Bachelor's Thesis

The Vowel Space of Suansu

Author
Ådne Skjæveland Jøssing
adne.jossing@student.uni-tuebingen.de

Supervisors
Jessica K. Ivani
jkivani@ucsb.edu
Gerhard Jäger
gerhard.jaeger@uni-tuebingen.de

A thesis submitted in partial fulfilment
of the requirements for the degree of

Bachelor of Arts
in
International Studies in Computational Linguistics

Seminar für Sprachwissenschaft
Eberhard Karls Universität Tübingen

November 2025

Antiplagiatserklärung

Ich erkläre hiermit,

dass ich die vorliegende Arbeit selbständig verfasst habe,

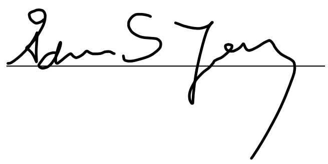
dass ich keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe,

dass ich alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe,

dass die Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist,

dass ich die Arbeit weder vollständig noch in wesentlichen Teilen bereits veröffentlicht habe,

dass das in Dateiform eingereichte Exemplar mit dem eingereichten gebundenen Exemplar übereinstimmt.

Datum: 30.11.2025 **Unterschrift:** 

Abstract

Suansu is an underdescribed Tibeto-Burman language of north-eastern India. This study presents a first instrumental phonetic analysis of its vowel system, using a subset of 186 monosyllabic and sesquisyllabic tokens elicited from a single speaker. Through analyses of this data, the acoustic properties of Suansu vowels are investigated with respect to syllable type, tone, nasality, and complex nuclei. Suansu displays a sesquisyllabic structure, with minor vowels forming a centralized and highly reduced set, whereas major vowels occupy a wider acoustic space, in addition to displaying contrastive tone and nasalization. Tonal analysis identifies a two-way contrast between a level/rising contour and a falling contour. Nasalization is shown to be phonemic, while also occurring as conditioned by nasal codas. The study further documents the distribution of onglides and offglides, though their phonological status remains inconclusive. The identified features align with broader Tibeto-Burman typology and provide a preliminary phonetic baseline for future descriptive work on the vowel space of Suansu.

Contents

List of Figures

List of Tables

1 Introduction

Suansu is a mostly undescribed Tibeto-Burman language spoken in the north-easternmost part of the Ukhrul district in Manipur, north-eastern India. Despite lying in an area of high linguistic diversity, Suansu has received little scholarly attention. The language was first described by Ivani, whose fieldwork constitutes the only existing primary documentation, and all published materials (Ivani, 2023; Ivani, 2019; Ivani, 2022; Ivani, 2024; Ivani and Zakharko, 2024) stem from her work and collaborations. The acoustic space of the vowel system remains mostly unanalysed. This thesis aims to provide the first primarily phonetic account of Suansu vowels based on a set of word recordings collected during Ivani's fieldwork. By describing the vowel system, this study contributes to a broader effort of preserving an endangered language, while also supporting future phonological and typological research.

Since a full phonetic analysis of Suansu's vowel system is still lacking, this initial characterization is valuable: it aids in establishing the phonemic inventory, allows for typological comparison with other underdescribed Tibeto-Burman languages, and provides a baseline for future projects. The scope of this study is necessarily limited, since it is based on a small set of recordings from just a single speaker, but still provides a foundation for future investigations.

The primary goals of this acoustic investigation is to systematically map the phonetic properties of Suansu's vowels, including the range of allophonic variation observed, tothen utilize these acoustic data to explore potential phonemic contrasts, specifically focusing on the base vowel inventory of minor and major syllables, tone, and nasality. More broadly, this work contributes to the documentation of an endangered and underdescribed language, and hopes to provide a basis for future analyses of Suansu, as well as giving direction to future elicitation efforts from Suansu speakers.

In order to address these goals, this study is guided by one primary research question, followed by a set of research objectives specifying how this question will be explored. They are as follows:

1. What are the structural and acoustic properties of the vowel space of Suansu?
    (a) Characterize the distinguishing properties and structure of the vowels in Suansu's minor and major syllables.
    (b) Analyse tonal contours and their place within the vowel system.
    (c) Investigate the distribution, properties, and phonemic status of nasal vowels.

The remainder of the thesis is structured as follows: Section 2 gives an overview of relevant background in phonetics and phonology. Section 3 describes the dataset used. Section 4 presents the background theory and methods used in this study. Section 5 goes over acoustic results, with a discussion of possible interpretations. Section 6 concludes the study.

This study undertakes analyses across several topics relating to Suansu's vowels, a broad and optimistic scope for a project of this size. This scope is however justified, as it would be fruitless to conduct in-depth analyses of any of the covered topics with a dataset of this limited size, as with this no conclusive result would be reachable regardless. Therefore, the general overview offered by this study leaves room for more insight.

2 Background

This chapter gives a necessary geographical, social, and typological context for the acoustic analysis presented in the rest of the study. It begins by outlining the sociolinguistic and geographical setting of the Suansu community, followed by a review of relevant vowel typology in the Tibeto-Burman family, and a summary of previous descriptive work.

2.1 Sociolinguistic Context

Suansu is spoken in a few villages in the north-east of the Ukhrul district: New Tusom (Tusom Khullen) and Tusom CV (Tusom Christian). These lie in the Naga hills about 500 meters above sea level, on the mountain ridge marking the border between India and Myanmar. The Myanmar border lies 4 km to their east, and the border with Nagaland 2 km to their north. Due to the remote location and limited infrastructure, the area is difficult to reach from the rest of the state, including neighbouring villages.

According to the 2011 Indian Census[1] and local accounts from the *Sarpanch* of New Tusom, Suansu is spoken by approximately 2200 people, almost all of whom belong to the Tangkhul Naga Tribe. *Naga* functions as an ethnolinguistic umbrella term for several communities in north-eastern India and north-western Myanmar, and does not represent a linguistic grouping. It is still not certain how Suansu should be classified within the Tibeto-Burman family, due to the complete lack of documentation until recently, further complicated by extensive language contact and historical migration in the region.

[1] http://censusindia.gov.in

Suansu is exclusively used in familial contexts and within the village community. There is no mutual intelligibility with neighbouring languages, and Tangkhul functions as the lingua franca in the district. Most Suansu speakers are also fluent in Meithei, the state language of Manipur. Some additionally speak Nagamese or Hindi, and some in the older generations have knowledge of Burmese. Most younger speakers know English, which is taught in schools. Linguistic pressure from more widely spoken languages is putting the language under threat, further fueled by exogamy and high levels of emigration to more centrally located states.

2.2 Vowel typology in Tibeto-Burman and the region

Languages within the northeast of India are incredibly typologically diverse, even looking only within Tibeto-Burman family. It can therefore be difficult to pinpoint specific linguistic properties shared by a majority of languages. This also holds true for Suansu, which often does not align neatly with other languages of the Ukhrul District. Nevertheless, certain general commonalities within languages of the area can still be observed.

The typical language ranges between having about five and eight vowels in its inventory, almost always containing a mid central vowel, often also a close central vowel. Rounding as a contrastive factor is rare, and contrastive vowel length even more so (Namkung, 1996).

One important feature to mention rather unique to the region and southeast Asia is sesquisyllabicity (Matisoff, 1989, p. 165). A sesquisyllable is a syllable and a half, comprised of one initial minor, and one major syllable. The minor syllable is shorter than a regular syllable, and allows for few or no consonants in coda position, as well almost always having phonetically reduced vowels, carrying no contrastive tone or nasalization.

Nasalization and tone are commonly found features. A large number languages in the region have a set of nasal vowels phonemically distinct from their oral vowels (Watkins, 2001). Tone very frequently occurs contrastively, although the majority of languages feature a minimal inventory of tones, often two, rarely more than three. It is also not uncommon for languages to feature no phonemic tone (Matisoff, 1999).

2.3 Previous work on Suansu vowels

The foundational documentation of Suansu vowels is provided in Ivani (2023), which offers an initial qualitative account based on fieldwork elicitation. This report successfully identifies vowels with distinct features and provides IPA transcriptions. However, as a first descriptive step, the study does not include a systematic phonemic analysis, instrumental acoustic measurements, or an in-depth discussion of allophony, nasalization, or tone. This valuable initial documentation establishes the necessity for an instrumentally grounded and detailed acoustic investigation, and the present study aims to be a first step towards this.

