vowel × gender	165.3	5	<0.001
vowel × context × tone	78.8	40	<0.001
vowel × context × gender	21.2	10	<0.05
Conditional R^2 = 0.86, Marginal R^2 =0.83			

Table 10. Results of an ANOVA conducted on an LME model with normalized F2 as the dependent variable and vowel, context, gender, tone, and their interactions as fixed effects. Only statistically significant interactions are shown in this table

	χ^2	df	*p*-value
vowel	2257.4	5	<0.001
context	13.5	2	<0.001
gender	0.19	1	0.65
tone	131.4	4	<0.001
vowel × tone	231.2	20	<0.001
vowel × context	59.0	10	<0.001
vowel × gender	229.0	5	<0.001
vowel × gender × tone	**80.8**	**20**	**<0.001**
Conditional R^2 = 0.32, Marginal R^2 =0.26			

Table 11. Results of an ANOVA conducted on an LME model with normalized F3 as the dependent variable and vowel, context, gender, tone, and their interactions as fixed effects. Only statistically significant interactions are shown in this table

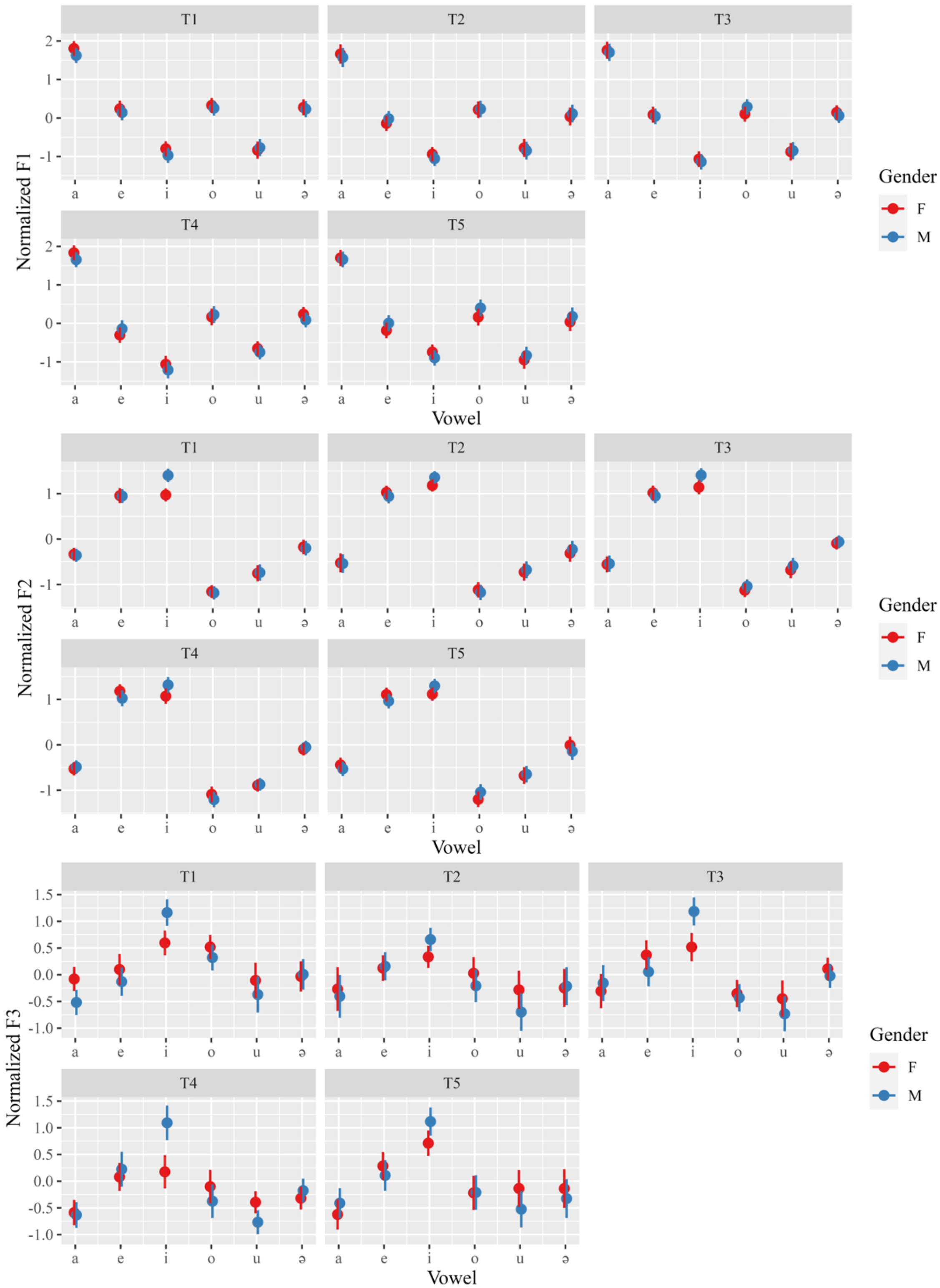

Figure 6. Interaction plot from the LME models where normalized F1 (top), F2 (mid), and F3 (bottom) are dependable variables, showing the interaction of gender, tone, and vowels

3.2.2 Monophthong duration

Apart from the vowel quality, we also extracted vowel quantity information as expressed by the duration of the vowels. A pretest was administered before conducting the vowel duration test. This preliminary test involved three female speakers producing the four meaningful vowels, namely, /a, ə, o, u/, that can occur in lexical words without an onset or coda. The vowels were in two contexts, i.e., in isolation and in meaningful sentence frames. They were repeated three times by each speaker, resulting in seventy-two tokens. This preliminary vowel duration study aimed to determine whether these four isolated vowels exhibited similar temporal characteristics or if there are differences among them. Table 12 provides the tokens used in the test.

