Fall 2023

The Sound of Dzongkha: *Recent Changes to its Tonal Paradigm*

Qiyou (Ark) Lu , '24

The Sound of Dzongkha

Recent Changes to its Tonal Paradigm

By
Ark Lu Qiyou
陸其右

A Thesis
Submitted to the Department of Linguistics
Of Swarthmore College
In Partial Fulfilment of the Requirements
For the Degree of Bachelor of Arts
In
December 2023

Abstract

This thesis explores the evolving phonetics and phonology of Dzongkha, a Tibetic language spoken in Bhutan, a small landlocked country in the eastern Himalayas, with a particular focus on the transformation of its tonal system over the past three decades. The country has seen noticeable changes to its official language in recent decades due to language policies and migration. This research aims to provide an updated analysis of the language's sound system based on newly collected field data.

The study covers various aspects of Dzongkha's sound system, including its vowel and consonant inventories, phonotactics and tones. In particular, the language is demonstrated to have a two-way phonation contrast for its onset stop, affricate, fricative series and some of its nasals and approximants. Compared to existing literature, this study shows that the language has been undergoing simplifications to its vowel and consonant inventories, as well as tonogenesis, where the previously redundant voicing contrast for its onset stop and affricate series have completely transphonologised as pitch in the following vowel. In addition, historical aspiration contrast for its voiced stop and affricate onsets have transphonologised as vowel phonation. Overall, the combinations of vowel pitch and phonation in relation to onset consonant aspiration can be categorised as tonal registers. However, at the same time, the language had lost tonal contour contrast, with previously minimal pairs being attested as allophones and the lexicon adjusting to avoid homophonic ambiguity.

By examining the phonological changes in Dzongkha, this thesis contributes to our understanding of language change and adaptation, especially in the context of contact with other languages. Additionally, it sheds light on the broader implications for linguistic research in the Eastern Himalayan region, where a rich diversity of languages coexist.

Acknowledgements

As I am not a speaker of Dzongkha, I am indebted to my language consultants for their continued support and patience over the course of my research. Their ability to quickly understand and use technical terms without prior knowledge of linguistics significantly helped my analysis. I am grateful to Professors Rikker Dockum and Jonathan Washington for their guidance in and out of class, whose expertise in field work, phonetics and phonology gave me the practical and theoretical tools to carry out this study. I am also grateful to Professors Amanda Payne and Jane Chandlee, who offered invaluable advice that made this thesis possible. In addition, I would like to thank my colleagues in Field Methods, Phonetics & Phonology and Thesis Seminar classes who helped me amass the elicitation data and offered suggestions during my writing process.

II. List of Figures

III. List of Tables

IV. List of Abbreviations

A	approximant consonant
C	consonant
CoG	centre of gravity
COP	copula
F	feminine
FRM	formal
F1	formant 1
F2	formant 2
GEN	genitive
Hz	hertz
INAN	inanimate
INFRM	informal
M	masculine
N	nasal consonant
NMZ	nominaliser
SD	standard deviation
SG	singular
TOP	topic
V	vowel
VOT	voice onset time

1. Introduction

Dzongkha (ISO: dzo) is a Tibetic language spoken primarily in the western half of Bhutan, a kingdom in the eastern Himalayas, by about 600,000 people, most of whom are not native speakers. The name of the language comprises two words: /t͡sǒŋ/, a distinctive type of fortified monastery building found in Bhutan, and /kʰá/, meaning “mouth” or “language”.

Tibetic languages, also called Central Bodish languages, contain the descendants of Old and Classical Tibetan. Geographically, the speech community of Dzongkha borders that of other Tibetic languages in the north, such as Central/Lhasa and Khams Tibetan, non-Tibetic Bodish languages in the east, such as Tsangla and Takpa, and Pahari languages, such as Nepali, in the south (Figure 1).

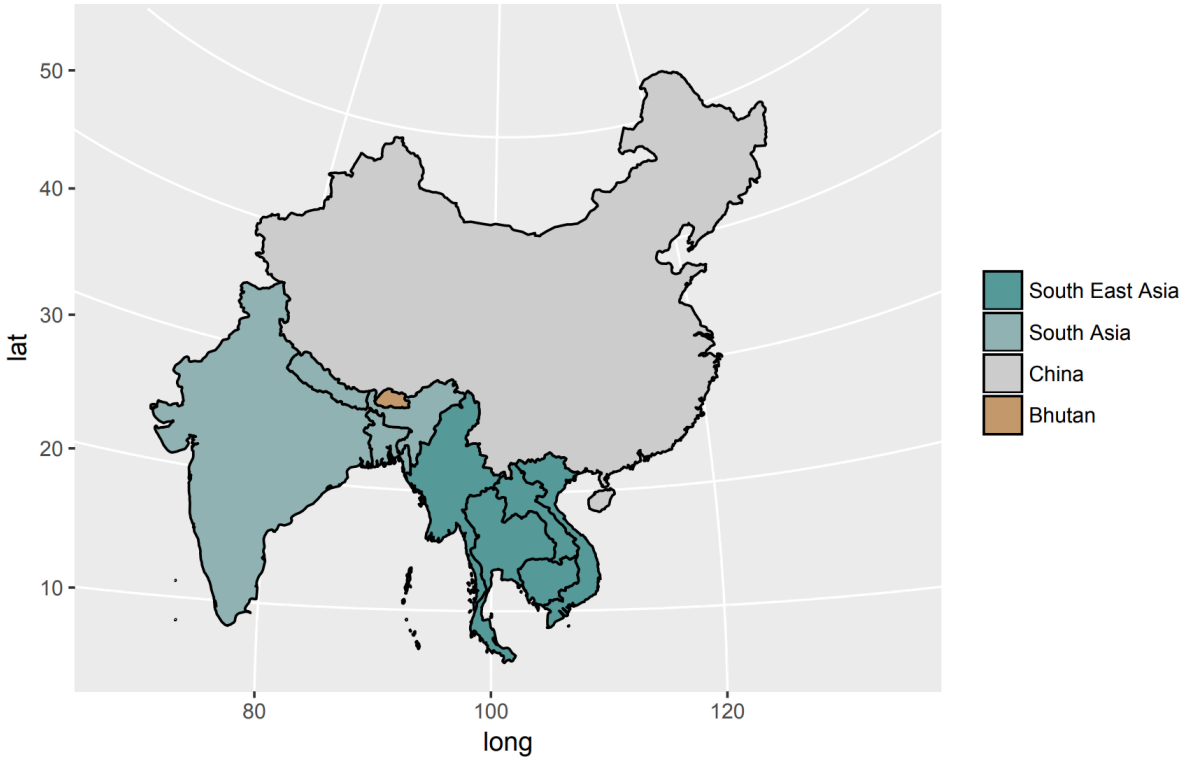

Figure 1. Bhutan in relation to other regions in Asia (Watters 2018). Edited by the author.

There are at least fifty Tibetic languages and around two hundred spoken varieties that belong to a geolinguistic continuum, where varieties closer on the continuum are more mutually intelligible (Tournadre 2014). In Bhutan, over a dozen Tibetic, other Bodish and non-Sino-Tibetan languages are spoken. Dzongkha was made the national language of the country in 1971 (van Driem 1991), and it is taught in schools alongside English. In the eastern and southern areas of Bhutan, Dzongkha serves as a lingua franca, where it is not natively spoken. Despite the fact that nearly all Bhutanese people have an understanding of Dzongkha, it is a primary L1 for only 28% of the country’s population. Tshangla, another non-Tibetic Bodish language, has 24% primary L1 speakers, while Nepali, an Indo-Aryan language, has 25%. The country’s national radio network, Bhutan Broadcasting Service, has all of its segments in these four languages, 60% of which are in Dzongkha, 16% in English, 13% in Tshangla and 8% in Nepali (Dorjee 2014). Multilingualism in Dzongkha and other languages, including English, can be contributing factors to the rapid change to the language (Figure 2).

Figure 2. Linguistic map of Bhutan by native language use (Babbage 2010). Edited by the author.

Dzongkha is written in Tibetan script, an abugida known for its ornate and distinct visual characteristics. The writing system not only reflects the language's linguistic heritage but also plays a pivotal role in the Tibetic cultural identity of Bhutan. It consists of thirty basic consonant letters and four vowel letters, with various combinations and modifiers to represent different sounds. Each character represents a syllable, with the exception of nasal codas, and letters are stacked above or below the root letter, similar to other abugidas (Appendix). However, literacy in the language is historically lower than its fluency in the country. In 2022, 98% of young people aged 15-24 are literate in Dzongkha, while only 65% of all people aged 15 and above are literate in Dzongkha. Nevertheless, literacy rate in Dzongkha or another language has been rising for the past decades (NSB, 2022).

Dzongkha plays a crucial role in Bhutanese society as the official language of the country, used in government documents, legal proceedings, and educational materials. The Bhutanese government places a strong emphasis on its preservation, making it a mandatory subject in schools. Public signage and official communication are typically conducted in both Dzongkha and English.

Dzongkha faces challenges in the modern era, particularly due to the influence of English. Dzongkha is not considered an endangered language, but the language is classified as "vulnerable to endangerment" by UNESCO (2015) and "threatened" by Glottolog's AES (Moseley 2010). As Bhutan is one of the most anglicised countries in Asia due to governmental efforts to modernise the country, Dzongkha is taught secondary to English, and students are prohibited from speaking Dzongkha in school outside of Dzongkha class. A generation of Bhutanese children have grown up speaking Dzongkha at home and English at school and work. Therefore, documenting the current change in the vocabulary, grammar, and sound system of Dzongkha would be beneficial as a case study of diglossia.

Dzongkha's unique linguistic environment also makes it an ideal subject for examining the intricate dynamics of language contact and change. Dzongkha is a useful case study for the effects of areal linguistics, as the Eastern Himalayas have one of the highest densities of languages, many of which belong to different language families and branches (Everett, Blasi & Roberts 2015). On the other hand, East and Southeast Asia have the highest density of languages with tonal contrast (Figure 3), and changes to the tonal system of Dzongkha can provide valuable information for the study of other languages in contact.