Ivani (2023) finds that Suansu exhibits a sesquisyllabic structure, aligning it with many of its neighbors. This finding is relevant to the present study as it means the language contains minor syllables, which are phonetically reduced, lack stress and tone contrasts, and typically feature a highly reduced vowel. Examples of this structure in Suansu include *əke* 'go up', and *məlaj* 'needle'.

The initial fieldwork suggests the presence of phonemically contrastive nasal vowels. Minimal pair evidence has been documented by Ivani, 2023, such as *ʔaθẽ* 'liver', and *ʔaθe* 'eat'.

Suansu has lexical tone at the word level, meaning pitch differences distinguish word meaning. Ivani, 2023 also demonstrates this contrast with minimal pairs, such as *θʎ́* 'feces' versus *θʎ̀* 'face'. With this the presence of tone is confirmed, but its full extent and interaction with the vowel system remains largely unknown.

2.4 Summary and Relevance to the Present Study

Previous work on Suansu vowels has provided a solid qualitative description, however a more in-depth analysis with instrumental and acoustic evidence is still lacking. This study aims to be a part of that, an initial step in providing acoustic evidence by analysing duration and formant measurements of vowels, in addition to mapping out the environments they occur in, in hopes of providing a solid, instrumental foundation that can serve as a baseline for subsequent comprehensive phonological analyses, as well as future fieldwork.

3 Data

This work is based on data obtained from Ivani, a set of recordings of approximately 900 items collected by her during a linguistic fieldwork trip to India in 2022. This is the same word list used by Coupe, 2020[2]. As this study is necessarily limited in scope, only a restricted number of words from this set could be analysed. The final analytical dataset comprised of 186 tokens selected based on the criteria of being monosyllabic, or sesquisyllabic with an onsetless minor syllable. This focuses resulting analyses, yet also restricts the amount of possibly occurring phonological environments, meaning that this study is not intended as a complete phonetic and phonological analysis of Suansu, but rather a presentation of the phonetic environments existing in tokens encountered in the limited dataset used for the study.

All words were elicited from a single speaker, an adult male native speaker of Suansu, also fluent in Tangkhul, Meithei, and English. He was born and raised in New Tusom, and moved to Maharashtra in his early 20's. All elicitations were recorded in the same environment with the same equipment, in 48 recording sessions spread over three days. The words were elicited from the speaker in the following manner: Ivani says an English word, followed by a brief definition if necessary, after which the speaker repeats the word in Suansu in isolation, then once more within the phrase *a [word] ɹəkè*, 'I write [word]', and lastly again in isolation. The instance of the word which will be analysed is the one uttered within the sentence frame. Using this kind of sentence frame is a standard procedure in elicitation, to prevent intonational cadence from interfering with the tone of the word (Coupe, 2014; Post, 2015).

The initial dataset already had all words of interest delimited from the full elicitation sessions, and were transcribed in IPA[3]. These transcriptions were occasionally revised. Within the words of the dataset 105 instances of minor syllable vowels were identified, and 186 instances of major syllable vowels, of which 115 were part of possible diphthongs or glide-vowel segments. This study is based on findings derived from formant and duration measurements, phonetic transcriptions, and environmental analyses of these vowels.

[2]https://dr.ntu.edu.sg/entities/publication/03e1027b-c79a-474b-bc05-9346ee1aec8f

[3]International phonetic alphabet, transcriptions by Grégoire Hansen

4 Theory and Methods

Methods employed in this study, as well as surrounding theory are presented together in this section, as most methodological decisions require immediate theoretical justification.

Acoustic vowel quality is described using the first two formants (F_1–F_2), where F_1 inversely correlates with vowel height and F_2 inversely correlates with vowel backness (Ladefoged and Johnson, 2015). Instrumental analysis of vowel spaces, through plotting measurements of these formants them in two-dimensional vowel plots, provides an objective method of identifying vowel phonemes, study allophonic variation, and compare systems across languages. Vowel duration, pitch, nasality, and rounding, also relevant to this study, can be acoustically measured as well. Sole interpretation of acoustic measurements may at times be insufficient, therefore subjective auditory perception will also have to be relied on. Additionally, normalization techniques are applied where necessary due to noisy data. Subsequent results then allows for an initial mapping of the Suansu vowel space, and a discussion including preliminary interpretations.

4.1 Segmentation and Acoustic Measurement Procedures

In order to conduct the analyses necessary for this study the already delimited relevant words were extracted from the recordings of full elicitation sessions into Praat[4], along with their IPA transcriptions. These were then checked manually, transcriptions occasionally adjusted, and recordings segmented on a phone-level, aligned with sound boundaries identified through a combination of auditory cues and visual interpretation of their broadband spectrograms in Praat. These segmentations were used to extract duration measurements, while measurements for the first two formants were taken at ten points equally spaced apart, starting 20% into the vowel, ending 80% into the vowel. The first and last 20% of the vowel were excluded from formant measurements in order to capture the stable core of the vowel, avoiding interference with formants from surrounding sounds affecting the vowel, a common practice when tracking formants (Boyle et al., 2024). Praat recognizes formant values automatically. Different settings were experimented with in order to minimize misidentification of formants, and eventually, identifying three formants with a formant ceiling of 3000 Hz was deemed to yield the most accurate results. This was still prone to error, so for any clearly mistracked formants (e.g., F_1 values above 850 Hz) a manual measurement was taken to override the

[4]All acoustic analysis, segmentation, and measurement were performed in Praat (Boersma and Weenink, 2025). Subsequent acoustic measurements were extracted using custom scripts written in Python.

initial automatic measurement. An example of this can be seen in Figure 1.

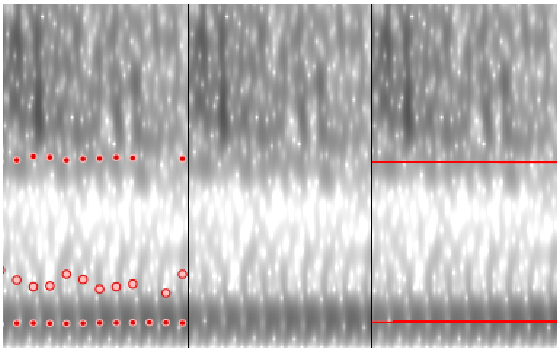


Figure 1: Example of mistracked formants on the left, in the middle no formants are marked, on the right the resulting manual correction is shown.

4.2 Terminological Clarification

To facilitate clear and concise discussion, The terms *minor vowel* and *major vowel* are adapted to refer to the vowel segments found in the respective syllable types. As this represents a fundamental distinction which must be brought up frequently, it is essential to establish clear and concise terminology to refer to these two categories.

The term *complex nucleus* is used to refer to a vowel core that includes adjacent elements such as a preceding semivowel/secondary articulation or diphthong onset, and/or an offglide [j]. The possible combinations can be seen in Table 5.6. These had to be carefully considered, as vowels within these behave differently from vowels not within a complex nucleus.

4.3 Minor-Syllable Vowels

Lobanov normalization is a common technique in acoustic phonetics that uses a modified Z-score transformation to mitigate inter-speaker variation (Lobanov, 1971). The technique calculates a normalized score for each vowel token based on the mean and standard deviation of that speaker's entire vowel space. This process standardizes vowel measurements relative to the individual speaker's acoustic center, which is defined as the zero point (0, 0) in the normalized $F_1 \times F_2$ space. In this study, Lobanov normalization was applied primarily as a measure of vowel centralization. This allowed for the objective demonstration of acoustic reduction in the minor vowels by quantifying their distance from the speaker's normalized mean. Experimentally, the method was also used to determine if it could more broadly reduce possible intra-speaker differences across multiple elicitations. This only affected token distribution very marginally, and would have no impact on final interpretations of the data.

K-means clustering is an unsupervised machine learning algorithm designed to partition a dataset into *k* distinct, non-overlapping clusters where each data point belongs to the cluster with the nearest centroid (MacQueen, 1967). In acoustic analysis, K-means is frequently used to find inherent structuring of a dense acoustic space (Renwick and Ladd, 2016; Bissell, 2021). For this study, the algorithm was applied to the highly centralized minor vowel tokens to empirically determine the most probable number of underlying acoustic categories, assuming any are present, guiding the subsequent phonemic interpretation.