Vowel	Meaning	Tone
/a/	*1st person pronoun*	T3
/ə/	*song, poetry*	T1
/o/	*okay, agreement*	T3
/u/	*3rd person pronoun*	T1

Table 12. Example of a minimal set of vowels, where they can appear in a lexical word without an onset or a coda

Figure 7 presents the results of the preliminary test on vowel duration, illustrating the average mean and standard errors of the four vowels when produced in isolation. The findings indicate that three out of four vowels exhibit similar durations, falling within the 260–280 milliseconds range, while the duration of vowel /u/ is at 190 milliseconds. The result showed that vowels, /a, ə, o/ have similar duration. However, the vowel /u/ is produced with an exceptionally short duration.

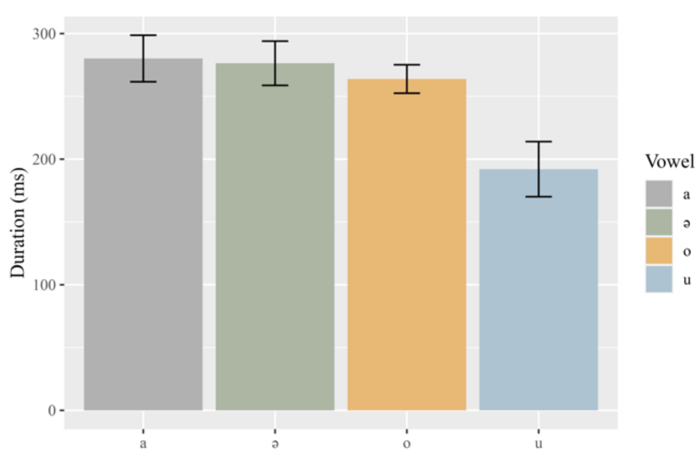

Figure 7. Duration of vowels where they can appear in a lexical word without an onset or a coda; number of vowels = 72

Contrasts	Estimates	SE	df	*t*-ratio	*p*-value
a – ə	2.31	21.9	57.1	0.106	1.0000
a – o	7.65	21.5	57.1	0.356	1.0000
a – u	79.44	21.5	57.1	3.701	< 0.01
ə – o	5.34	14.2	55.0	0.377	1.0000
ə – u	77.13	14.2	55.0	5.441	<.0001
o – u	71.79	13.7	54.8	5.237	<.0001

Table 13. Pairwise comparison of the vowels that appear as a lexical word without an onset or a coda

To confirm the results statistically, an LME model was constructed, with vowel duration as the dependent variable and vowel as the fixed effect, with speaker, context, and repetition as random effects. A Type II Wald χ^2 test for analysis of deviance revealed that vowel duration is indeed significant across the four vowels ($\chi^2 = 43.59$, $p < 0.001$). The LME model was subjected to a post-hoc Bonferroni test to compare the vowels in pairwise manner. The result of the pairwise test is presented in Table 13, which indicates that the vowel duration of /a, ə, o/ are not statistically significantly different. However, the vowel /u/ is significantly shorter than the other three vowels.

The second part of the vowel duration study involved analyzing 12,066 tokens. Figure 8 illustrates the average duration of the six vowels, and the average values and standard deviations are provided in Table 14. As seen in the figure and table, vowels /a, e, i, o/ are in the mid-range with a duration ranging between 185–195 ms, while vowels /u/ (175 ms) and /ə/ (202 ms) are the shortest and longest, respectively. We conducted further analysis to explore vowel duration differences in terms of gender and context. The results, as shown in Figure 9 with the duration values in Table 15, indicate that vowels produced in isolation exhibit longer durations compared to when they are produced within sentences or carrier frames. On the other hand, vowels produced in isolation are longer. The duration of the vowels is gender-specific too, as we see in Figure 9, the female speakers have longer duration in both sentence and carrier frames. In the isolation frame, the duration of /ə/ is relatively longer in females than males.

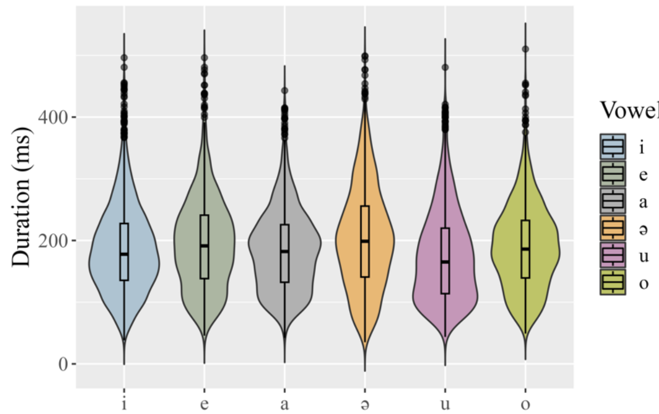

Figure 8. Violin plot of the average duration of the six Angami vowels (n = 12066)

Vowels	Mean	SD
a	184.2	66.0
e	194.0	74.1
ə	202.3	83.2
i	186.0	71.0
o	190.0	69.0
u	175.0	75.0

Table 14. Average vowel duration means and standard deviations of six Angami vowels

An LME (Linear Mixed Effects) model was constructed to assess the significance of vowel duration within and across various contexts. This model had duration as the dependent variable, with vowels, tone, context and gender as fixed effects. Speaker, words, syllable types and syllable structure types (mono or disyllabic) as the random effects, as indicated in (2). The results, as shown in Table 16, revealed that vowel, tone and context all had significant effects on the vowel duration. The χ^2 values showed that the context had the most significant effect, followed by vowel types and tones. All the interactions of the fixed effects had significant effects on vowel duration.