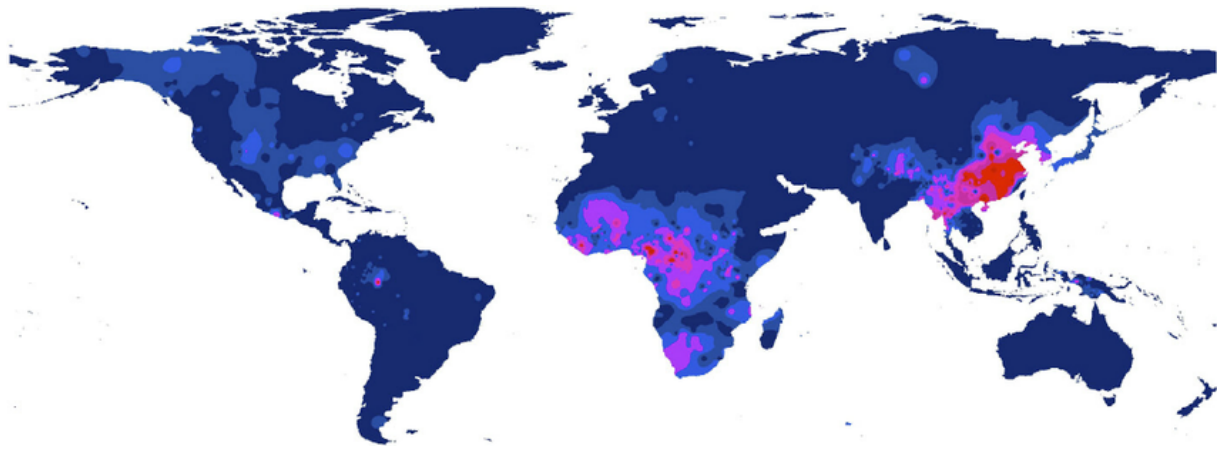

Figure 3. Heat map of language tonality complexity density (Everett, Blasi & Roberts 2015).

The evolving phonological landscape of Dzongkha presents both challenges and opportunities for linguistic research, with the need for updated analysis. Likely due to the small amount of fieldwork done relative to the rate in which the sound system of the language is changing, there are only three comprehensive studies on the subject (Mazaudon & Michailovsky 1988, van Driem 1991, Watters 2018) with known gaps in knowledge.

This thesis aims to provide an overview of the phonetics and phonology of Dzongkha using newly collected field data, with a focus on the paradigm change to the tonal system in the past three decades. The study is based on spoken Dzongkha, as opposed to the language used for text-reading with a more conservative phonology. The scope of this study focuses on the language spoken in western Bhutan, specifically Thimphu and Paro, where the two consultants are from. Overall, the sound system of the language had undergone simplifications in at least the past three decades, where the number of both vowel and consonant phonemes had decreased, vowel length contrast had been lost, and the multiple distinct tonal contours had shifted to a two-way pitch contrast (van Driem 2015; Lee & Kawahara 2018; Watters 2018).

Particularly, the loss of voicing contrast in exchange for pitch contrast can be characterised as a tonogenesis process. Dzongkha was understood as one of the only Bodish languages that retained a four-way phonation contrast for its onset stop and affricate series (van Driem 1991), as its neighbours in the East, such as Tshangla and Kürtop, have three-way contrasts, and Lhasa Tibetan to the North has only a two-way contrast. The voicing contrast in Dzongkha encoded overlapping information with pitch, as high and low vowels only surfaced following voiceless and voiced onset consonants, respectively (Mazaudon & Michailovsky 1988; van Driem 1991). In more recent studies, the language's stop and affricate series were attested to have a three-way phonation contrast, as the voiced aspirated series were noted to surface as either voiced or voiceless unaspirated. Pitch, however, remained contrastive (Watters 2018). This study demonstrates that onset phonation has further transphonologised as vowel pitch contrast.

By delving into the evolving sounds system of Dzongkha, this thesis hopes to contribute to our understanding of language change, shedding light on the broader implications for linguistic research in the eastern Himalayan region.

2. Background

There are existing descriptions of the phonetic and phonological system of Dzongkha, three of which are comprehensive and based on original fieldwork to my knowledge.

The earliest of which was done by Mazaudon & Michailovsky (1988), who categorised the phonological inventory of the language, which includes eight vowels (Table 1) and forty-six consonants (Table 2). Only diphthongs that begin or end with /i/ and /u/ are attested. The place and manner of articulation of /r/ is not specified in the paper. The voiceless nasals /m̥/ and /n̥/ are only found in some dialects and are noted to have merged with /h/ in Thimphu. Examples including /m̥/ or /n̥/ are not provided. The study finds vowel length and nasalisation contrasts in the language, and it notes that nasalised vowels are always long, without specifying which of the vowels can surface as long or nasalised. The authors also provide acoustic evidence for a high/low pitch contrast for all syllables and a tone contour contrast for open long and closed syllables.

	Front	**Back**
High	i | y	u
High-Mid	e | ø	o
Low-Mid	ɛ	
Low	a	

Table 1. Vowels attested by Mazaudon & Michailovsky (1988).

Manner of Articulation	Phonation	Place of Articulation					
		Bilabial	Alveolar	Retroflex	(Alveolo-) Palatal	Velar	Glottal
Stop	Voiced Unaspirated	b	d	ɖ	ɟ	g	
	Voiced Aspirated	bʰ	dʰ	ɖʰ	ɟʰ	gʰ	
	Voiceless Unaspirated	p	t	ʈ	c	k	
	Voiceless Aspirated	pʰ	tʰ	ʈʰ	cʰ	kʰ	
Affricate	Voiced Unaspirated	bʑ	d͡z				
	Voiced Aspirated	bʑʰ	d͡zʰ				
	Voiceless Unaspirated	pɕ	t͡s				
	Voiceless Aspirated	pɕʰ	t͡sʰ				
Fricative	Voiced Unaspirated		z		ʑ		
	Voiced Aspirated		zʰ		ʑʰ		
	Voiceless		s		ɕ		h
Lateral	Voiced		l				
	Voiceless		l̥				
Approximant	Voiced	w	r	j		w	
Nasal	Voiced	m	n		ɲ	ŋ	
	Voiceless	m̥	n̥				

Table 2. Consonants attested by Mazaudon & Michailovsky (1988).

In 1991, van Driem developed an official romanisation scheme, which is still used in the country along with Wylie romanisation (Appendix), in cooperation with the Royal Government of Bhutan that includes eight vowels and forty-four consonants. Since the paper did not use IPA transcription, the phonetic values of some of the sounds are estimated based on place and manner descriptions. The system uses a diacritic (̦) for the “devoiced” series of stops, which are described as voiced and murmured, understood here as voiced aspirated. Van Driem categorises the tonal system into two registers, high and low, where only voiceless onset consonants can surface in a high-pitched syllable, and only their voiced counterparts can surface in a low syllable. Meanwhile, the consonants that can surface in any syllable regardless of pitch (/m; n; ɲ; ŋ; w; j/) are marked with a diacritic (‘). The paper has a note on tone contours that are contrastive in closed and long syllables, which are only attested in central Dzongkha dialects that were not in the scope of the study. The system also includes diacritics for long vowels (^) that can be applied to five of the vowels (/i, e, a, o, u/), based on the finding that length is contrastive for them.

	Front	**Back**
High	i | y	u
High-Mid	e | ø	o
Low-Mid	ɛ	
Low	a	

Table 3. Vowels attested by van Driem (1991).

Manner of Articulation	Phonation	Place of Articulation					
		Bilabial	Alveolar	Retroflex	(Alveolo-) Palatal	Velar	Glottal
Stop	Voiced Unaspirated	b	d	ɖ	ɟ	g	
	Voiced Aspirated	bʰ	dʰ	ɖʰ	ɟʰ	gʰ	
	Voiceless Unaspirated	p	t	ʈ	c	k	
	Voiceless Aspirated	pʰ	tʰ	ʈʰ	cʰ	kʰ	
Affricate	Voiced Unaspirated	bʑ	d͡z				
	Voiced Aspirated	bʑʰ					
	Voiceless Unaspirated	pɕ	t͡s				
	Voiceless Aspirated	pɕʰ	t͡sʰ				
Fricative	Voiced Unaspirated		z		ʑ		
	Voiced Aspirated		zʰ		ʑʰ		
	Voiceless		s		ɕ		h
Lateral	Voiced		l				
	Voiceless		l̥				
Approximant	Voiced	w	r	j		w	
	Voiceless		r̥				
Nasal	Voiced	m	n		ɲ	ŋ	

Table 4. Consonants attested by van Driem (1991).

Watters's (2018) grammar offers an overview of the phonetic, phonological, morphological and syntactic systems of the language based on fieldwork with six speakers, whose age ranged from thirty-five to sixty-two, in addition to conversation data collected from 18 speakers. The paper lists nine vowels and forty-five consonants and consonant clusters. The grammar uses van Driem's "devoiced" diacritic (̥) for the voiceless semi-aspirated series, which has a mean VOT in between that of the voiceless aspirated and unaspirated sounds.

Watters also notes vowel length contrast in all but the two rounded vowels (/y, ø/), while the central vowels (/ɜ, ɐ:/) form a length contrast pair. A notable difference of Watters (2018) is the post-alveolar affricate series, which were attested as palatal stops in previous studies. Similar to van Driem (1991), the paper divides the tonal system into two registers that correlate voiceless onsets with high-pitched syllables and voiced onsets with low-pitched syllables. Additionally, Watters notes the effect of aspiration and phonation on pitch. In particular, aspiration and breath tend to raise pitch, whereas glottalization and creak drop pitch.

	Front	**Central**	**Back**
High	i | y		u
High-Mid	e | ø		o
Low-Mid	ɛ	ɜ	
Low		ɐ:	

Table 5. Vowels attested by Watters (2018).

Manner of Articulation	Phonation	Place of Articulation					
		Bilabial	Alveolar	Retroflex	(Alveolo-) Palatal	Velar	Glottal
Stop	Voiced Unaspirated	b	d	ɖ		g	
	Voiced Aspirated	b̤	d̤	ɖ̤		g̤	
	Voiceless Unaspirated	p	t	ʈ		k	
	Voiceless Aspirated	pʰ	tʰ	ʈʰ		kʰ	
Affricate	Voiced Unaspirated	bd͡ʒ	d͡z		d͡ʒ		
	Voiced Aspirated	bd͡ʒʰ			d͡ʒʰ		
	Voiceless Unaspirated	pt͡ʃ	t͡s		t͡ʃ		
	Voiceless Aspirated	pt͡ʃʰ	t͡sʰ		t͡ʃʰ		
Fricative	Voiced Unaspirated		z		ʒ		
	Voiced Aspirated		z̤		ʒ̤		
	Voiceless		s		ʃ		h
Lateral	Voiced		l				
	Voiceless		l̥				
Approximant	Voiced	w	ɾ	j		w	
	Voiceless		ɾ̥				
Nasal	Voiced	m	n		ɲ	ŋ	
	Voiceless		n̥				

Table 6. Consonants attested by Watters (2018).

3. Methods

Data was collected from elicitation sessions with two language consultants from Bhutan, both of whom were university students in the United States, living part time in each country.

Speaker 1 lived in Thimphu, the capital and largest city of Bhutan. Her mother being a Nepalese immigrant, Nepali was regularly heard in her home in addition to Dzongkha and English. Elicitations with her took place in the U.S. over the course of three months in 2021, during which about one thousand hours of audio recordings were made. The elicitations ranged from individual words for phonetic measurements to longer phrases and monologues. Phonetic measurements were then taken in Praat and then charted and plotted on Google Sheets.