4.4 Tonal Analysis

Tone in this study is analysed using the fundamental frequency (f_0), which reflects the pitch of the voice and is used to characterize tonal contours (Gussenhoven, 2004; Ladd, 2008).

Preprocessing of f_0 measurements included the use of the Median Absolute Deviation (MAD) (Hampel et al., 1986) to remove outliers, often used in f_0 contour analysis (Chien, Borský, and Guðnason, 2019). Gaps in the f_0 contour caused by voiceless segments were filled using linear interpolation (Suni et al., 2013). Cubic interpolation was also experimented with, but it produced unnatural sinusoidal contours, so linear interpolation was preferred.

A preliminary cluster analysis, similar to that performed for minor-syllable vowels, was applied to explore whether the f_0 data would independently recover tonal categories. This analysis suggested two tonal categories, although both were falling with subtle differences in mid-contour shape. This classification was considered unlikely and less reliable than auditory assessment, so manual perceptual classification was prioritized.

Tonal contours were identified by ear for each word and added to the annotations, these were not included in the original transcriptions. A simple binary system was assumed, based on perceived relative pitch changes between the start, mid, and endpoints of words. An initial ternary classification was attempted but later abandoned due to insufficient evidence for systematic distinction. Tonal contours are represented with diacritics: /á/ for Tone 1 (level/rising) and /à/ for Tone 2 (falling).

Perceptual classification by a non-speaker must be approached with caution, as non-distinctive intonational variation or other speech properties can be misinterpreted as tonal differences (Schwanhäußer and Burnham, 2005). Nevertheless, in the context of this study, perceptual classification was the only feasible approach to identifying phonemic tonal contrasts. Previous research shows that non-native listeners can identify tone contrasts above chance (Manyah, 2019), suggesting that careful perceptual classification can yield informative, though not conclusive, results.

4.5 Identifying Nasal Vowels

The primary cues for the presence of nasal vowels in a spectrogram are: a slightly raised F_1, the presence of nasal formants, and a wider bandwidth for oral F_1. (Johnson, 2011) Though all should be considered when identifying nasal vowels, the latter is often the clearest visual sign of nasality in a spectrogram. This increased breadth of the band surrounding the first formant, the darkened area towards the bottom of the spectrogram, can be seen compared with that of an oral vowel in Figure 2. For various reasons these cues may however not always be identifiable, meaning auditory perception remains an instrumental tool.

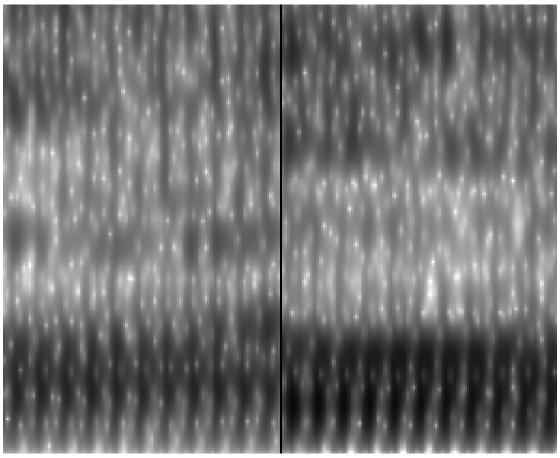


Figure 2: A comparison between oral [ʌ] (left), and nasal [ʌ̃] (right) in a broadband spectrogram. Here an increased bandwidth F_1, the most easily identifiable sign of nasalization, is clearly visible.

4.6 Monophthongal Vowel Analysis

Of words within the data 43 contain onglides which together with the following vowel could possibly constitute *complex nuclei*, displayed in Table 5.6. Aiming for a more concentrated analysis, all such words were excluded from the plots drawn of nasal and oral vowels in Sections 5.4 and 5.5, as the formant values in these tended to deviate considerably from the ellipses drawn to outline the area these occupied in the vowel space.

Vowels with an offglide [j] were however still included in these plots. Additionally, [ɔ̃] only occurred preceding [j] in this dataset, meaning that no instances of [ɔ̃] would have been included in the plots if vowel tokens before [j] had been excluded.

To show the area oral and nasal vowels occupied in the vowel space, ellipses were drawn around the F_1–F_2 mid-point clouds for each vowel category. These ellipses are computed from the sample mean and covariance matrix of F_1 and F_2 for the tokens of each vowel, using an eigen-decomposition of the covariance matrix and scaling the resulting axes by the critical value from the χ^2 distribution for two dimensions, thus representing a 95% confidence region under a bivariate-normal assumption (Harrington, 2018). Before ellipse calculation, tokens with F_1 or F_2 values exceeding a predefined Z-score threshold of ±2 were excluded, reducing the influence of outliers or tracking errors. The resulting ellipses thus visually summarize both the central tendency and the dispersion of each vowel's formant distribution. This is an established method used in other acoustic vowel-space analyses to represent the vowel space of a language (Wang et al., 2013).

4.7 Complex Nuclei

This section concerns sequences involving potential glides, diphthongs, or vowel–vowel combinations. The analysis relies on standard segment-based acoustic methods. Each token was segmented, and the duration of each extracted. A duration ratio between the non-nuclear segment and the vowel nucleus was used as a heuristic for distinguishing glide-like articulations from vocalic elements. Ratios below 1.0 were interpreted as vocalic, ratios at or above 1.0 as semivocalic. This procedure follows general duration-based approaches for distinguishing glides from vowels in phonetic segmentation.

Formant values (F_1–F_2) were also measured at the midpoint of each segment to compare suspected glides to the vowel space. Overlap in formant space was taken as potential evidence of a close relationship to vowels.

4.8 Computational Tools

In order to further analyse and visualize the data, a series of scripts needed to be written.[5] Using Praat's own scripting language a common solution, and the most direct when already working with Praat itself. However, due to a higher level of familiarity with the Python programming language, and some of the scripts needed going beyond the scope

[5]All Python analysis scripts transcriptions, and measurements used for this study are available at https://github.com/adnesj/suansu

covered by Praat scripts, Python was instead used, which with the Parselmouth library[6] allows for the same functionality as scripts written directly in Praat, in addition to the extended flexibility of Python.

5 Results and Discussion

This chapter presents the results of the acoustic analysis of Suansu vowels described in the previous sections. It focuses on descriptive findings, including vowel duration, the distinction between minor and major syllables, the distribution of vowel qualities, and observations of nasality, tone, and complex nuclei. For clarity and consistency across the acoustic plots, all IPA annotations and vowel groupings presented in this chapter represent the author's revised, acoustically-informed interpretation.

A staged approach is adopted, beginning with *Duration and Syllable Type* to objectively establish the fundamental distinction between major and minor syllables. This structural segmentation allows for a targeted analysis of the *Acoustic Space of Minor Vowels* before shifting focus to the major syllable. Within the major syllable, *Tonal Contours* are addressed first, as tone is a feature independent of specific segmental qualities. Subsequently, *Nasal Vowels* are examined to resolve issues of environmentally conditioned nasalization before getting to *Oral Vowels*. Finally, the remaining *Complex Nuclei* are analysed. These steps progressively refine the dataset, allowing for a coherent characterization of the vowel space of Suansu.

The presentation of acoustic results are integrated with their immediate discussion and preliminary interpretation. The first paragraphs of each section presents key descriptive findings from the acoustic analysis. Subsequent paragraphs then analyse and discuss these results, including their implications, their alignment with or divergence from previous research, and how they might justify the approaches taken in following analytical sections.

5.1 Duration and Syllable Type

The first analysis conducted was measuring the duration of vowels in Suansu. Their average duration is shown in Figure 3. These measurements are based on vowels found in both minor and major syllables. As these are expected to behave differently and to display significantly different features, it is important to make an initial clear distinction

[6]The analysis relied on the open-source Python library Parselmouth, which provides an interface to Praat's core C/C++ routines for acoustic analysis in Python.

between these two categories of vowels. One straight-forward way of obtaining an objective measure for this is by classifying them by duration. Such a measure also serves the purpose of further proving the existence of sesquisyllabic structures in Suansu.