Duration ~ vowels + context + gender + tones + vowels*context*gender*tones + (1|speaker) + (1|syllable) + (1|structure) + (1|code) (2)

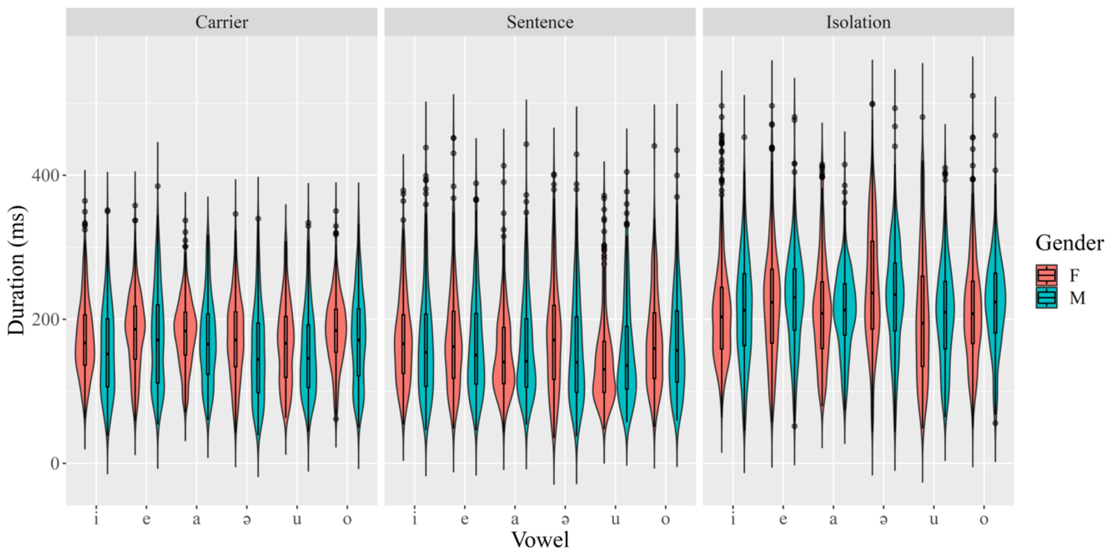

Figure 9. Violin plots of the average duration of the six Angami vowels by contexts and gender; number of vowels =12066

	Female			Male		
	Sentence	Carrier	Isolation	Sentence	Carrier	Isolation
Vowels	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
a	154.5 (59)	181.8 (47)	212.5 (71)	172.5 (71)	169.7 (59)	212.3 (57)
e	172.5 (71)	182.2 (54)	223.3 (79)	163.2 (69)	174.0 (68)	228.4 (69)
ə	175.0 (76)	174.0 (57)	246.0 (84)	158.0 (76)	153.0 (66)	230.2 (73)
i	171.1 (61)	174.5 (55)	211.9 (72)	166.1 (75)	157.6 (62)	212.4 (73)
o	169.1 (65)	184.0 (49)	212.3 (73)	167.2 (71)	172.0 (63)	221.0 (64)
u	144.2 (63)	165.7 (56)	203.2 (89)	154.0 (71)	155.0 (59)	207.0 (70)

Table 15. Average vowel duration means and standard deviation in six Angami vowels by three contexts and gender

A post-hoc Bonferroni test was conducted on the model where duration was the dependent variable and vowels served as the fixed effects for further examination of pairwise differences in duration across vowels. The results are shown in Table 17 which shows that the back vowels /o/ and /u/ are significantly shorter than the non-back vowels and the vowel /u/ is the shortest vowel as discussed earlier.

	χ^2	df	*p*-value
vowel	99.1	5	<0.001
context	1595.1	2	<0.001
gender	1.3	1	0.24
tone	53.5	4	<0.001
vowel × tone	845.3	20	<0.001
vowel × context	67.1	10	<0.001
vowel × gender	36.5	5	<0.001
vowel × context × tone	200.5	40	<0.001
vowel × gender × tone	45.4	20	<0.001
vowel × context × gender × tone	62.3	40	<0.05

Conditional $R^2 = 0.47$, Marginal $R^2 = 0.17$

Table 16. Results of an ANOVA conducted on an LME model with vowel duration as the dependent variable and vowel, context, gender, tone, and their interactions as fixed effects. Only statistically significant interactions are shown in this table

Contrasts	Estimate	SE	*z*-ratio	*p*-value
a – e	-6.46	2.35	-2.75	0.0894
a – ə	-0.08	2.34	-0.03	1.0000
a – i	-5.45	2.36	-2.31	0.3135
a – o	3.87	2.46	1.57	1.0000
a – u	**13.68**	**2.51**	**5.44**	**< 0.0001**
e – ə	6.38	2.20	2.90	0.0556
e – i	1.01	2.14	0.47	1.0000
e – o	**10.33**	**2.39**	**4.31**	**< 0.001**
e – u	**20.13**	**2.51**	**8.03**	**< 0.0001**
ə – i	-5.37	2.08	-2.58	0.1473
ə – o	3.95	2.35	1.67	1.0000
ə – u	**13.75**	**2.40**	**5.72**	**< 0.0001**
i – o	**9.32**	**2.31**	**4.03**	**0.0008**
i – u	**19.13**	**2.52**	**7.57**	**< 0.0001**
o – u	**9.81**	**2.54**	**3.86**	**< 0.01**

Table 17. Results of a Bonferroni post hoc test conducted on an LME model with vowel duration as the dependent variable and vowel as fixed factors

3.3 Acoustic features of diphthongs in Angami

As discussed in Section 1.2, there are two diphthongs in Angami, namely, /ie/ and /uo/. The first three discrete cosine transforms (DCT) coefficients of the F1 and F2 formant contours were extracted to examine the diphthong characteristics. In contrast to the approach used for monophthongs, where vowel data is plotted from the midpoint of the vowel, the analysis of diphthongs involved extracting F1 and F2 values at every 5% interval of the total duration of the diphthong formant contour. This method was employed to capture the entirety of the diphthong glides. Additionally, to minimize contextual effects, 10% of the onset and offset were excluded from the analysis.