Speaker 2 grew up in Paro, the third-largest Bhutanese city, and moved to Thimphu recently. She spoke and heard Dzongkha and English in her home. Elicitations with her were conducted over the internet asynchronously over the course of one week in 2023, during which about one hour of audio recordings were collected. The elicitations included individual words and short phrases for phonological analysis.

4. Results

In this section, the findings from the elicitation sessions are presented, focusing on the vowels, consonants, phonotactics, and tonal characteristics of the sound system. Analysis on the sounds and their variations observed in this study are discussed along with the tonal system and its interactions with the sounds. Throughout the presentation of the results, they are compared with data from existing literature, and discrepancies are analysed.

4.1. Vowels

Vowels in Dzongkha are sounds that function as syllabic nuclei. Formant measurements suggest that there are five vowels: /i, e, a, o, u/. In terms of frontness, /i, e/ are front vowels, and /a, o, u/ are back. In terms of height, /i, u/ are high, while /e, o/ are mid, and /a/ is low.

	Front	**Back**
High	/i/	/u/
Mid	/e/	/o/
Low	/a/	

Table 7. Vowels with place of articulation.

Vowel	**Example**	**Gloss**
/i/	[ʔí]	“village”
/e/	[ʔékó]	“neck”
/a/	[ʔèp]	“man”
/o/	[ʔòm]	“milk, breasts”
/u/	[ʔú]	“hair.FRM”

Table 8. Vowels with example words and gloss.

Each vowel has a range of realisations. Formant measurements were taken on forty-five tokens. Three repetitions were done for each word, lest their prosodic positions skew the results.

/i/ centers around [i]. While /i/ is usually pronounced unrounded, it can be pronounced rounded as [y] in some words in text-reading, reflecting the historical merger with */y/. For the measurements, our consultants pronounced the vowel as [i] consistently.

/e/ centers around [ɐ̟]. While /e/ is usually pronounced unrounded, it can be pronounced rounded as [ø] in some words in text-reading, reflecting the historical merger with */ø/. For the measurements, our consultants pronounced the vowel as [e] consistently.

/a/ centers around [ɐ], and it is raised to around [ə] before an approximant coda. See Phonotactics.

/o/ centers around [o̞].

/u/ centers around [ʊ̟], more front than the other back vowel, /o/.

Vowel	Example	Gloss	Mean ± SD F1 (Hz)	Mean ± SD F2 (Hz)
/i/	/ʔí/	“village”	272 ± 21	2359 ± 120
	/ʑì/	“four”		
	/mì/	“fall”		
/e/	/pè/	“do”	433 ± 19	2063 ± 40
	/t͡ɕʰé/	2.SG.INFRM “you”		
	/lè/	“than”		
/a/	/tà/	“arrow”	581 ± 29	1227 ± 80
	/kʰá/	“mouth, language”		
	/t͡sá/	“grass, vein”		
/o/	/pʰó/	“belly”	416 ± 28	804 ± 38
	/mó/	3.SG.F.INFRM “she, her”		
	/t͡sʰó/	“lake, dinner”		
/u/	/ʈù/	"dragon, six"	345 ± 27	1190 ± 81
	/zù/	"body"		
	/t͡ɕʰú/	"water"		

Table 9. Formant 1 and 2 measurements of each vowel.

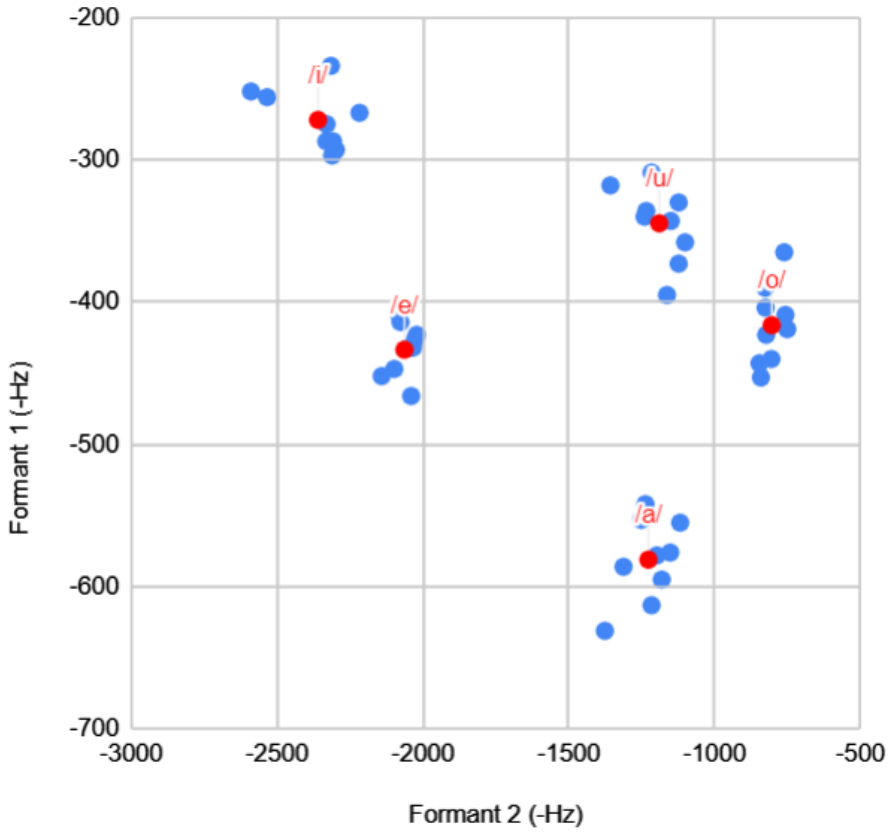

Figure 4. Scatterplot of Formant 2 versus Formant 1 frequencies of forty-five tokens. The orientation of the axes are reversed to match the IPA vowel quadrilateral. The mean value of each vowel is in red.

Vowel length is not attested to be contrastive in this study, since no minimal pair was found. Length measurements were taken on the first vowel of twenty-eight words. There is not a significant division between any two groups of lengths, though monosyllabic words tend to have a longer vowel.

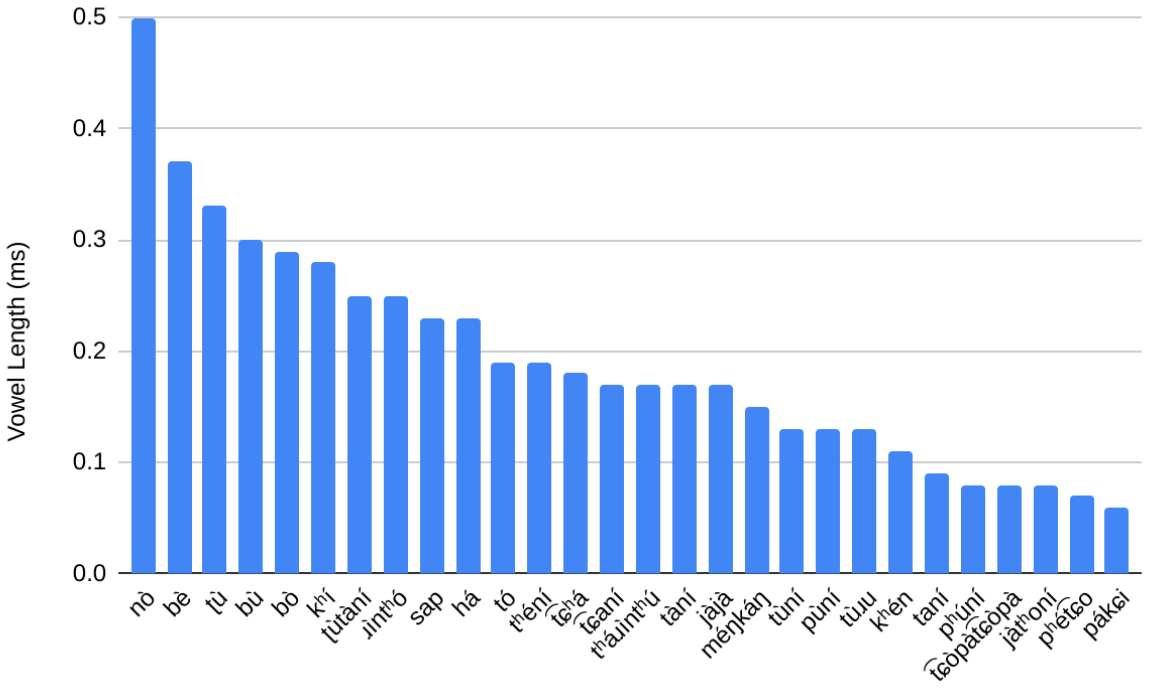

Figure 5. Lengths of the first vowel of twenty-eight words.

4.2. Consonants

Consonants in Dzongkha are phonemes that serve as syllable onsets or codas. Thirty consonant phonemes are attested in this study, each with various manners of articulation, including stops, affricates, fricatives, nasals, approximants, and laterals. All thirty are attested at syllable onset, but only ten (/-p, -t, -k; -m, -n, -ŋ; -s; -ɹ, -j, -w/) are attested at syllable coda.

Manner of Articulation	Phonation	Place of Articulation				
		Bilabial	Alveolar	Retroflex / (Alveolo-) Palatal	Velar	Glottal
Stop	Unaspirated	p	t	ʈ	k	ʔ
	Aspirated	pʰ	tʰ	ʈʰ	kʰ	
Affricate	Unaspirated		t͡s	t͡ɕ		
	Aspirated		t͡sʰ	t͡ɕʰ		
Fricative	Voiced		z	ʑ		ɦ
	Voiceless	ɸ	s	ɕ		h
Lateral	Voiced		l			
	Voiceless		l̥			
Approximant	Voiced	w	ɹ	j	w	
Nasal	Voiced	m	n	ɲ	ŋ	
	Voiceless		n̥			

Table 10. Consonants with place and manner of articulation.

Consonant	Example	Gloss
/p/	[pékó]	“skin”
/pʰ/	[pʰó]	“belly”
/t/	[té]	“horse”
/tʰ/	[tʰép]	“rope”
/ʈ/	[ʈéméɕép]	“naughty”
/ʈʰ/	[ʈʰé]	“blood”
/k/	[kép]	“white”
/kʰ/	[kʰé]	“mouth”
/ʔ/	[ʔéj]	“mother”
/t͡s/	[t͡sé]	“grass, vein”
/t͡sʰ/	[t͡sʰé]	“salt”
/t͡ɕ/	[t͡ɕép]	“do”
/t͡ɕʰ/	[t͡ɕʰé]	“hair.INFRM, monkey”
/ɸ/	[káɸí]	“coffee”
/s/	[sḗm]	“soul, mind”
/z/	[zè]	“drip”
/ɕ/	[ɕé]	“meat”
/ʑ/	[ʑè]	“eat.FRM”
/ɦ/	[ɦòlé]	“under”
/h/	[hémé]	“before (in time)”
/m/	[mò]	3.SG.F “she, her”
/n/	[nḗŋkʰé]	“sky”
/n̥/	[n̥épé, l̥épé, hépé]	“nose.INFRM’
/ɲ/	[ɲè]	“fish”
/ŋ/	[ŋé]	“five”
/w/	[tèwé]	“moon”
/r/	[rò͡tɕʰí]	“dog”
/j/	[jé]	“up”
/l/	[lèp]	“hand”
/l̥/	[l̥é]	“deity”

Table 11. Consonants with example words and gloss.