As can be seen in Figure 3, there exists a significant contrast in duration between vowels in minor (marked (min.)) and major syllables. On average, major vowels are 37% longer than minor vowels, with means of 140 and 60 ms, respectively. The notably shorter duration of minor vowels, along with the prevalence of sesquisyllabicity in languages of the region, further solidifies the claim of their existence made in Ivani, 2023. Between vowel tokens of the same quality no significant durational difference can be found, and no further evidence for a phonemic vowel length distinction is present in Suansu.

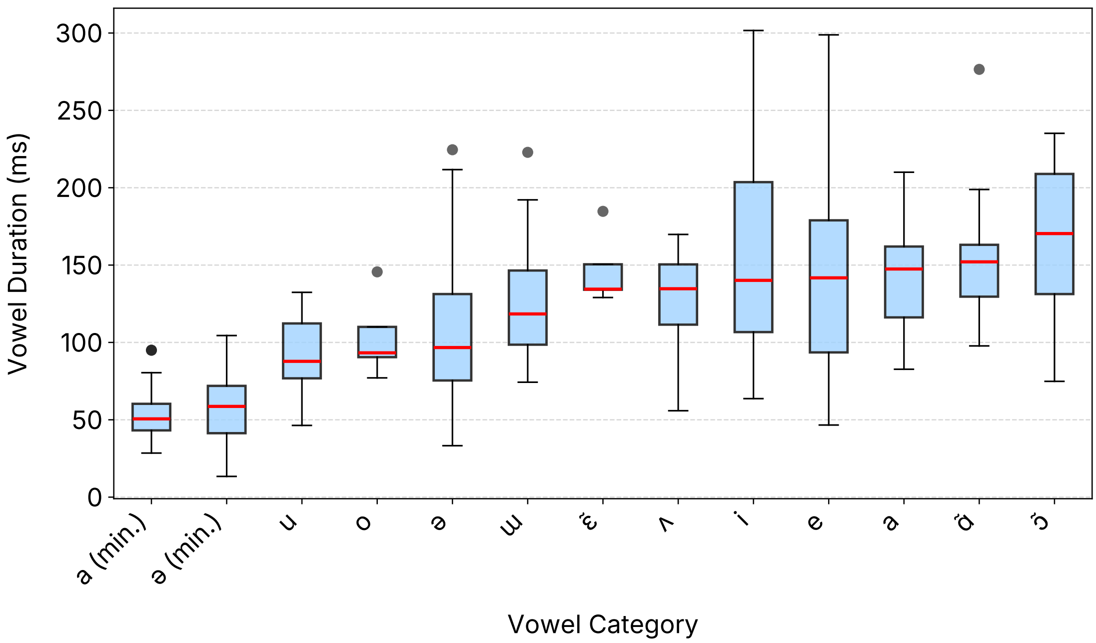


Figure 3: Duration in milliseconds of minor (min.) and major vowels in Suansu, sorted by median duration, and visualized through box plots. Outliers are plotted individually, and the median is indicated by the red line.

5.2 Acoustic Space of Minor Vowels

Due to the short duration and low amplitude characteristic of these phones, the acoustic data for minor vowels exhibited considerable noise, which limits the possible range of interpretation. Average formant values from vowels in minor syllables were collected and plotted in the $F_1 \times F_2$ space (Figure 4).

The plot shows the tokens forming a tight cluster in the center of the acoustic space, indicating a significantly reduced vowel inventory compared to the wide dispersion of major vowels. This centralization can be quantified by normalized formant values: minor syllable tokens exhibit a significantly higher average normalized F_1 (mean: +0.15) and a less extreme average normalized F_2 (mean: -0.05), clustering closely around the speaker's overall normalized origin (0, 0).

One significant outlier was observed: the vowel in the minor syllable of *eɹe*, 'agree'. With an average F_2 of 2168 Hz and a duration comparable to major vowels, it aligns acoustically with major [e]. Due to this, it was excluded. Other less notable outliers caused by measurement errors were also excluded in the following cluster analysis. To investigate potential distinctions within the remaining dense cluster, a K-means cluster analysis was performed, which suggested an optimal partitioning into three categories (Figure 5).

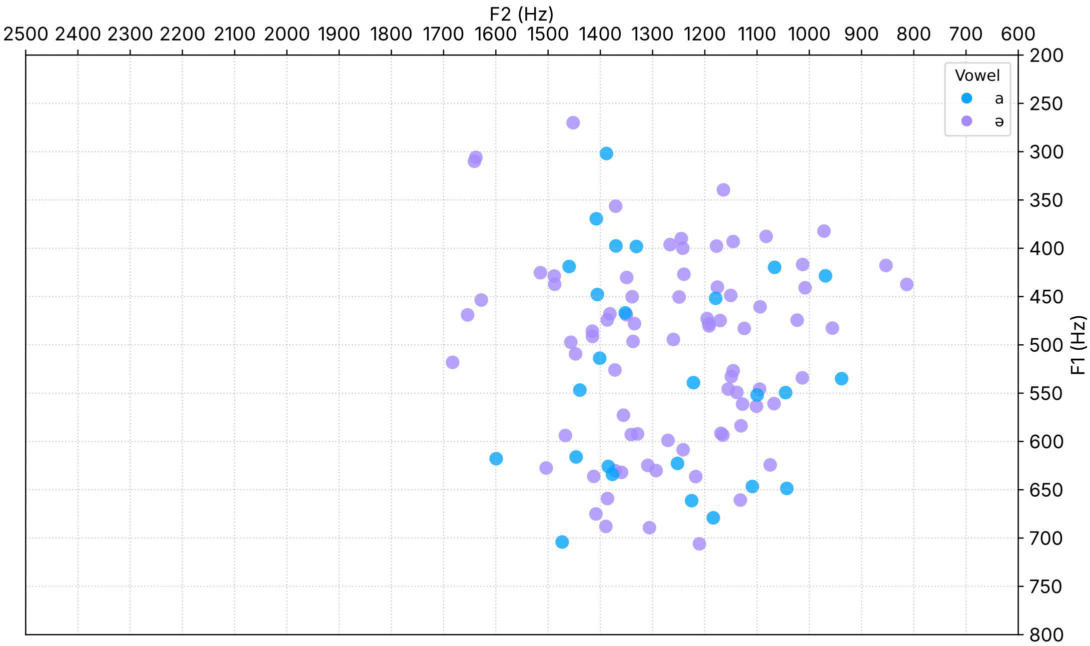


Figure 4: Minor vowels plotted in the $F_1 \times F_2$ space, with original transcriptions retained.

Given the dense clustering and the acoustic overlap visible in the K-means results, determining the phonemic inventory requires interpreting these patterns in light of coarticulatory effects. While the cluster analysis predicts a ternary distinction (high-front, low-front, and back), the *back* cluster appears to be largely conditioned by phonetic environment. As shown in Figure 6, minor vowels followed by labial consonants are fairly consistently retracted in the vowel space. This correlation could suggests that the observed backing is a coarticulatory effect rather than the realization of a distinct back vowel phoneme in minor syllables. Having the opposite effect, minor vowels followed by palatal and palato-alveolar sounds were observed further in the front of the cluster than the average.