To visually represent the location of the diphthongs within the vowel space, the F1 and F2 values of the vowels were plotted, overlaid by 80% of the F1 and F2 values of the diphthong vowels, as depicted in Figure 10. The figure illustrates that the onset of the diphthong /ie/ occupies the vowel space typically associated with the vowel /i/, while its offset aligns with the vowel /e/. Similarly, the onset of the diphthong /uo/ corresponds to the vowel space of /u/, with its offset aligning with the vowel /o/. The figure effectively highlights the trajectory of the diphthongs within the vowel space. The first three DCT coefficients, as discussed in Section 2.5, also allow one to reconstruct the original smoothened formant contour. Hence, the inverse discrete cosine transform (IDCT) method was used to reconstruct the formant contours of the two diphthongs as seen in Figure 11. Table 18 displays the mean values and standard deviations (SD) of F1 and F2 at 30% and 70% of the two diphthong vowels in Angami.

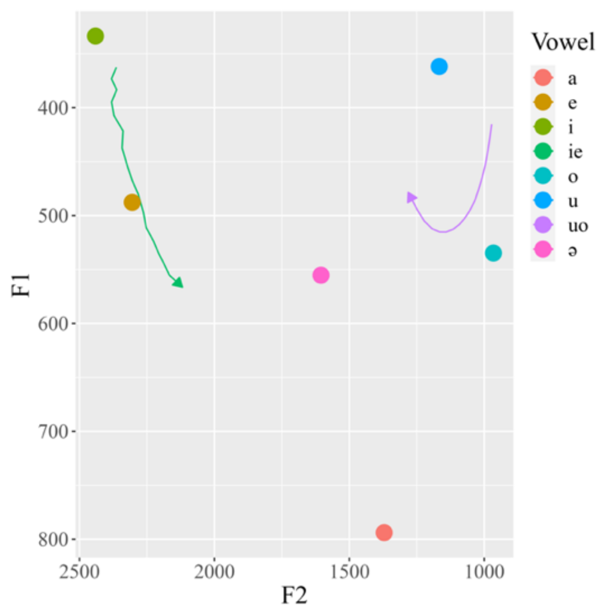

Figure 10. The formant values of F1 and F2 for monophthongs and diphthongs showing the space the diphthongs occupy; number of diphthongs = 734

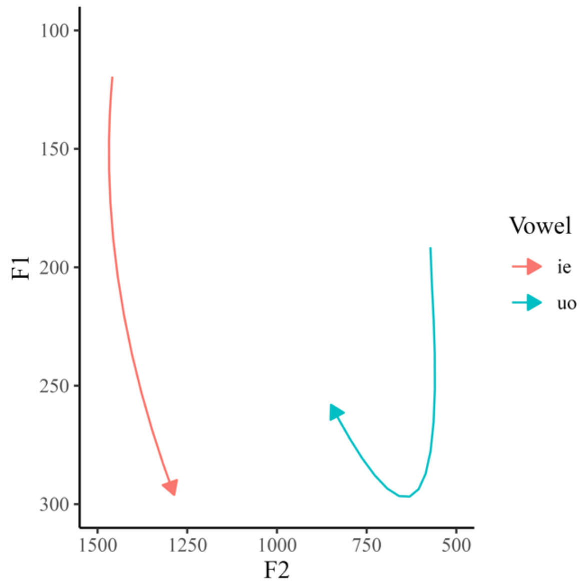

Figure 11. Inverse discrete cosine transform contours of F1 and F2, calculated from the mid-80% of the total duration of the two diphthongs in Angami; number of diphthongs = 734

	30%		**70%**	
Vowels	F1	F2	F1	F2
ie	407 (51)	2371 (342)	523 (100)	2224 (312)
uo	464.5 (59)	1016 (142)	518 (93)	1169 (251)

Table 18. The mean values and standard deviation (SD) of F1 and F2 of the two diphthong vowels in Angami

The variation in diphthong production by male and female speakers was further examined by plotting the IDCT values of F1 and F2 in Figure 12. Additionally, the mean values are summarized in Table 19. As observed in both the figure and the table, male speakers tend to have lower F1 values compared to female speakers at 30% and 70% of the total duration for both diphthongs. Conversely, in both diphthongs, the F2 values of female speakers are higher than those of male speakers. These findings underscore the gender-related differences in diphthong production within the Angami language, particularly in terms of F1 and F2 characteristics.