4.2.1. Stops and Affricates

Thirteen stops and affricates are attested in this study: /p/, /pʰ/; /t/, /tʰ/; /ʈ/, /ʈʰ/; /k/, /kʰ/; /t͡s/, /t͡sʰ/; /t͡ɕ/, /t͡ɕʰ/; /ʔ/. Each of the six non-glottal stop and affricate series have a two-way phonation contrast, although they are written with a three-way contrast.

Voice onset time (VOT) measurements were taken on five words for each stop and affricate onset. Voiced stops and affricates are expected to have negative VOTs, while their voiceless counterparts are expected to have positive VOTs. Larger positive VOTs indicate aspiration. The results suggest that both written voice and voiceless unaspirated stops and affricates have small positive VOTs, while written voiceless aspirated stops and affricates have large positive VOTs. Overall, all stops and affricates are voiceless, but they contrast in aspiration.

Stop Onset	Written	Transcription	Gloss	Mean ± SD VOT (ms)
ད <d> /t-/	མདའ་ <mda’>	[tè]	"arrow"	15.52 ± 4.29
	གང་འདི་ <ga ‘di>	[kètì]	"which"	
	དེ་ལས་ན་ <de lasan>	[tèlè]	"afterward"	
	རྡོ་ <rdo>	[tò]	"rock"	
	དུད་པ་ <dud pa>	[tùpé]	"smoke"	
ཏ <t> /t-/	རྟ་ <rta>	[té]	"horse"	14.24 ± 2.99
	ཨ་རྟག་རང་ <a rtag rang>	[ʔétéɹè]	"always"	
	རྟག་བུ་རང་ <rtag bu rang>	[tépùɹè]	"always"	
	སྟུག་རས་ <stug ras>	[túɹè]	"rag"	
	ལྟོ་ <lto>	[tó]	"cooked rice, food"	
ཐ <th> /tʰ-/	ཐག་པ་ <thag pa>	[tʰép]	"rope"	68.94 ± 20.76
	ཐོག་ཁར་ <thog khar>	[tʰókʰé]	"top"	
	ཐུ་རམ་ <thuram>	[tʰúm]	"spoon"	
	འཐུ་ནི་ <’thu ni>	[tʰűní]	"pick up, gather"	
	ཐུག་པ་ <thug pa>	[tʰúp]	"soup"	

Table 12. Voice onset times of the alveolar stop series, /t/ and /tʰ/, organised by their three written forms, <d>, <t> and <th>.

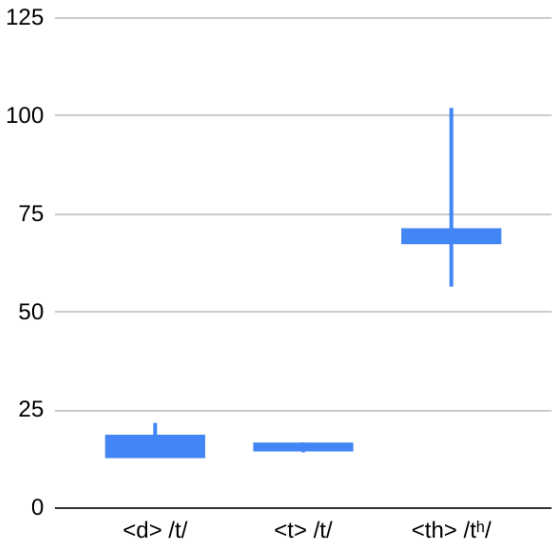

Figure 6. Candlestick plot of voice onset times of the onset alveolar stop series, /t/ and /tʰ/, organised by their three written forms, <d>, <t> and <th>.

4.2.2. Fricatives

Seven fricatives are attested in this study, each belonging to a pair that contrasts in voicing, except for the bilabial fricative: /ɸ/; /s/, /z/; /ɕ/, /ʑ/; /h/, /ɦ/. The bilabial /ɸ/ is only attested for English loanwords containing [f].

Spectral centre of gravity (CoG) measurements were taken on five words for each fricative onset. CoG, the weighted average of frequencies, can be used to acoustically distinguish between fricative sounds. Front fricatives have higher CoGs, while back fricatives have lower CoGs. For example, among /s/, /ɕ/ and /h/, /s/ has the highest mean CoG at 6336.8 Hz, while /ɕ/ has a lower mean CoG at 5106.8 Hz, and /h/ has the lowest mean CoG at 3168 Hz.

Fricative Onset	Written	Transcription	Gloss	Mean ± SD CoG (Hz)
<s> /s-/	སེམས་ <sems>	[sḗm]	"mind.INFRM"	6336.8 ± 350.0
	གསར་པ་ <gsarpa>	[sép]	"new"	
	སེར་པོ་ <serpo>	[sép]	"yellow"	
	གསུམ་ <gsum>	[sṹm]	"three"	
	གསོ་ལཝ་ <gsolawa>	[séw]	"lunch"	
<sh> /ɕ-/	ཤ་ <sha>	[ɕé]	"meat"	5106.8 ± 582.0
	གཤོག་སྒྲོ་ <gshog sgro>	[ɕóʈò]	"feather"	
	ཤིང་ <shing>	[ɕḯŋ]	"tree"	
	ལེ་ཤཱ་ <le shā>	[lèɕé]	"much"	
	མང་ཤོས་རང་ <mang shos rang>	[mɛ̀ŋɕóɹɛ̀ŋ]	"mostly"	
<h> /h-/	ཧེ་མ་ <hen ma>	[hɛ̰́mé]	"before"	3168.0 ± 375.0
	ཧོན་མ་ <honma>	[hɛ̰́m]	"blue"	
	ཧ་གོ་ནི་ <ha go ni>	[hɛ̰́kɔ̃́ní]	"understand"	
	ཧཱ་ <hā>	[hɛ̰́]	"Haa District"	

Table 13. Centres of gravity (CoG) of three Dzongkha fricative onsets, /s/, /ɕ/ and /h/, written as <s>, <sh> and <h>, respectively.

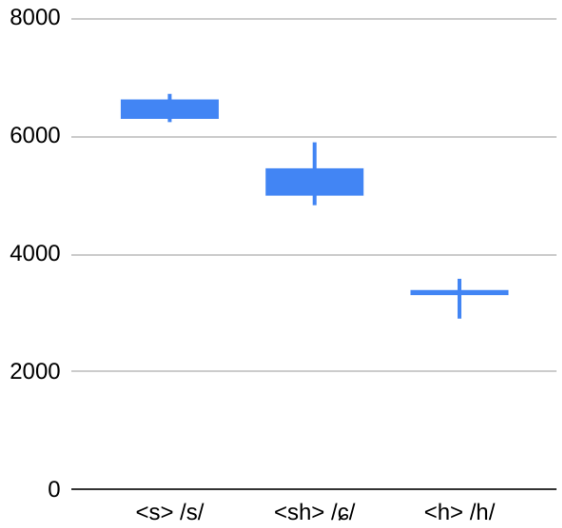

Figure 7. Candlestick plot of the centres of gravity of three fricatives, /s/, /ɕ/ and /h/, written as <s>, <sh>, <h>, respectively.

4.2.3. Nasals

Five nasals are attested in this study: /m/; /n/, /n̥/; /ɲ/; /ŋ/. The voiced bilabial, alveolar and velar nasals are attested at both syllable onset and coda positions, while the palatal and voiceless alveolar nasals are only attested in the onset. In addition, the voiceless alveolar /n̥/ surfaces as [h] for Speaker 1, from Thimphu, and as [n̥] and [l̥] for Speaker 2, from Paro.

4.2.4. Approximants

Five approximants are attested in this study. There is one alveolar lateral approximant series, consisting of the voiced /l/ and the voiceless /l̥/. Additionally, there are three non-lateral approximant phonemes: the voiced labial-velar approximant /w/, the alveolar approximant /ɹ/, and the palatal approximant /j/.

4.3. Phonotactics

Syllables are in the form of /$C_1V(C_2)$/.

Ten consonants are attested at syllable coda, the six most common of which are all stops or nasals: /-p, -t, -k; -m, -n, -ŋ/. These consonants likely survived as codas due to the ease of pronunciation and perception. The fricative /s/ and approximants /w/ and /j/ are also attested as coda consonants, but they are less frequent.

Coda Consonant		**Example**	
Manner of Articulation	**Transcription**	**Transcription**	**Gloss**
Stop	/-p/	/sáp/	“new”
	/-t/	/pàsámèt/	“blunt”
	/-k/	/pákɞí/	“bamboo”
Nasal	/-m/	/sṹm/	“three”
	/-n/	/ɲn̂̀n/	“ear.FRM”
	/-ŋ/	/mĩ̀ŋ/	“name”
Fricative	/-s/	/pùs/	“boy”
Approximant	/-ɹ/	/tʰé̯kʰá̯ɹ/	“blunt”
	/-j/	/ʔáj/	“mother”
	/-w/	/tàw/	“month”

Table 14. Syllable coda consonants with manner of articulation, example words and gloss.

/a/ is raised to around [ə] before an approximant coda. The following example shows the comparison of /dàw/ “month”, where /a/ is raised, with /dàwá/, where /a/ is not raised, since the /w/ is the onset of the following syllable.

5. a → ə \ _A|
    a. “month” /dàw/ → [dəw]
    b. “moon” /dàwá/ → [dɛ̀wɛ́]

The following three rules use the derivation of the same example phrase, /pòm ní/ [pɔ̀ ní] “to/will swell”.

Vowels preceding nasal codas are prenasalised.

6. $V \rightarrow \tilde{V}_{[+nasal]}$ \ _N|
    a. /pòm ní/ → /pɔ̀̃m ní/
    big INF/FTR
    “to/will swell”

Nasal codas assimilate to the place of the onset consonant of the following syllable.

7. $N \rightarrow N_{[\alpha\ place]}$ \ $\_C_{[\alpha\ place]}$
    a. /pɔ̀̃m ní/ → /pɔ̀̃n ní/
    big INF/FTR
    “to/will swell”

Geminate consonants, i.e. paused before onset or lengthened, are not attested, and they are resistant to forming at syllable boundaries.