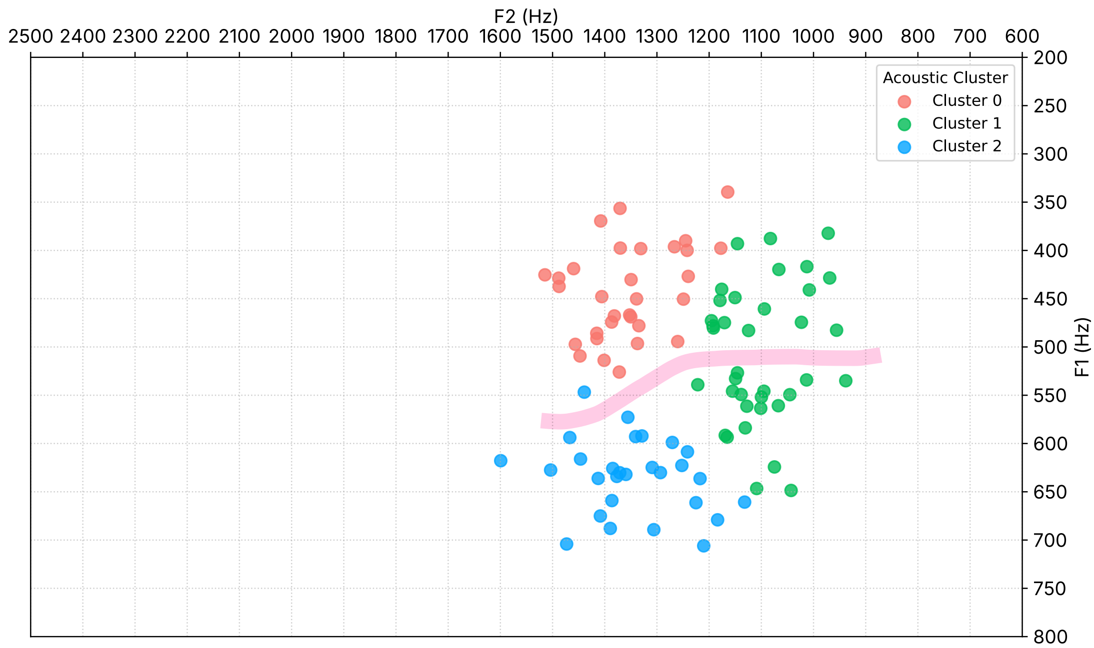


Figure 5: Minor syllables grouped by colour from a K-means cluster analysis showing a potential three-category distinction. The pink line shows a potential two-category distinction.

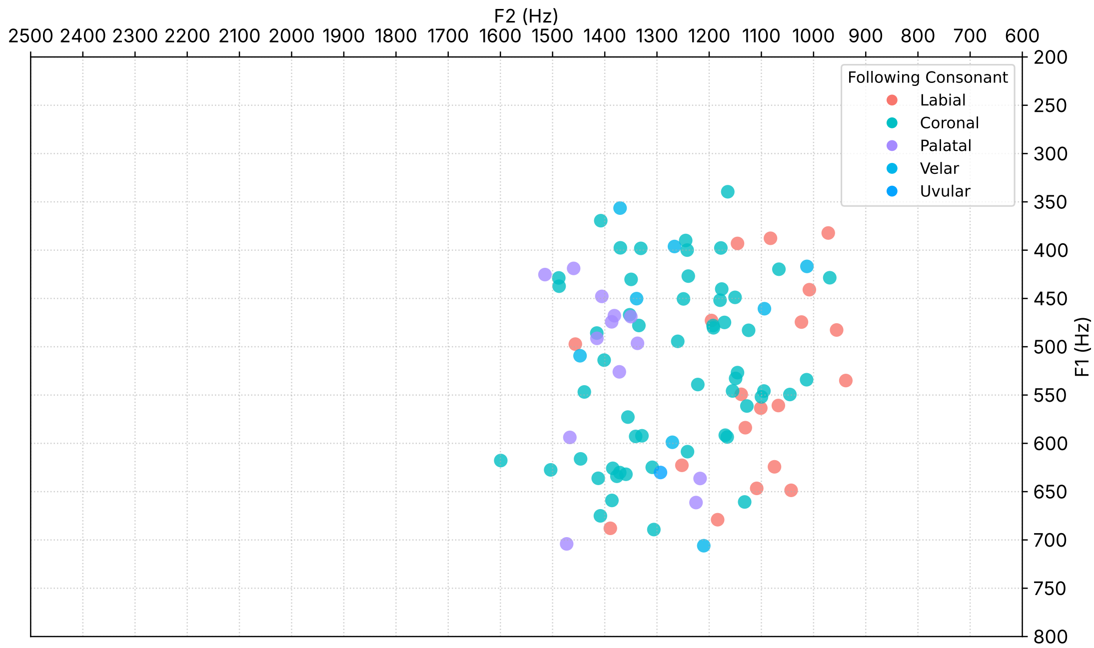


Figure 6: Minor vowels plotted by following consonant group, showing potential retraction triggered by labial contexts, and fronting triggered by palatal contexts.

Consequently, a binary distinction based primarily on vowel height represents a more parsimonious interpretation of the system. This leaves a contrast between a non-low central vowel /ə/ and a low central vowel /a/, shown by the pink line in Figure 5. This classification assumes the slight vertical separation observed is more significant than the noisy data makes it seem, while treating horizontal variation as arbitrary or allophonic. However, the tight clustering also makes interpreting a single, undifferentiated phoneme a likely outcome.

As minor syllables do not appear to carry any of the further contrasts investigated in this study, the following sections focus on phonemic features tied to the major syllable.

5.3 Tonal Contours

Auditory assessment of the data suggested a contrast between two distinct pitch contours. Based on the established sesquisyllabic structure, tone is assumed to be carried exclusively by the major syllable, so that sesquisyllabic words possess the same tonal distinctions as monosyllabic ones. Preprocessing was applied to the f_0 data following the methods described in Chapter 4.

Figure 7 visualizes the mean f_0 contour for the two primary tonal categories in monosyllabic words, and Figure 8 does the same for sesquisyllabic words. Both contours exhibit a falling or level trajectory in the first half of the word, diverging in the final half: Tone 1 remains relatively level or shows a slight terminal rise, while Tone 2 continues to fall throughout. A third tonal category was initially considered but discarded due to insufficient token count and acoustic similarity to Tone 2. For visualization, mean f_0 contours are plotted with ±1 standard deviation shading to illustrate the variability across individual tokens (Xu, 2001).

The pitch range for the high versus low end of the tonal contours for the Suansu speaker was approximately 4.26 semitones ($st = 12 \times \log_2(f_{high}/f_{low})$), a value which provides a quantitative characterization of the observed contrast.

Minimal pairs are predicted within the data, such as *sʌ́*, 'deer' versus *sʌ̀*, 'come out', and *t͡θʌ́*, 'excrement' versus *t͡θʌ̀*, 'face', the latter also documented in Ivani, 2023.

Looking at tone distribution across lexical categories shows that verbs favor the falling tone (61 out of 78 tokens), while nouns show a slight preference against the falling tone (30 out of 75). If the tonal contours identified are representative, this pattern suggests a possible link between tone and morphological or lexical category, as observed in other Tibeto-Burman languages (Teo, 2014; Lotven, 2023).

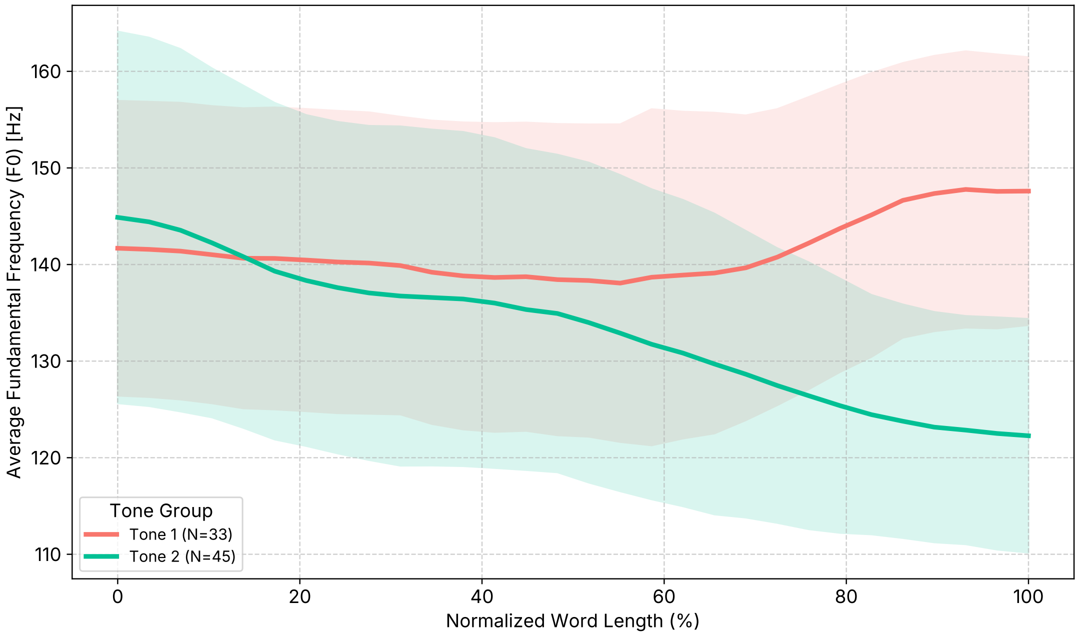


Figure 7: Averages in monosyllabic words of the two tonal contours identified by ear, with word length normalized.