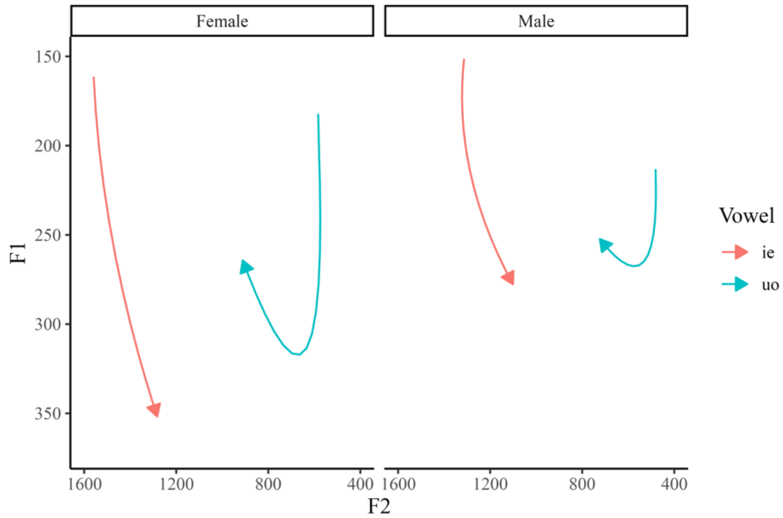

Figure 12. Inverse discrete cosine transform contours of F1 and F2 calculated from the mid 80% of the total duration of the two diphthongs for male (left) and female (right) Angami speakers (n= 734)

		30%		**70%**	
	Vowels	F1	F2	F1	F2
Male	ie	381 (36)	2155 (138)	454 (64)	2044 (165)
	uo	431 (51)	942 (101)	458 (54)	1060 (185)
Female	ie	425 (53)	2516 (362)	570 (93)	2346 (330)
	uo	483 (55)	1057 (145)	552 (94)	1230 (262)

Table 19. The mean values and standard deviation (SD) of F1 and F2 of the two diphthong vowels in Angami as produced by Male and Female speakers

To analyze the differences in formant values across the diphthongs, we constructed six Linear Mixed Effects (LME) models. These models were built with C0, C1, and C2 as dependent variables for both F1 and F2. Vowel, context, gender, and their interactions were the fixed effects, while speaker and unique words were the random effects. As C0 merely represents the average formant value of a formant contour, for F1, there were no significant effects of the vowel (diphthong) type on the C0 of F1, as seen in Table 20. However, the two diphthongs have distinct F2 spaces, which resulted in significant differences of the C0 of F2 for vowel types. For C0 of F1, context and gender both had significant effects with their significant interactions with the vowel types. In terms of C0 of F2, there was no effect of the context, but gender did have a significant effect which also interacted significantly with the vowel types. This clearly showed that average F1 and F2 varied significantly by gender. In Table 21 and Table 22, the dependent variables are C2 of the DCT coefficient, which represents values when a formant contour is compared to a half and a full cycle of a cosine curve, respectively. As the formant contours of the two diphthongs are distinct, it is expected that both C1 and C2 will have a significant effect of the vowels (diphthongs). The C1 values for both F1 and F2, are significantly different, as seen in Table 21. These shapes also have significant effects of the context but no effect of gender. Nevertheless, the shape of the contours does change within each vowel category by gender as indicated by the significant effect of the interaction of vowel and gender in the table. In case of C2, vowel types and context have significant effects on the C2 values. The results reported in these tables show that the formant contours are significantly different for the two diphthongs, and they are affected by the context (isolation or sentence). On the other hand, gender did not have consistent effects on the diphthong formant contours.

		F1		**F2**	
	df	χ^2	*p*-value	χ^2	*p*-value
vowel	1	3.1	0.1	1250.3	<0.001
context	2	181.6	<0.001	2.3	0.31
gender	1	26.8	<0.001	37.3	<0.001
vowel × context	2	56.5	<0.001	0.0	0.99
vowel × gender	1	31.2	<0.001	71.9	<0.001
Conditional R^2 = 0.66 (F1), 0.95 (F2); Marginal R^2 =0.36 (F1), 0.89 (F2)					

Table 20. Results of an ANOVA conducted on an LME model with C0 (F1 & F2) as the dependent variable and vowel, context, gender, tone, and their interactions as fixed effects. Only statistically significant interactions are shown in this table

		F1		**F2**	
	df	χ^2	*p*-value	χ^2	*p*-value
vowel	1	81.6	<0.001	48.0	<0.001
context	2	116.9	<0.001	20.4	<0.001
gender	1	30.0	0.53	0.3	0.53
vowel × context	2	27.3	<0.001	25.2	<0.001
vowel × gender	1	33.0	<0.001	15.1	<0.001
Conditional R^2 = 0.54 (F1), 0.62 (F2); Marginal R^2 = 0.43 (F1), 0.28 (F2)					

Table 21. Results of an ANOVA conducted on an LME model with C1 (F1 & F2) as the dependent variable and vowel, context, gender, tone, and their interactions as fixed effects. Only statistically significant interactions are shown in this table

		F1		**F2**	
	df	χ^2	*p*-value	χ^2	*p*-value
vowel	1	93.3	<0.001	24.5	<0.001
context	2	15.6	<0.001	98.4	<0.001
gender	1	9.6	<0.01	0.4	0.52
vowel × context	2	5.4	0.06	46.2	<0.001
vowel × gender	1	33.2	<0.001	2.9	0.09
Conditional R^2 = 0.39 (F1), 0.62 (F2); Marginal R^2 = 0.29 (F1), 0.28 (F2)					

Table 22. Results of an ANOVA conducted on an LME model with C2 (F1 & F2) as the dependent variable and vowel, context, gender, tone, and their interactions as fixed effect. Only statistically significant interactions are shown in this table

The average duration of the diphthongs for both male and female speakers is plotted in Figure 13 and Figure 14, with corresponding mean values detailed in Table 23 and Table 24. As depicted in the figure, the average duration of the /ie/ diphthong is longer than that of the /uo/ diphthong.

Additionally, for both diphthongs, the duration tends to be relatively longer in female speakers than in male speakers. Further diphthong durations were analyzed based on three contexts and two genders. As illustrated in Figure 14 and detailed in Table 26, diphthongs produced in isolation exhibit longer durations compared to those produced in carrier and sentence frames. Notably, except for the /ie/ diphthongs produced in an isolation context, diphthongs articulated by female speakers tend to have longer durations across all contexts. These findings about the temporal characteristics of diphthong production in the Angami show that context and speaker gender play an important role.