8. $C_\alpha \rightarrow \varnothing$ \ _C_α
    a. /pǒn ní/ → [pǒ ní]
        big INF/FTR
        “to/will swell”

/a/ is raised to around [ə] before an approximant coda. The following example shows the comparison of /dàw/ “month”, where /a/ is raised, with /dàwá/, where /a/ is not raised, since the /w/ is the onset of the following syllable.

9. a → ə \ _A|
    a. “month” /dàw/ → [də̀w]
    b. “moon” /dàwá/ → [də̀wé]

Consonant cluster onsets are not attested, not even the sonorically more preferable ones involving glides. However, the conservative orthography of Dzongkha often leads to the more formal spelling-pronunciation of the heterorganic stop-affricate clusters, [pt͡ɕ] and [pt͡ɕʰ]. In spoken language, the stops at the onset of these clusters are omitted.

Onset Consonant Cluster			**Example Word**			
Written	**Text-reading**	**Spoken**	**Written**	**Text-reading**	**Spoken**	**Gloss**
བྱ <by>	[pt͡ɕ-]	[t͡ɕ-]	བྱ་ <bya>	[pt͡ɕè]	[t͡ɕè]	“bird”
པྱ <py>	[pt͡ɕ-]	[t͡ɕ-]	སྤྱ་ <spya>	[pt͡ɕé]	[t͡ɕé]	“monkey”
ཕྱ <phy>	[pt͡ɕʰ-]	[t͡ɕʰ-]	འཕྱག་ <'phyag>	[pt͡ɕʰé]	[t͡ɕʰé]	“wipe, sweep”

Table 15. Consonant clusters used in text-reading that are absent in informal speech with written form, example word and gloss.

4.4. Tone

Two contrastive pitches are attested in this study, high and low, for all onset series. High vowels are attested following both aspirated and unaspirated onsets, while low vowels are only attested following unaspirated onsets. In addition, two vowel phonations are attested, modal and breathy, which are contrastive for some onsets. Breathy vowels are attested following both aspirated and unaspirated onsets, while modal vowels are only attested following unaspirated onsets.

<table>
<tr><th colspan="2">Historical Onset Consonant</th><th colspan="2">Onset Consonant</th><th colspan="2">Vowel</th></tr>
<tr><th>Phonation</th><th colspan="2">Manner of Articulation</th><th>Phonation</th><th>Pitch</th><th>Phonation</th></tr>
<tr><td>Voiced Unaspirated</td><td colspan="2" rowspan="4">Stop, Affricate</td><td rowspan="3">Voiceless Unaspirated</td><td rowspan="2">Low</td><td>Modal</td></tr>
<tr><td>Voiced Aspirated</td><td>Breathy</td></tr>
<tr><td>Voiceless Unaspirated</td><td rowspan="2">High</td><td>Modal</td></tr>
<tr><td>Voiceless Aspirated</td><td>Voiceless Aspirated</td><td>Breathy</td></tr>
<tr><td>Voiced Unaspirated</td><td colspan="2" rowspan="3">Fricative</td><td>Voiced</td><td rowspan="2">Low</td><td>Modal</td></tr>
<tr><td>Voiced Aspirated</td><td rowspan="2">Voiceless</td><td>Breathy</td></tr>
<tr><td>Voiceless</td><td>High</td><td>Modal</td></tr>
<tr><td>Voiced</td><td colspan="2" rowspan="2">Nasal, Approximant</td><td>Voiced</td><td>High, Low</td><td>Modal</td></tr>
<tr><td>Voiceless</td><td>Voiceless</td><td>High</td><td>Breathy</td></tr>
</table>

Table 16. Vowel pitch and phonation in relation to historical and modern onset consonant phonation.

Onset Consonant		Vowel		Example	
Written	**Transcription**	**Pitch**	**Phonation**	**Transcription**	**Gloss**
ཨ <b>	/p-/	Low	Modal	[pè]	“release”
			Breathy	[pe̤]	“cow”
ཨ <p>		High	Modal	[pé]	“attach”
ཕ <ph>	/pʰ-/		Breathy	[pʰé̤]	“there”
ད <d>	/t-/	Low	Modal	[tò]	“stone”
			Breathy	[tò̤]	“hole”
ཏ <t>		High	Modal	[tó]	“cooked rice”
ཐ <th>	/tʰ-/		Breathy	[tʰó̤]	“see”
ག <g>	/k-/	Low	Modal	[kò]	“hear”
			Breathy	[kò̤]	“door”
ཀ <k>		High	Modal	[kó]	“throw”
ཁ <kh>	/kʰ-/		Breathy	[kʰó̤]	3.SG.M “he, him”
ཛ <dz>	/t͡s-/	Low	Modal	[t͡sè]	“drip”
			Breathy	?	?
ཙ <ts>		High	Modal	[t͡sé]	“grass, vein”
ཚ <tsh>	/t͡sʰ-/		Breathy	[t͡sʰé̤]	“salt”
ཇ <j>	/t͡ɕ-/	Low	Modal	[t͡ɕè]	“sticky”
			Breathy	[t͡ɕè̤]	“tea”
ཅ <c>		High	Modal	[t͡ɕé]	“hair”
ཆ <ch>	/t͡ɕʰ-/		Breathy	[t͡ɕʰé̤]	“wipe”

Table 17. Pitch and phonation minimal and near-minimal pairs with orthography and gloss for various stop and affricate onset consonants.

Onset Consonant		Vowel		Example	
Written	**Transcription**	**Pitch**	**Phonation**	**Transcription**	**Gloss**
ꔹ <z>	/z-/	Low	Modal	[zè]	“rainbow”
	/s-/		Breathy	[sɛ̰̀]	“eat.INFRM”
ꔤ <s>		High	Modal	[sé]	“land, soil”
ꔻ <ʑ>	/ʑ-/	Low	Modal	[ʑè]	“keep”
	/ɕ-/		Breathy	[ɕɛ̰̀]	“day and night”
ꔼ <ɕ>		High	Modal	[ɕé]	"meat"
ꔦ <m>	/m-/	High	Modal	[mí]	“person”
		Low		[mì]	“fire”
ꔧ <n>	/n-/	High	Modal	[nó]	“think”
		Low		[nò]	“cattle”
ꔨ <lh>	/n̥-/	High	Breathy	[n̥épé]	“nose”
ꔩ <ɲ>	/ɲ-/	High	Modal	[ɲé]	“borrow”
		Low		[ɲè]	“fish”
ꔪ <ŋ>	/ŋ-/	High	Modal	[ŋé]	“five”
		Low		[ŋè]	1.SG “I, me”
ꔫ <l>	/l-/	High	Modal	[ló]	“cough”
		Low		[lò]	“year, age”
ꔨ <lh>	/l̥-/	High	Breathy	[l̥ɛ́]	“deity”
ꔬ <y>	/j-/	High	Modal	[jé]	“itch”
		Low		[jè]	“up”

Table 18. Pitch and phonation minimal and near-minimal pairs with orthography and gloss for various fricative, nasal and approximant onset consonants.

Tonal contour is not attested to be contrastive in this study. The pitches at the onset and coda of the first syllable of twenty-eight words at the end of prosodic phrases were measured in Praat and plotted in Google Sheets. At the onset, there is a clear division between the high and low pitches at 210 to 215 Hz for our consultants. Although there is not a clear division between two groups of coda pitches, no minimal pairs between them are attested.

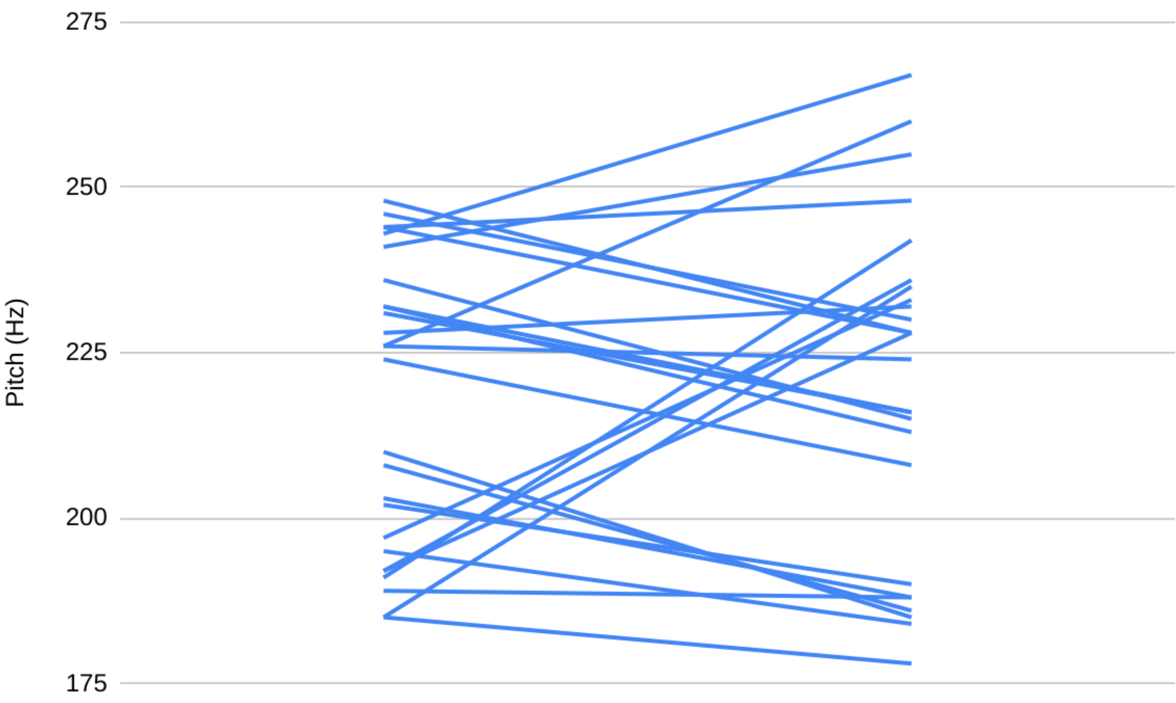

Figure 8. Onset and coda pitches of the first syllable of twenty-eight words.

Additionally, the pitch of the syllable at the end of a prosodic phrase is attested to rise regardless of underlying vowel pitch. Elsewhere, syllables have a falling pitch.

10. $V \rightarrow V_{[+falling]}$
    a. *big shop*

"shop	big"
/t͡sʰóŋkʰáŋ	pòm/
[t͡sʰõ̤ŋ˦kʰẽ̤ŋ˦	põm˧]

    b. *The shop is big.*

"shop	TOP	big	COP"
/t͡sʰóŋkʰáŋ	tì	pòm	jì/
[t͡sʰõ̤ŋ˦kʰãŋ˦	ti˩	põm˩	ji˧]

11. $V \rightarrow V_{[+rising]}$ \ _(C)||

The pitch difference between the syllable onset and coda of the first syllable of twenty-eight words at the end of prosodic phrases were measured and plotted. With the exception of "hospital" /méŋkáŋ/, all syllables with positive pitch changes are at the end of phrases. Even so, /méŋkáŋ/ has the smallest positive pitch change of 4 Hz.