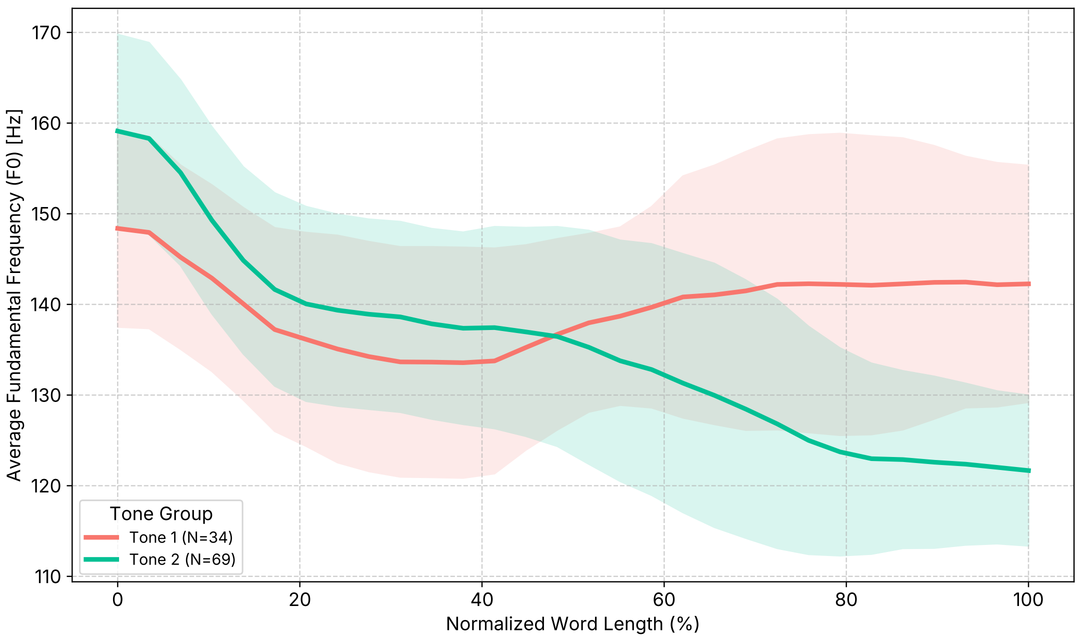


Figure 8: Averages in sesquisyllabic words of the two tonal contours identified by ear, with word length normalized.

5.4 Nasal Vowels

The distribution and phonemic status of nasality needs to be investigated, as it is commonly a phonemic feature in Tibeto-Burman languages. This is handled before oral vowels, in order to clarify which vowels with surface nasalization are environmentally conditioned, and can therefore be considered allophones of oral vowels.

Nasal vowels occur in nine open syllables, in 15 syllables closed by a and coda [j], and in twelve syllables closed by a nasal coda [m] or [ŋ]. Analysing the acoustic profile of the nasalisation in these two contexts suggests a functional difference, where the former is phonemic, while the latter is coarticulatory. Tokens found in open syllables show the strongest acoustic signature of nasality, such as a consistent and significantly widened F_1 bandwidth as shown in Figure 2. The vowels [ə] and [u] show marked nasalization when followed by a nasal consonant, becoming [ə̃] and [ũ], respectively, occurring in the environments [ə̃m], [ə̃ŋ], and [ũŋ]. Since the nasal variants of these have formants in comparable environments to their oral equivalents, and never occur without a nasal coda, they will be treated as allophones, and are therefore plotted with oral vowels in Figure 10 instead of in this chapter. This leaves us with the nasal vowels [ʌ̃, ɔ̃, ɛ̃] as the only tokens that can be safely identified as having distinctive nasalization. Comparing these with oral vowels in Figure 10, these nasal vowels cannot be trivially derived as nasal variants of their oral counterparts, which is typologically unusual (Maddieson, 2007).

The nasal tokens [ɔ̃] and [ɛ̃] occupy areas in the vowel space unique to those occupied by oral vowels, however, nasal [ʌ̃] overlaps extensively with the oral [ʌ], making the $F_1 \times F_2$ space alone insufficient to show a phonemic distinction between the two. This leaves us with minimal pairs such as *t͡θʌ̀*, 'face', and *t͡θʌ̃̀*, 'come in'. Other minimal pairs were also previously noted by Ivani, 2023. Considering this, in addition to having identified the primary acoustic cues for nasal vowels, nasality can be reasonably interpreted as a contrastive feature in the major syllable vowels of Suansu.

Looking at Table 5.6, the lack of a coda [n] is apparent, given the presence of both [m] and [ŋ]. See Chen and Guion-Anderson, 2011 and Johnson, 2011, pp. 202–204 for an experiment-backed discussion of mergers of nasal codas based on their relative confusability. A possible explanation for both this gap in permissible nasal codas and the overall presence of phonemic nasal vowels could be that historically coda [n] was allowed, nasalized preceding vowels, then disappeared, leaving behind coda-less nasal vowels: **Vn* > *Ṽ*.

None of the three possible nasal phonemes identified interact with the top of the vowel

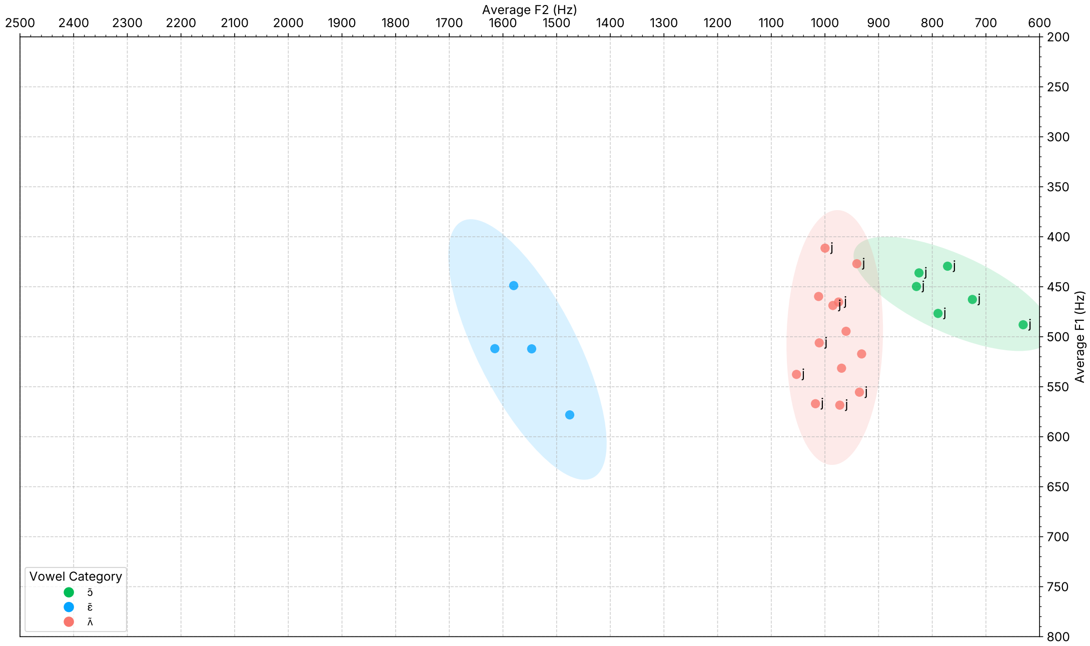


Figure 9: Nasal vowels. Nasal vowels with coda [m] or [ŋ] are assumed non-phonemically nasal and excluded from the plot. The following segment, if any, is displayed to the right of each point.

space, and we also have notably fewer nasal than oral vowel phonemes. This is not unusual, it is a cross-linguistic tendency, as high nasal vowels tend to be the most unstable, and often either lower, or denasalize, and the acoustic effect of nasalization is known to affect the perceptual distinctiveness of vowels (Beddor, 1993).