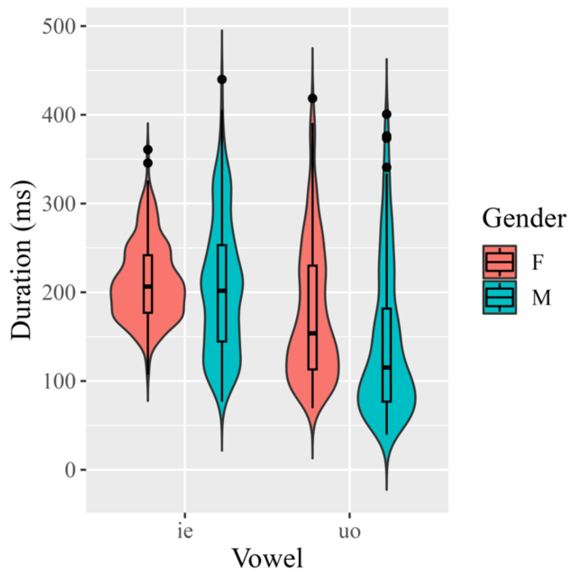

Figure 13. Violin plots of the average diphthong duration across male and female speakers; number of diphthongs = 734

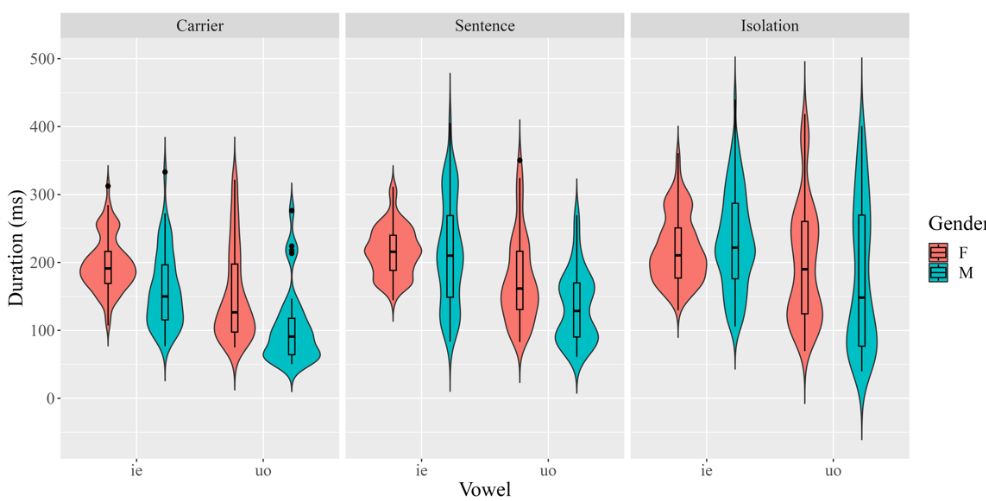

Figure 14. Violin plots of the average diphthong duration by the three contexts and gender; number of diphthongs = 734

Vowel	Mean (SD)
ie	209.4 (59)
uo	163.4 (82)

Table 23. Average duration of the two diphthongs in Angami with their standard deviation

	Female			Male		
	Sentence	Carrier	Isolation	Sentence	Carrier	Isolation
Vowels	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
ie	216.0 (36)	197.5 (40)	219.9 (49)	214.7 (79)	159.2 (55)	231.5 (72)
uo	174.2 (64)	151.3 (68)	199.2 (87)	134.2 (52)	103.5 (53)	(106)

Table 24. Average duration of the two diphthongs with standard deviation (SD) in the three contexts, i.e., sentence, carrier and isolation divided by gender

4 Discussion and Conclusion

In our study, we examined the vowel system of Angami, encompassing both monophthongs and diphthongs. Our analysis confirms the presence of six phonemic monophthongs and two phonemic diphthongs in the language. Cross-linguistically, the most common vowel systems languages exhibit are inventories containing 5 to 6 distinct vowels. This observation is further supported by the data from the UPSID database (317 entries) and subsequent studies on vowels in the World Atlas of Language Structures (WALs) (564 entries). In both these databases, languages most reported 5 to 6 vowels (Maddieson 1984; 2013). This holds true for several Naga languages in particular and for Tibeto-Burman languages in general. For example, in case of other Naga languages, such as, Khezha (Yabu 1994), Mao (Marrison 1967), Lotha (Acharya 1975), Sema (Teo 2014), and Ao (Coupe 2003), it becomes apparent that Naga languages commonly have vowel systems consisting of 6 to 7 vowels. Notably, Yabu (1994) documented that Khezha specifically features a vowel inventory encompassing seven vowels. In case of Angami, contrary to Kuolie (2006) and Suokhrie (2015), this study found only a six-vowel system. The vowels /ü/ and /ə/ as reported in Kuolie (2006) and Suokhrie (2015), did not exhibit any distinctive spectral or temporal characteristics. Hence, in the current study, these two vowels were considered as belonging to the same category. Nevertheless, it was also noticed that the vowel /ə/ can be used as an affix[10] where it acquires substantially longer duration.