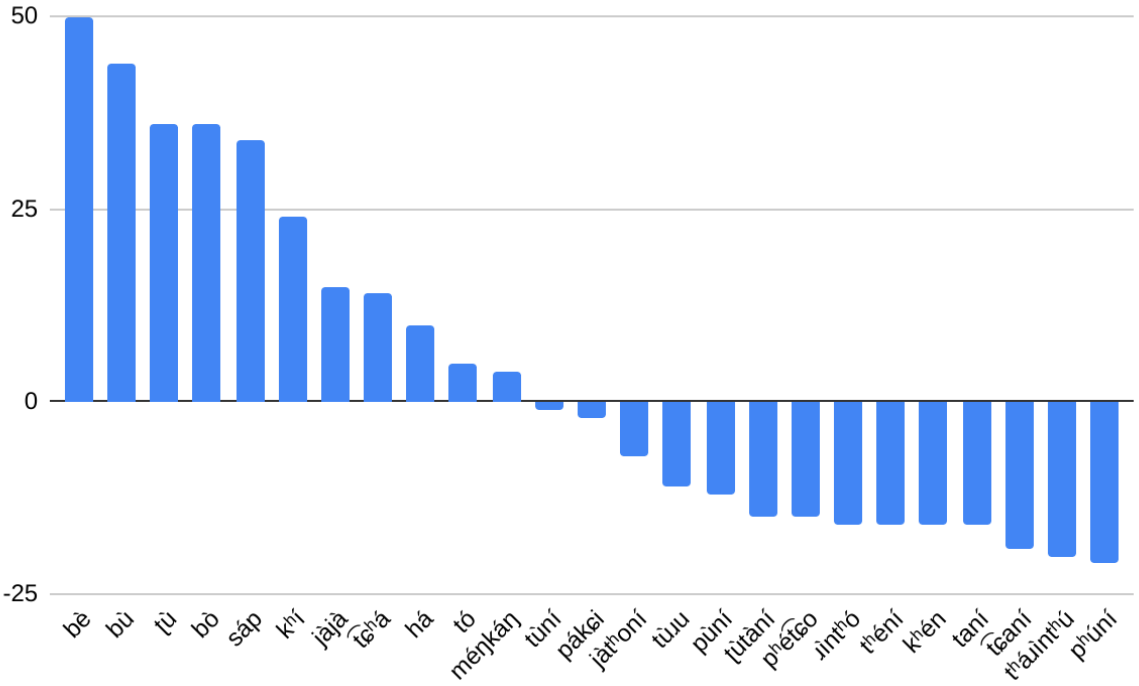

Figure 9. Onset to coda pitch changes of the first syllable of twenty-eight words at the end of prosodic phrases.

As a product of the underlying pitch of the vowel of a syllable and the prosodic position of the syllable, four syllable tones are attested: high falling, high rising, low falling, low rising.

Vowel Pitch	Prosodic Position	Prosodic Tonal Contour	Syllable Tone
High	Beginning, Middle	Falling	High Falling
	End	Rising	High Rising
Low	Beginning, Middle	Falling	Low Falling
	End	Rising	Low Rising

Table 19. Syllable tones as a result of vowel pitch, prosodic position and tonal contour.

5. Discussion

In this section, results from this study are compared with existing literature, and differences between them are discussed.

5.1. Vowels

Eight vowel phonemes were attested in existing literature (Mazaudon & Michailovsky 1988; van Driem 1991; Watters 2018). The front vowel space, which five phonemes used to occupy, had merged into two phonemes, while the back vowel space remained unchanged. In particular, words with /i, y/ attested in past studies are attested as /i/ in this study, and those with /e, ø, ɛ/ are attested as /e/. The remaining three vowels, /a, o, u/, remain relatively unchanged.

Mazaudon & Michailovsky (1988)	van Driem (1991)	Watters (2018)	This study
i	i	i	i
y	y	y	
e	e	e	e
ø	ø	ø	
ɛ	ɛ	ɛ	
a	a	a	a
o	o	o	o
u	u	u	u

Table 20. Comparison of vowels attested in existing literature and this study.

Mazaudon & Michailovsky (1988)	Van Driem (1991)	Watters (2018)	This study	Gloss
[tɕʰi˥m]	[tɕʰím]	[t͡ʃʰɪ́m]	[tɕʰím]	“house”
[byˬː]	[bỳ]	[bỳ]	[pì]	“snake”
[se˥m]	[sém]	[sém]	[sém]	“soul, mind”
[cʰø˥ː]	[cʰǿ]	[t͡ʃʰǿ]	[t͡ɕʰé]	2.SG.INFRM “you”
[gɛˬː]	[gɛ̀ː]	[gɛ̀ʔ]	[kè]	“eight”
[gʰà]	[g̥à:]	[g̥ɜ̀]	[kè]	“who”
[só]	[só]	[só]	[só]	“tooth”
[sum˥]	[súm]	[súm]	[súm]	“three”

Table 21. Comparison of vowel examples attested in existing literature and this study.

5.2. Consonants

In this study, the voiced onsets given by existing literature are attested as voiceless unaspirated with a low pitch. The progression from 1988 to today suggests the language had undergone a transphonologisation of onset voicing as pitches, a similar process as the Sinospheric Tonogenesis. Many of the other Bodish languages spoken in the East of the country, such as Tshangla (Andvik 2010) and Kürtop (Hyslop 2011), have a partial three-way phonation contrast. Watters (2018) claims that the adaptation of Dzongkha by native speakers of those languages and their migration to urban centres in western Bhutan are factors for the loss of the four-way phonation contrast. In a similar vein, language contact with languages with a two-way phonation contrast, such as English, Lhasa Tibetan and Standard Modern Chinese, may be a reason for the additional loss of the three-way phonation contrast. Particularly, the English fluency of both our consultants may be the cause for their two-way phonation contrast.

5.2.1. Stops and Affricates

Earlier literature (Mazaudon & Michailovsky 1988; van Driem 1991) attests twenty-four stops and affricates, with the addition of the voiced aspirated series (/bʰ/; /dʰ/; /ɖʰ/; /gʰ/; /d͡zʰ]; [d͡ʑʰ]). In Watters's grammar (2008), these sounds were analysed as a distinct series of underlying phonemes that surfaced as either voiced or voiceless semi-aspirated, varying based on the speaker, with a mean VOT in between that of the voiceless aspirated and unaspirated stops and affricates. In this study, all voiced stops and affricates are attested as voiceless unaspirated that condition a low pitch in the following vowel. In addition, none of the existing literature attests the glottal stop in the onset.

Written	Pitch	Mazaudon & Michailovsky (1988); van Driem (1991)	Watters (2018)	This study
བ <b>	Low	[b]	[b]	[p]
		[bʰ]	[b̥]	
པ <p>	High	[p]	[p]	
ཕ <ph>		[pʰ]	[pʰ]	[pʰ]
ད <d>	Low	[d]	[d]	[t]
		[dʰ]	[d̥]	
ཏ <t>	High	[t]	[t]	
ཐ <th>		[tʰ]	[tʰ]	[tʰ]
ག <g>	Low	[g]	[g]	[k]
		[gʰ]	[g̥]	
ཀ <k>	High	[k]	[k]	
ཁ <kh>		[kʰ]	[kʰ]	[kʰ]
ཛ <dz>	Low	[d͡z]	[d͡z]	[t͡s]
		[d͡zʰ]	[d̥͡z̥]	
ཙ <ts>	High	[t͡s]	[t͡s]	
ཚ <tsh>		[t͡sʰ]	[t͡sʰ]	[t͡sʰ]
ཇ <j>	Low	[ɟ]	[d͡ʒ]	[t͡ɕ]
		[ɟʰ]	[d̥͡ʒ̥]	
ཅ <c>	High	[c]	[t͡ʃ]	
ཆ <ch>		[cʰ]	[t͡ʃʰ]	[t͡ɕʰ]

Table 22. Comparison of stop and affricate onsets attested in existing literature and this study with written form, conditioned pitch and gloss.

Written	Mazaudon & Michailovsky (1988)	Van Driem (1991)	Watters (2018)	This study	Gloss
སྦལ་པ <sbalpa>	[bɛ:pɻ]	[bè:p]	[bèp, bè:p, bà:p]	[pèp]	“frog”
བ <ba>	[bʰà]	[bʰà]	[b̥à]	[pà]	“cow”
པགས་ཀོ <pags ko>	?	[pákó]	[pɜ́k ko]	[pékó]	“skin”
ཕག་པ <phag pa>	[pʰapɻ, pʰák]	[pʰáp]	[pʰɜ́p, pʰɜ́k]	[pʰép]	“pig”
སྨྱུག་མ <mjug ma>	[ɟumɻ, ɟùmà]	[ɟùmà]	[d͡ʒùm]	[t͡ɕùmè]	“tail”
ཇ <ja>	[ɟʰà]	[ɟʰà]	[d̥͡ʒ̥à]	[t͡ɕè]	“tea”
ལྕེ <lce>	[cé]	[cé]	[t͡ʃéʰ]	[t͡ɕé]	“tongue”
ཆར་པ <char pa>	[cʰa:pɻ]	[cʰá:p]	[t͡ʃʰá:p]	[t͡ɕʰɛ́p]	“rain”

Table 23. Comparison of example words with various stop and affricate onsets attested in existing literature and this study with written form and gloss.

5.2.2. Fricatives

Earlier literature (Mazaudon & Michailovsky 1988; van Driem 1991) attests eight fricatives, with the addition of the voiced aspirated series ([zʰ]; [ʑʰ]). In Watters’s (2008) grammar, these sounds were attested as voiceless semi-aspirated, with a VOT in between that of the voiced and voiceless fricatives, denoted by the [̥] diacritic. In this study, the same sounds are attested as voiceless with a VOT similar to other voiceless fricatives. Only Watters (2018) attests the bilabial fricative [ɸ] that are also found in this study. However, comparable examples between the studies cannot be found. In addition, Watters (2018) attests the voiced counterpart [β], which is not found in this study. The voiced glottal fricative [ɦ] is not attested in existing literature, but its voiced counterpart /h/ is.

Written	Pitch	Mazaudon & Michailovsky (1988); van Driem (1991)	Watters (2018)	This study
བ <b>	Low	?	?	[β]
ཕ <ph>	High	?	?	[ɸ]
ཟ <z>	Low	[z]	[z]	[z]
		[zʰ]	[z̥]	[s]
ས <s>	High	[s]	[s]	
ཞ <zh>	Low	[ʑ]	[ʒ]	[ʑ]
		[ʑʰ]	[ʑ̥]	[ɕ]
ཤ <sh>	High	[ɕ]	[ʃ]	
འ <’>	Low	∅	∅	[ɦ]
ཧ <h>	High	[h]	[h]	[h]

Table 24. Comparison of fricative onsets attested in existing literature and this study with written form, conditioned pitch and gloss.