Overall, Suansu's nasal vowels form a small, distinct subset of the vowel inventory, confined to mid positions, and occupying regions that do not directly correspond to oral counterparts. [ə̃] and [ũ] are identified as likely allophones of oral vowels conditioned by a nasal coda, which sets the stage for understanding the broader system of oral vowels.

5.5 Oral Vowels

The remaining vowels were plotted in the $F_1 \times F_2$ space, including those conditionally nasalized by nasal codas, as described in Section 5.4. The tokens containing onglides or vowel-vowel sequences were excluded from this plot, see Section 5.6 for a discussion on those. Ellipses represent the area in the vowel space each identified token type occupies.

The overall vowel space has a triangular structure, with a large gap left open between [e] and [a]. (The nasal ɛ̃ however occurs within this gap.) More generally, a clear gap between front vowels [i] and [e], and the other non-front vowels can be observed. In the back and center of the vowel space there is considerable crowding, with some overlap between ellipses. This follows what is often typical of Tibeto-Burman languages, which tend to have more crowding in the rear end of their vowel spaces.

Figure 10: Major vowel phonemes with individual points for each vowel, and ellipses marking the approximate vowel space of each cluster. The following segment, if any, is displayed to the right of each point.

One notable observation from figure 10 is that the vowels [ɯ, i, e] are only found in fully open syllables, and the only additional environment [a, ʌ] allow is with an offglide [j]. Importantly, it should be noted that [o] is only observed followed by [q]. This makes it a likely candidate for allophony, possibly as an allophone of [u] in syllables with uvular codas, a phenomenon also observed in other languages (Hagerup, 2011). The vowel [u] can also be seen to be split in two smaller clusters, when followed by [ŋ] it is further backed and opened than in open syllables. Given that [u] is considerably nasalized when followed by [ŋ], this observation follows how we would expect a vowel's formants to behave when nasalized.

The environments that most vowels occur in are quite restrictive. The vowel displaying by far the most variation in its following environment is [ə]. It is uncertain how this should be interpreted, but it could be linked to neutralization of possible vowel contrasts in coda environments displayed by [ə], which other vowels seem to lack. An additional observation regarding [ə] is that when followed by [j], it is noticeably fronted compared to its other occurrences.

F_1 dispersion was the largest for [a], which is consistent with its lowered quality. [ɯ] forms a fairly tight cluster, except two outliers. Formant measurements for this vowel were the most troublesome, as properties of preceding consonants bled considerably into the vowel. This was especially noticeable with fricatives, where most parts of the following [ɯ] would be heavily fricated. Due to the low occurrence of [ɯ] in the dataset no further interpretation of this phenomenon could be made, but it is something which should be investigated further in future work on Suansu's vowels.

The overall patterning of oral vowels thus suggests an inventory of /a, ʌ, ə, u, ɯ, i, e/, not atypical of a Tibeto-Burman language. Vowels seem strictly conditioned by their segmental environment: front vowels remain confined to open syllables, [ə] shows the widest positional range, allowing all occurring codas. The realization of [u] is heavily conditioned by the presence of a nasal or uvular coda. Non-high, non-front vowels also allow a transition into an offglide [j], and the exact way in which this is best treated will be discussed in the next section.

5.6 Complex Nuclei

In addition to the monophthongs treated in the last two sections, exactly half of the words in the dataset contain *complex nuclei*, consisting of a vowel that includes adjacent elements such as a preceding semivowel/secondary articulation or diphthong onset, and/or an offglide. These were identified in the following categories:

1. Prevocalic
   (a) Vowel: [i, u, ɯ] (example: *a**niè*** 'bake')
   (b) Semivowel: [j, w, ɰ] (example: *an**jə́*** 'be fresh')
2. Postvocalic semivowel: [j]. (example: *at**àj*** 'hold')

Segments preceding the vowel nucleus were quantified by measuring their duration. It was found that those with a duration shorter than the following vowel could be considered semivocalic, and those with a duration longer than the following vowel could be considered vocalic. This adheres with findings from previous research (Crystal, 2008). The exception is intervocalic semivowels, which were longer in duration than the following vowel. (Example: *ají* 'hold')

Table 5.6 shows the observed distribution of semivocalic and vocalic segments cooccurring with regular vowels. Note that combinations consisting of both an onglide and a coda were only observed in *əjàj*, 'shine', *apiàj*, 'diverge', and *əmwàj*, 'itch'.

Onglide						Nucleus	Coda				
w	ɰ	j	u	ɯ	i		∅	j	q	m	ŋ
X		X			X	a	X	X			
X						ʌ	X	X			
X						ə	X	X	X	X	X
						o			X		
						u	X				X
X	X					i	X				
X	X	X	X		X	e	X				
						ɑ̃	X	X			
						ɔ̃		X			
X						ɛ̃	X				

Table 1: Distribution of onglides and nuclei across coda environments.

Measurements of F_1 and F_2 were taken of tokens for onglides [j, w, ɰ], and when plotted, these tokens overlap respectively with the elliptical vowel space drawn for the vowels [i, u, ɯ]. Though it is sometimes argued that semivowels are always non-syllabic realizations of underlying vowels (Kaye and Lowenstamm, 1984) as this is a common occurrence cross-linguistically (Rosenthall and Horn, 2013), it is not always the case (Levi, 2008).

One of these glides, [j], is additionally observed in coda position. Given the cross-linguistic manner hierarchy that tends to be preferred for coda consonants (nasal > liquid > obstruent > glide) (VanDam, 2004), vowel segments with an offglide [j] would be best assumed to be diphthongs. This analysis is further supported by the observed distribution of nasal vowels: nasal vowels can occur with an offglide [j], and if these were analysed as V+j sequences, the historical development _*Vn_ > _Ṽ_ would require positing an underlying coda sequence _*VGn_. Suansu otherwise permits only single-segment codas, and no other independent evidence motivates multisegment codas in the language. Treating _Vj_ as diphthongs therefore preserves the generalization that Suansu syllables allow only one coda segment, while allowing nasal vowels to be derived straightforwardly from earlier Vn sequences in line with the rest of the dataset.

The purpose of including a section such as this was as a contextual analysis to broaden the perspective from the rest of this work, and explore whether the tokens analysed could be considered part of the vowel space. As the data used in this analysis is too sparse, and the acoustic analysis conducted too shallow, no conclusive interpretation can be made, and further discussion on the topic is left for later work on the phonetics of Suansu.

6 Conclusion

This study set out to investigate the structural and acoustic properties of the vowel space of Suansu, focusing on minor versus major syllable types and their properties, tone, and nasality. Across these domains, the analyses reveal a vowel space that is largely consistent with typological expectations for Tibeto-Burman languages of north-eastern India.

Minor syllable vowels show reduced, centralized qualities, while major syllable vowels display a broader range of contrastive features, including tone and nasality. Tone seems to be contrastive, in a simple system, employing a binary distinction. Nasal vowels occur in more restricted positions with reduced phonological contrasts compared to oral vowels. Lastly, half of all major syllables contain complex nuclei, consisting of vowel–vowel combinations or combinations of vowels and semivowels, rather than simple monophthongs. Their distribution was briefly explored, but a proper classification must be left for later, more comprehensive studies.

Together, these findings give an overarching picture of Suansu's vowel space, along with the distribution of surrounding segments. The interpretations collectively approach an answer to the all aspects of the research question, showing that Suansu's vowel space contains a combination of features typical of Tibeto-Burman languages of north-east India.

6.1 Limitations

This study is primarily limited by its small dataset of 186 mono- and sesquisyllabic tokens, as well as by the restricted time and resources available for analysis. This leaves room for erroneous interpretation, as it is unlikely that all possible vowel environments were present in the data, biasing interpretation towards environments which were present. Entire phonemes could easily be missed in such an analysis, as too few occurrences would likely lead to them being wrongly categorized. It is also a possibility that entire phonemes by chance simply do not occur in the data analysed. Due to this, the data discussed in this study should be considered preliminary. All aspects of phonology discussed, especially the phonemic status of diphthongs, require larger datasets and deeper investigation.