Furthermore, in addition to the number of vowels, cross-linguistically, a preference is also noticed for specific vowel qualities. Among the languages that have five vowel systems, there is a preference for the inventory of /i, e, a, o, u/. Similarly for six vowel systems, there is an overwhelming preference for /i, e, a, o, u, ə / (Schwartz et al. 1997; Recasens 2022). In the UPSID database, among 54 languages that have six vowel systems, 26 follow the inventory with schwa in it. Only 12 out of 54 languages prefer having a /ü/ instead of a /ə/ (Maddieson 2013). In case of North-East Indian Tibeto-Burman (TB) languages, it is noticed that the sixth vowel is described with various symbols and descriptions in the literature. Positional restrictions and poor orthographic representation have always marred the identity of the sixth vowel in the TB languages (see Sarmah et.al 2015 and Burling 2013 for more discussion). Even in the current study it was shown that some of the discrepancies in

[10] For example, as a prefix, /ə-tsali/ for 'folk-song'; as a suffix /dzə- ə/ 'water-draw'.

the description of segments in Angami (probably also in several other TB languages of NE India) arise due to the inadequate orthographic conventions adopted for the language.

Abbreviations

CV	consonant vowel
DEF	definite
NA	Not applicable
NE	North East
NOM	nominative
SD	standard deviation
SE	standard error
SG	singular
TB	Tibeto-Burman
V	vowel
VC	vowel consonant

References

Acharya, Katti Panduranga. 1975. *Lotha phonetic reader*. Mysore: Central Institute of Indian Languages.

Adank, Patti; Smits, Roel; and van Hout, Roeland. 2004. "A comparison of vowel normalization procedures for language variation research". *The Journal of the Acoustical Society of America* 116.5: 3099–3107. Acoustical Society of America (ASA). https://doi.org/10.1121/1.1795335

Bates, Douglas; Mächler, Martin; Bolker, Ben; and Walker, Steve. 2014. "Fitting linear mixed-effects models using lme4". *Journal of Statistical Software 67.1: 1-48*. Foundation for Open Access Statistics. https://doi.org/10.18637/jss.v067.i01

Becker-Kristal, Roy. 2010. *Acoustic typology of vowel inventories and Dispersion Theory: Insights from a large cross-linguistic corpus*. Ph.D. diss., University of California.

Bhaskararao, Peri; and Ladefoged, Peter. 1991. "Two types of voiceless nasals". *Journal of the International Phonetic Association* 21.2: 80–88. Cambridge University Press (CUP). https://doi.org/10.1017/s0025100300004424

Blankenship, Barbara; Ladefoged, Peter; Bhaskararao, Peri; and Chase, Nichumeno. 1993. "Phonetic structures of Khonoma Angami". *Linguistics of the Tibeto-Burman area* 16. 2: 69-88.

Boersma, Paul; and Weenink, David. 2020. *Praat: doing phonetics by computer* [Computer program] 6:1.

Burling, Robbins. 1960. "Angami Naga phonemics and word list". *Indian Linguistics* 21.

Burling, Robbins. 2003. "The Tibeto-Burman languages of Northeastern India". In: Thurgood, Graham; and Lapolla, Randy J. (eds.), *The Sino-Tibetan Languages*. London: Routledge, 169–192.

Burling, Robbins. 2013. "The 'sixth' vowel in the Boro-Garo languages". In: Hyslop, Gwendolyn; Morey, Stephen; and Post, Mark W. (eds.), *North East Indian Linguistics*. Foundation Books, 271–280. https://doi.org/10.1017/9789382993285.014

Chase, Nichumeno. 1992. *Descriptive analysis of Khwunomia dialect*. Deccan College Post Graduate and Research Institute Pune. Ph.D. diss., Savitribai Phule Pune University.

Coupe, Alexander R. 2003. *A phonetic and phonological description of Ao: A Tibeto-Burman language of Nagaland north-east India*. Pacific Linguistics, The Australian National University.

Fant, Gunnar. 1970. *Acoustic theory of speech production* 2. De Gruyter.

Flemming, Edward; and Johnson, Stephanie. 2007. "Rosa's roses: reduced vowels in American English". *Journal of the International Phonetic Association* 37.1: 83. Cambridge: Cambridge University Press (CUP). https://doi.org/10.1017/s0025100306002817

Flemming, Edward. "The phonetics of schwa vowels". 2009. *Phonological weakness in English* 493: 78–95.

Fox, John; and Weisberg, Sanford. 2019. *An R Companion to Applied Regression*, Third Edition, Thousand Oaks CA: Sage.

Giridhar, Puttushetra Puttuswamy. 1980. *Angami grammar*. Mysore: Central Institute of Indian Languages.

Koopmans-van Beinum, Florien J. 1994. "What's in a Schwa?" *Phonetica* 51.1–3: 68–79. De Gruyter GmbH. https://doi.org/10.1159/000261959

Kuolie, Duovituo. 2006. *Structural description of Tenyidie: a Tibeto-Burman language of Nagaland*. Kohima: Ura Academy, Publication Division.

Ladefoged, Peter. 1996. *Elements of acoustic phonetics*. Chicago: University of Chicago Press.

Lalhminghlui, Wendy; Terhiija, Viyazonuo; and Sarmah, Priyankoo. 2019. "Vowel-Tone Interaction in Two Tibeto-Burman Languages". In: *Interspeech*, ISCA, 3970–3974. https://doi.org/10.21437/interspeech.2019-2808

Lenth, Russell V. 2017. *Emmeans: Estimated Marginal Means, Aka Least-Squares Means*. CRAN: Contributed Packages, The R Foundation. https://doi.org/10.32614/cran.package.emmeans

Lobanov, Boris M. 1971. "Classification of Russian Vowels Spoken by Different Speakers". *The Journal of the Acoustical Society of America* 49.2B: 606–608. Acoustical Society of America (ASA). https://doi.org/10.1121/1.1912396

Maddieson, Ian. 1984. "Vowel". *Patterns of Sounds*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511753459

Maddieson, Ian. 2013. "Vowel Quality Inventories". In: Dryer, Matthew S.; and Haspelmath, Martin (eds.)., *The World Atlas of Language Structures Online* 2020.3 [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7385533

Mahapatra, Bijay Prasad; McConnell, Grant D.; Padmanabha P.; and Verma V. S. 1989. "Book 2: Non-Constitutional Languages". In: Kloss, Heinz; and McConnell, Grant D. (eds.), *The Written Languages of the World: A Survey of the Degree and Modes of Use*, *Registrar General and Census Commissioner, India*. 611–1359.