Written	Mazaudon & Michailovsky (1988)	Van Driem (1991)	Watters (2018)	This study	Gloss
ཀཱ་ཕི་ <kaa phi>	?	?	?	[kɛ́ɸí]	“coffee”
གཟུགས <gzugs>	[zu:˩]	[zù:]	[zù:ʔ]	[zù]	“body”
ཟོག་པ་ <zogpa>	[zʰop˩]	[zʰòp]	[ʐ̩òp]	[sòp]	“bad”
གསར་པ་ <gsarpa>	[sa:p˥]	[sá:p]	[sá:p]	[sɛ́p]	“new”
བཞི་ <bzhi>	[ʑì]	[ʑì]	[ʒì]	[ʑì]	“four”
ཞྭམོ་ <zhwamo>	[ʑʰàm˥]	[ʑʰàm]	[ʐ̩ə̀m]	[ɕɛ̀m]	“hat”
ཤིང་ <shing>	[ɕĩː˩]	[ɕíŋ]	[ʃĩ́ː]	[ɕĩ́ŋ]	“tree, wood”
འོག་ལུ་ <’og lu>	?	?	[ɔ̀ lu]	[ɦòlé]	“under”
ཧིང་ <hing>	[hĩː˥]	[hĩ́ː]	[hĩ́ː]	[hĩ́ŋ]	“heart”

Table 25. Comparison of example words with various fricative onsets attested in existing literature and this study with written form and gloss.

5.2.3. Nasals

Mazaudon & Michailovsky 1988; Watters 2018) attests voiceless nasals in some dialects other than Thimphu that are realised as [h] in Thimphu. Mazaudon & Michailovsky (1988) attest six nasals, with the addition of /m̥/, /n̥/. Van Driem (1991) only attests the voiced nasals, while Watters (2018) attests five nasals, with the addition of /n̥/, the same ones as this study. Neither the 1988 nor the 1991 study provides any examples for the voiceless nasal onsets. In this study, all nasals are attested as voiced.

Written	Pitch	Mazaudon & Michailovsky (1988)	van Driem (1991)	Watters (2018)	This study
མ <m>	Low, High	[m]	[m]	[m]	[m]
ན <n>	Low, High	[n]	[n]	[n]	[n]
ལྷ <lh>	High	[n̥, h]	?	[n̥]	[n̥, h]
ང <ng>	Low, High	[ŋ]	[ŋ]	[ŋ]	[ŋ]
ཉ <ny>	Low, High	[ɲ]	[ɲ]	[ɲ]	[ɲ]

Table 26. Comparison of nasal onsets attested in existing literature and this study with written form, conditioned pitch and gloss.

Written	Mazaudon & Michailovsky (1988)	Van Driem (1991)	Watters (2018)	This study	Gloss
སྨན་ <sman>	[mén]	[mén]	[mén]	[mə́n]	“medicine”
གནགཔོ་ <gnagpo>	[napˀ]	[náp]	[nɜ́p]	[nép]	“black”
ལྷ་པ་ <lha pa>	?	?	[n̥ápa, hápa]	[n̥épé, l̥épé, hépé]	“nose.INFRM”
ལྔ་ <lnga>	[ŋá]	[ŋá]	[ŋá]	[ŋé]	“five”
ཉ་ <nya>	[ɲà]	[ɲà]	[ɲà]	[ɲà]	“fish”

Table 27. Comparison of example words with various nasal onsets attested in existing literature and this study with written form and gloss.

5.2.4. Approximants

Mazaudon & Michailovsky (1988) attest the same approximants as this study, while two other sources (van Driem 1991; Watters 2018) attest the addition of a voiceless rhotic /r̥/. None of the example words containing the voiceless rhotic were attested in the 1988 and this study. The rhotic consonants are attested as taps in Watters (2018) but as an approximant in this study.

Written	Pitch	Mazaudon & Michailovsky (1988)	Van Driem (1991)	Watters (2018)	This study
ལ <l>	Low, High	[l-]	[l-]	[l-]	[l-]
ལྷ <lh>	High	[l̥-, h-]	[l̥-, h-]	[l̥-, h-]	[l̥-, h-]
ཝ <w>, བ <b>	Low, High	[w-]	[w-]	[w-]	[w-]
ར <r>	Low, High	[r-]	[r-]	[ɾ-]	[ɹ-]
ཧྲ <hr>	High	?	[r̥-]	[ɾ̥-]	?
ཡ་ <y>	Low, High	[j-]	[j-]	[j-]	[j-]

Table 28. Comparison of approximant onsets attested in existing literature and this study with written form, conditioned pitch and gloss.

Written	Mazaudon & Michailovsky (1988)	Van Driem (1991)	Watters (2018)	This study	Gloss
རླུང་ <rlung>	[lũː˥]	[lúŋ]	[lṹː]	[lṹŋ]	“air, wind”
ལྷ་ <lha>	?	[l̥á]	[l̥á]	[l̥é]	“deity”
དབང་ <dbang>	[wãː˥]	[wáŋ]	[wáŋ]	[wɐ́ŋ]	“empowerment ritual”
རྭོ་ <rawo>	[rauˬ]	[rào]	[ràu]	[rəw]	“horn”
ཧྲལ་ <hral>	?	[r̥é]	[r̥éː]	?	“tear (apart)”
གཡག་ <gayag>	[jaː˥]	[jáː]	[jáːʔ]	[jé]	“yak”

Table 29. Comparison of example words with various approximant onsets attested in existing literature and this study with written form and gloss.

5.3. Phonotactics

All existing literature attests the heterorganic onset cluster series, written as <by>, <py> and <phy>. In Mazaudon & Michailovsky, they are attested as stop-stop clusters for the voiced series and stop-fricative clusters for the voiceless series. Van Driem (1991) does not provide articulatory descriptions, so the sounds are assumed to be stop-stop clusters. Watters (2018) attests the series as stop-affricate clusters.

Mazaudon & Michailovsky (1988) note that the clusters had merged with their non-cluster counterparts in Paro, but they remained contrastive in other parts of Bhutan, including Thimphu. This study finds that two onset series had merged as affricates, [t͡ɕ-] and [t͡ɕʰ-], for both of our language consultants, who were from Thimphu and Paro, respectively. However, they still pronounced the cluster onsets, as [pt͡ɕ-] and [pt͡ɕʰ-], during text-reading and careful enunciation.

Written	Pitch	Mazaudon & Michailovsky (1988)	Van Driem (1991)	Watters (2018)	This study
ཀྱ <by>	Low	[bʲ]	[bʲ]	[bd͡ʒ]	[t͡ɕ]
	Low	[bʲʰ]	[bʲʰ]	[bd͡ʒʰ]	
པྱ <py>	High	[pɕ]	[pc]	[pt͡ʃ]	
ཕྱ <phy>	High	[pɕʰ]	[pcʰ]	[pt͡ʃʰ]	[t͡ɕʰ]

Table 30. Comparison of onset clusters attested in existing literature and this study with written form, conditioned pitch and gloss.

Written	Mazaudon & Michailovsky (1988)	Van Driem (1991)	Watters (2018)	This study	Gloss
གནམ་བྱ་ <gnam bya’>	[namꜜ bɟaːꜛ]	[námbɟàː]	[nɜ́m bd͡ʒaː]	[némt͡ɕà]	“summer”
བྱམོ་ <byamo>	[bɟʰamꜛ]	[bɟʰàm]	[bd̥ʲd̥ʒ̊ɜ̀m]	[t͡ɕʰèm]	“hen”
པྱིང་ <pying>	[pɕín]	[pcíŋ]	[pt͡ʃíŋ]	[t͡ɕʰíŋ]	“glue”
ཕྱག་མ་ <phyagma>	[pɕʰaːmꜜ]	[pcʰáːm]	[ptʃʰáːm]	[t͡ɕʰém]	“broom”

Table 31. Comparison of example words with various approximant onsets attested in existing literature and this study with written form and gloss.

5.4. Tone

The relation between historical onset VOT and pitch is consistent with languages in the Sinospheric Tonbund (Matisoff 2001), the linguistic area categorised by the contact-induced cross-linguistic transphonologization of onset voicing contrasts as tones in East and Southeast Asia over the past millennium (Dockum & Gehrmann 2021). Onset voicing quality had first been identified as a conditioning environment for tonogenesis in this region by Karlgren (1915), who focused on Sinitic languages. Chang (1947) described the same process for tonogenesis in Hmong-Mien languages, Li (1943) in Tai languages, and Haudricourt (1946, 1954) in Karenic and Vietic languages.

Tonal contrasts in Middle Chinese, proto-Hmong-Mien, proto-Tai, proto-Karenic and proto-Vietic emerged in two distinct stages. During the first stage, before about 1000 CE, coda consonants conditioned three tonal contours, where [-ʔ] triggered rising pitch due to glottal constriction, [-h] and [-s] triggered falling pitch due to glottal relaxation, and [-∅] did not trigger any pitch variation and surfaced as level pitch. Subsequently, the codas were deleted. During the second stage, after about 1000 CE, onset consonants conditioned the splitting of the three tonal contours into two distinct registers, high and low, where voiced onsets lowered the starting pitch of syllables, while voiceless onsets did the opposite. The combination of three tonal contours and two registers surfaced as six contrastive tones. Subsequently, the onsets lost voicing contrast (Norman 1988).

Tibetic languages are unique in the linguistic area, as their proto language did not develop contrastive tones. Old Tibetan, spoken around the same period as the first stage of Sinospheric Tonogenesis, had a syllable structure of */(C_1C_2)C_3(G_1G_2)V(C_4C_5)/ (Hill 2010). Classical Tibetan, spoken around the same period as the second stage of Sinospheric Tonogenesis, had a simplified syllable structure of (*/(C_1C_2)C_3(G)V(C_4C_5)/ (Beyer 1992). Among the consonants, only */C3/, written as the root letter of a character in Tibetan script, had been contrastive in voicing. Over the course of historical Tibetan, the voicing of */C_3/ stayed contrastive, and codas remained. Therefore, neither contrastive pitch nor tonal contour appeared in the language.