6.2 Future Directions

This initial phonetic analysis provides a methodological baseline for the study of Suansu vowels. Its interpretations, along with challenges encountered, define the immediate objectives for future fieldwork efforts on this underdescribed Tibeto-Burman language, and paves the way for subsequent investigations into the vowel space of Suansu.

7 Acknowledgments

This thesis owes a great deal to my supervisor, Jessica Ivani. I am deeply grateful for her commitment to data collection in the field, which yielded the foundational Suansu fieldwork data used in this study, and for giving me the opportunity to analyse this material for my thesis, all of which has ultimately inspired me to further pursue field linguistics. I am also grateful to Nikita Beklemishev for sharing materials that supported this research, and for our insightful discussions on the topics covered in my thesis.

References

Beddor, Patrice Speeter (1993). "The perception of nasal vowels". In: *Nasals, nasalization, and the velum*. Elsevier, pp. 171–196.

Bissell, Marie (2021). "Automatic phonetic classification of vocalic allophones in Tol". In: *Proceedings of the Linguistic Society of America* 6.1, pp. 403–410.

Boersma, Paul and David Weenink (2025). *Praat: doing phonetics by computer*. Version 6.4.27. Retrieved from `http://www.praat.org/`. University of Amsterdam.

Boyle, John P, Jiaang Dong, Armik Mirzayan, and VB Scott (2024). "An Acoustic–Phonetic Description of Hidatsa Vowels". In: *Languages* 9.10, p. 315.

Chen, Ying and Susan Guion-Anderson (2011). "Perceptual Confusabiltiy of Word-final Nasals in Southern Min and Mandarin: Implications for Coda Nasal Mergers in Chinese." In: *ICPhS*, pp. 464–467.

Chien, Yu-Ren, Michal Borský, and Jón Guðnason (2019). "F0 Variability Measures Based on Glottal Closure Instants". In: *Proceedings of Interspeech 2019*. Graz, Austria: ISCA, pp. 1986–1989. DOI: `10.21437/Interspeech.2019-1326`. URL: `https://www.isca-archive.org/interspeech_2019/chien19c_interspeech.pdf`.

Coupe, Alexander (2014). "Problems and Strategies in the Elicitation of Tone in Northeast Indian Languages". In: *Linguistics of the Tibeto-Burman Area* 37.1, pp. 1–25.

Coupe, Alexander R (2020). "Northern Sangtam phonetics, phonology and word list". In: *Linguistics of the Tibeto-Burman Area* 43.1, pp. 147–189.

Crystal, David (2008). *A dictionary of linguistics and phonetics*. John Wiley & Sons, pp. 431–432.

Gussenhoven, Carlos (2004). "The phonology of tone and intonation". In.

Hagerup, Asger (2011). “A phonological analysis of vowel allophony in West Greenlandic”. MA thesis. Norges teknisk-naturvitenskapelige universitet, Det humanistiske fakultet ..., p. 21.

Hampel, Frank R., Elvezio M. Ronchetti, Peter J. Rousseeuw, and Werner A. Stahel (1986). *Robust Statistics: The Approach Based on Influence Functions*. John Wiley & Sons, p. 105.

Hardcastle, William J, John Laver, and Fiona E Gibbon (2012). *The handbook of phonetic sciences*. John Wiley & Sons.

Harrington, Jonathan (2018). *Classification and Discriminant Analysis*. Accessed: 2025-11-29. URL: `https://www.phonetik.uni-muenchen.de/~jmh/lehre/Rdf/EMU-SDMS/lesson10/10_Classification.html`.

Ivani, Jessica K (2019). “A first overview on Suansu, a Tibeto-Burman language from Northeastern India”. In: *Talk, Held at the 29th Conference of the Southeast Asian Linguistic Society*.

Ivani, Jessica K (2022). “Comparative constructions in Suansu and the languages of northeastern India”. In: *Linguistic Typology at the Crossroads* 2.1, pp. 65–93.

Ivani, Jessica K (2023). “Suansu language from northeastern India: A field report”. In: *Linguistics of the Tibeto-Burman Area* 46.1, pp. 138–163.

Ivani, Jessica K and Taras Zakharko (2024). “Phasal polarity in Suansu”. In: *Linguistics of the Tibeto-Burman Area* 47.2, pp. 318–342.

Ivani, Jessica Katiuscia (2024). “Caritive expression in Suansu”. In: *Himalayan Linguistics* 23.3.

Johnson, Keith (2011). *Acoustic and auditory phonetics*. John Wiley & Sons.

Kaye, Jonathan and Jean Lowenstamm (1984). “De la syllabicité”. In: *Forme sonore du langage* 123, p. 59.

Ladd, D Robert (2008). *Intonational phonology*. Cambridge University Press.

Ladefoged, Peter and Keith Johnson (2015). *Vowels and Consonants*. 3rd. Wiley-Blackwell.

Levi, Susannah V (2008). “Phonemic vs. derived glides”. In: *Lingua* 118.12, pp. 1956–1978.

Lobanov, Boris M (1971). “Classification of Russian vowels spoken by different speakers”. In: *The Journal of the Acoustical Society of America* 49.2B, pp. 606–608.

Lotven, Samson (2023). “Tone in South Central Tibeto-Burman”. In: *Himalayan Linguistics* 22.1.

MacQueen, J (1967). “Multivariate observations”. In: *Proceedings ofthe 5th Berkeley Symposium on Mathematical Statisticsand Probability*. Vol. 1, pp. 281–297.

Maddieson, Ian (2007). “Areal distribution of nasalized vowels”. In: *Proceedings of the 16th International congress of phonetic sciences*. Universität des Saarlandes Saarbrücken, pp. 1381–84.

Manyah, Kofi Adu (2019). "NON-NATIVE SPEAKERS'IDENTIFICATION OF LEXICAL TONE CONTRASTS". In.

Matisoff, James A (1999). "Tibeto-Burman tonology in an areal context". In: *Proceedings of the symposium "Crosslinguistic studies of tonal phenomena: Tonogenesis, Japanese Accentology, and Other Topics*. Tokyo: Tokyo University of Foreign Studies, Institute for the Study of ..., pp. 3–31.

Matisoff, James A. (1989). "The bulging monosyllable, or the mora the merrier: echo-vowel adverbialization in Lahu". In: *Southeast Asian Linguistics: Essays in honour of Eugénie J.A. Henderson*. Ed. by Jeremy H.C.S. Davidson. London: School of Oriental & African Studies, University of London, pp. 163–197.

Namkung, Jun (1996). *Phonological inventories of Tibeto-Burman languages*. Sino-Tibetan Etymological Dictionary and Thesaurus Project, Center for ...

Post, Mark W (2015). "Tones in Northeast Indian languages, with a focus on Tani: A field-worker's guide". In: *Language and culture in Northeast India and beyond: In honour of Robbins Burling*, pp. 182–210.

Renwick, Margaret EL and D Robert Ladd (2016). "Phonetic distinctiveness vs. lexical contrastiveness in non-robust phonemic contrasts". In: *Laboratory Phonology* 7.1.

Rosenthall, Samuel and Laurence Horn (2013). *Vowel/glide alternation in a theory of constraint interaction*. Routledge.

Schwanhäußer, Barbara and Denis Burnham (2005). "Lexical tone and pitch perception in tone and non-tone language speakers." In: *INTERSPEECH*, pp. 1701–1704.

Suni, Antti, Daniel Aalto, Tuomo Raitio, Paavo Alku, and Martti Vainio (2013). "Wavelets for intonation modeling in HMM speech synthesis". In: *Proc. 8th ISCA Speech Synthesis Workshop, 2013*. ISCA-INST SPEECH COMMUNICATION ASSOC.

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

Thurgood, Graham and Randy J LaPolla (2016). *The sino-tibetan languages*. Routledge.

VanDam, Mark (2004). "Word final coda typology". In: *Journal of Universal Language* 5.1, pp. 119–148.

Wang, Yuguang et al. (2013). "An MRI-based acoustic study of Mandarin vowels." In: *INTERSPEECH*, pp. 568–571.

Watkins, Justin W (2001). "Burmese". In: *Journal of the International Phonetic Association* 31.2, pp. 291–295.

Xu, Yi (2001). "Fundamental frequency peak delay in Mandarin". In: *Phonetica* 58.1-2, pp. 26–52.