Marrison, Geoffrey Edward. 1967. *The classification of the Naga languages of north-east India.* School of Oriental and African Studies. Ph.D. diss., University of London.

McCabe, Robert Blair. 1887. *Outline Grammar of the Angāmi Nāgā Language: with a Vocabulary and Illustrative Sentences*. Superintendent of Government Printing.

Meyase, Savio M. 2016. *A Morphological Investigation of the Tonal Phonology of Tenyidie*. Unpublished Ph.D. diss., The English and Foreign Language University, Hyderabad.

Meyase, Savio M. 2021. "Polarity in a four-level tone language: tone features in Tenyidie". *Phonology* 38.1: 123–146. Cambridge: Cambridge University Press (CUP). https://doi.org/10.1017/s0952675721000063

Meyase, Savio M. 2024. "Dissimilation and phonological conspiracy in Tenyidie tone". *Journal of Linguistics*, 1–16. Cambridge: Cambridge University Press (CUP). https://doi.org/10.1017/s0022226724000392

Mezhüvilie, Liezietsu; Kesiezie, Keduosievi; and Liezietsu, Shürhozelie. 2019. *MKS Dieda*, Third Edition. Ura Academy. Kohima: N.V. Press.

Nienu, Vikuosa. 1990. "Body Parts Questionnaire (Angami Naga)". Unpublished manuscript contributed to *the Sino-Tibetan Etymological Dictionary and Thesaurus* (STEDT).

Office of the Registrar General & Census Commissioner of India. 2011. *Statement-1: Abstract of Speakers' Strength of Languages and Mother Tongues*.

Peterson, Gordon E.; and Barney, Harold L. 1952. "Control Methods Used in a Study of the Vowels". *The Journal of the Acoustical Society of America* 24.2: 175–184. Acoustical Society of America (ASA). https://doi.org/10.1121/1.1906875

R Core Team. 2022. *R: A Language and Environment for Statistical Computing, Foundation for Statistical Computing*. Vienna, Austria.

Ravindran, N. 1974. *Angami phonetic reader* 10. Central Institute of Indian Languages.

Recasens, Daniel. 2022. "Acoustic characteristics and placement within vowel space of full schwa in the world's languages: A survey". *Journal of the International Phonetic Association* 52.1: 59–94. Cambridge: Cambridge University Press (CUP). https://doi.org/10.1017/s0025100320000122

Rose, Phil. 1987. "Considerations in the normalisation of the fundamental frequency of linguistic tone". *Speech Communication* 6.4: 343–352. Elsevier BV.. https://doi.org/10.1016/0167-6393(87)90009-4

Sarmah, Priyankoo; Dihingia, Leena; and Choudhury, Dimple. 2015. "An Acoustic Study of Bodo Vowels". *Himalayan Linguistic*s 14.1. https://doi.org/10.5070/h914125288

Schwartz, Jean-Luc; Boë, Louis-Jean; Vallée, Nathalie; and Abry, Christian. 1997. "Major trends in vowel system inventories". *Journal of Phonetics* 25.3: 233–253. Elsevier BV.. https://doi.org/10.1006/jpho.1997.0044

Suokhrie, Kelhouvinuo. 2015. *Internal variation in Angami: A case study of Kohima village*. Unpublished M. Phil. diss., University of Delhi.

Suokhrie, Kelhouvinuo. 2016. "Clans and clanlectal contact". *Asia-Pacific Language Variation* 2.2: 188–214. John Benjamins. https://doi.org/10.1075/aplv.2.2.04suo

Teo, Amos B. 2014. A phonological and phonetic description of Sumi, a Tibeto-Burman language of Nagaland. Asia-Pacific Linguistics.

Terhiija, Viyazonuo; Sarmah, Priyankoo; and Vijaya, Samudra. 2018. "Acoustic Analysis of Vowels in Two Southern Angami Dialects". In: *21st Oriental COCOSDA*, 7-8.

Terhiija, Viyazonuo; Sarmah, Priyankoo; and Vijaya, Samudra. 2019. "Development of Speech Corpus and Automatic Speech Recognition of Angami". In: *22nd Oriental-COCOSDA*, Cebu, Philippines.

Terhiija, Viyazonuo; and Sarmah, Priyankoo. 2020. "Aspiration in voiceless nasals in Angami". *Proceedings of Meetings on Acoustics* 42, ASA, 060008. https://doi.org/10.1121/2.0001403

Terhiija, Viyazonuo; and Sarmah, Priyankoo. 2024(a). "Voice onset time in Angami". *Linguistics of the Tibeto-Burman Area* 47.1: 32–56. John Benjamins. https://doi.org/10.1075/ltba.00016.ter

Terhiija, Viyazonuo; and Sarmah, Priyankoo. 2024(b). "Acoustics of Voiceless Nasals in Angami". *Journal of Asian and African Studies*, Supplement, 4: 129-144. https://doi.org/10.15026/0002000631

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Yabu, Shiro. 1994. "Body Parts Questionnaire (Khözha)". Unpublished manuscript contributed to the *Sino-Tibetan Etymological Dictionary and Thesaurus* (STEDT).

Viyazonuo Terhiija
viyazonuo@alumni.iitg.ac.in

Priyankoo Sarmah
priyankoo@iitg.ac.in