As Tibetic languages diverged from their common ancestor, some developed contrastive tones, while others did not. Modern Tibetic languages range from having no tones in Amdo to as many as six in Khams (Tournadre 2014). Dzongkha had developed pitch contrast for its nasal and approximant onsets (Mazaudon & Michailovsky 1988; van Driem 1991; Watters 2018), where onset clusters involving nasals and approximants evolved into simple nasal and approximant onsets that condition a high pitch. This study demonstrates that the same transphonologisation of onset consonant features as vowel pitch has expanded to encompass all but one (/h/) voiceless onset as well. Dzongkha now has a similar tonal system as Lhasa Tibetan, which has two contrastive pitches and a two-way phonation contrast for its onset stop, affricate and fricative series. However, its high and low pitches are fully contrastive, including for syllables with aspirated onsets. In Lhasa, minimal pairs between aspirated high and aspirated low syllables are attested, likely due to the fact that historical voiced aspirated onsets had become aspirated low syllables, rather than the unaspirated low syllables found in Dzongkha.

On the other hand, Dzongkha appears to have lost tonal contour contrast over the same period as its pitch contrast expanded. Mazaudon & Michailovsky (1988) attest four tones for some coda consonants that comprise combinations of pitches and tonal contours: high level, high falling, low level, low falling. Specifically, flat words are observed to be monosyllabic in Old Tibetan, and falling words are historically multisyllabic. Van Driem (1991) attests two tonal contours as well, rising and slight falling, for some dialects of Dzongkha, corresponding to the flat and falling tones attested in Mazaudon & Michailovsky (1988), respectively. In Watters (2018) and this study, the high minimal pairs given in the former studies are attested with no contour specification, while the low minimal pairs are not attested. The formerly high level words maintained their morphology, while the formerly high falling words are now multisyllabic with different second-syllable morphemes. The data suggests that the tonal contour had become no longer contrastive at the phonemic level, and that the lexicon had adjusted to this change in the tonal paradigm to avoid ambiguity.

Mazaudon & Michailovsky (1988)	**Van Driem (1991)**	**Watters (2018); This study**
High Flat	High Rising	High
High Falling	High Slight-Falling	
Low Flat	Low Rising	Low
Low Falling	Low Slight-Falling	

Table 32. Comparison between tones attested in existing literature and this study.

Tone	**Mazaudon & Michailovsky (1988)**	**Van Driem (1991)**	**Watters (2018)**	**This Study**	**Gloss**
High Flat / High Rising / High	ཤོབ་ <shob> [ɕop˦] “lie”	ཤོབ་ <shob> [ɕop˦] “lie”	ཤོབ་ <shob> [ɕóp] “lie”	ཤོབ་ <shob> [ɕóp] “lie”	“lie”
	གསུམ་ <gsum> [sum˦] “three”	གསུམ་ <gsum> [sum˦] “three”	གསུམ་ <gsum> [súm] “three”	གསུམ་ <gsum> [súm] “three”	“three”
High Falling / High Slight-Rising / High	གཤོག་པ་ <gshog pa> [ɕop˧˨] “wing-NMZ”	གཤོག་པ་ <gshog pa> [ɕop˦] “wing-NMZ”	གཤོག་མཐིལ་ <gshog mthil> [ɕó thɛː] “wing bottom”	གཤོག་མཐིལ་ <gshog mthil> [ɕó tʰḛ̂] “wing bottom”	“wing”
	སྲུང་བ་ <srung ba> [sum˧˨] “guard-NMZ”	སྲུང་བ་ <srung ba> [sum˦] “guard-NMZ”	སྲུང་ཀི་ <srung ki> [sű́ːki] “guard NMZ”	སྲུང་ཀི་ <srung ki> [sű́ːkʰí] “guard NMZ”	“charm, amulet”

Table 33. Comparison between tone data attested in existing literature and this study, including example words and their written form, transcription and gloss.

Lexicon change as a result of phonological change is well documented across genealogically and geographically distant languages. In Modern Standard Chinese, the Sino-Tibetan language with the most native speakers, 3% of words are one-syllable words, 64% are two-syllable words, 18% are three-syllable words, and the remainder are four- or more syllable words (Cao, Yu & Dong 2021). The average syllable length of Sinitic languages increased from closer to one in Old Chinese to more than two in modern varieties. At the same time, they gained tonal contrasts, and their consonant inventories, especially for coda consonants, shrank. Modern Sinitic languages that retained more complex syllable structures and tonal contrasts, such as Min and Yue, have shorter prosodic words. Duanmu (2007) attributes the lengthening of Chinese prosodic words to the combination of a number of reasons, namely for homophone avoidance, the subsequent change in metrical structure and the introduction of new disyllabic vocabulary. In Southern US English where the pin-pen merger is common, *ink pen* is routinely used instead of *pen* to avoid homophonic ambiguity with *pin* (Brown 1991).

a. *pin* /pʰɪn/ [pʰɪn]
b. *pen* /pʰɛn/ [pʰɪn] → *ink pen* [ɪnkʰpʰɪn]

Comparing data from Mazaudon & Michailovsky (1988) to Watters (2018) and this study, a similar process may have occurred in Dzongkha. Due to the loss of the tonal contour contrast, homophones emerged, and speakers may have begun to use different synonymous words to avoid homophonic ambiguity.

6. Conclusion

This thesis delved into the evolving phonetics and phonology of Dzongkha, a Tibetic language spoken in Bhutan, with a particular focus on the transformation of its consonant inventory and tonal system over the past three decades. The study, motivated by noticeable changes possibly influenced by language policies and migration, aimed to provide an updated analysis of the language's sound system based on newly collected field data.

The vowel analysis revealed five distinct vowels with variations in frontness and height. The consonant inventory comprised thirty phonemes with various manners of articulation, exhibiting differences in syllable onset and coda positions. The phonotactic structure of Dzongkha syllables was identified as /C1V(C2)/, with specific consonants more prevalent at the syllable coda.

The tonal system exhibited a contrast between high and low pitches, where low vowels did not appear following an aspirated onset, including /h/. The discussion section provided comparisons with existing literature, highlighting changes in vowel phonemes, consonant voicing patterns, and the merging of heterorganic onset clusters.

Notably, the study identified a shift in the tonal system linked to historical onset voicing, drawing parallels with languages in the Sinospheric Tonbund. The findings suggested a transphonologisation process similar to the Sinospheric Tonogenesis, influenced by factors such as language contact and migration to urban centers.

More investigation is necessary to get a more complete and accurate perspective on the sound system of Dzongkha. Since only two consultants participated in the study, the results cannot represent the entire Dzongkha speech community. Therefore, more elicitation sessions should be conducted to gather data from more speakers that differ in age, gender, social class, and language experience to reveal the sociolinguistic and diachronic aspects of the language that are not covered in this study. Further documentation on the evolving language can prove beneficial for the theoretical understanding of tonogenesis, areal linguistics in the Himalayas, as well as educational and revitalisation work.

7. Appendix: Orthography

Dzongkha is typically printed and displayed in Uchen “with head”, an abugida inherited from Old Tibetan that is also used to write other modern Tibetic languages. Dzongkha is usually handwritten in Joyig “speedy”, which resembles speedily written Uchen and is primarily used in Bhutan. Both scripts are read from left to right with word separators (') and sentence separators (|). Each character represents a syllable, where both onset and coda consonants can be visually stacked to represent clusters. Vowels are written as diacritics above or below the character.

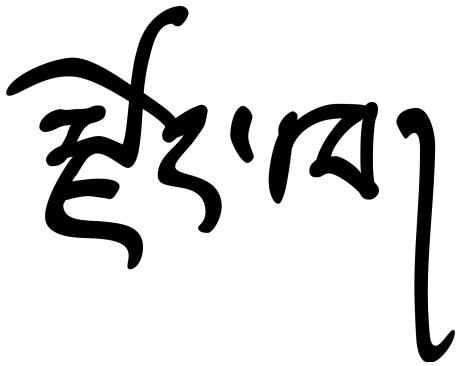

Figure 10. རྫོང་ཁ <rdzong kha> handwritten in Joyig script (Fynn 2019.)

The earliest archaeological evidence of Uchen is from the 8th century, likely based on the indic-Brahmic scripts of the time from the Indian subcontinent. Other Tibetic scripts have emerged since then, but Uchen remains the most popular for printing.

The orthography of Tibetic languages became mostly fixed in the 11th century, so it does not reflect modern pronunciation. For example, /ʔù/ “hair.FRM” is spelled <dbu>. The cultural prestige of the very conservative orthography leads to instances of spelling-pronunciation. For instance, [ɲìkì ~ ɲì] 1.SG.GEN “my” is much more common in speech, but [ŋèkì] is preferred instead when reading text, because the word is spelled <ŋa gi> in writing. As a result, the latter pronunciation has a more formal and careful connotation than the former.

Wylie transliteration, created in 1959, is a method for transliterating Tibetic script using ASCII letters. The system aims to transcribe the text exactly as written, instead of reflecting actual pronunciation. Therefore, since all languages that diverged from Classical Tibetan since the 11th century are spelled nearly identically, they are also transliterated nearly identically in Wylie.

U.	W.	IPA	U.	W.	IPA	U.	W.	IPA	U.	W.	IPA
ཀ	<ka>	[ké]	ཁ	<kha>	[kʰé]	ག	<ga>	[kè]	ང	<nga>	[ŋè]
ཅ	<ca>	[tɕé]	ཆ	<cha>	[tɕʰé]	ཇ	<ja>	[tɕè]	ཉ	<nya>	[ɲè]
ཏ	<ta>	[té]	ཐ	<tha>	[tʰé]	ད	<da>	[tè]	ན	<na>	[nè]
པ	<pa>	[pé]	ཕ	<pha>	[pʰé]	བ	<ba>	[pè]	མ	<ma>	[mè]
ཙ	<tsa>	[tsé]	ཚ	<tsha>	[tsʰé]	ཛ	<dza>	[tsè]	ཝ	<wa>	[wè]
ཞ	<zha>	[ʑè]	ཟ	<za>	[zè]	འ	<'a>	[ɦè]	ཡ	<ya>	[jè]
ར	<ra>	[rè]	ལ	<la>	[lè]	ཤ	<sha>	[ɕé]	ས	<sa>	[sé]
ཧ	<ha>	[hé]	ཨ	<a>	[ʔa]						

Table 34. The Tibetic alphabet Uchen, Wylie transliteration and IPA transcription. The stop and affricate series are ordered by place of articulation from back to front in the alphabet. [t̪], [t̪ʰ] and [l̪], which are not included in the alphabet, are represented as the digraphs of <kr> and <khr> in both Uchen and Wylie.

8. References

Andvik, E. (2010). *A grammar of Tshangla (Vol. 5).* Brill.

Babbage (2010). “Languages of Bhutan with labels.” The Wikimedia Foundation.

Beyer, S. V. (1992). *The Classical Tibetan Language.* State University of New York Press.

Brown, V. R. (1991, Autumn). *Evolution of the Merger of /l/ and /ɛ/ before Nasals in Tennessee. American Speech, Vol. 66,* No. 3, pp. 303-315.

Cao, Y., Yu, G. & Dong, L. (2021). Chinese Lexicon. In *Teachers Talking about their Classrooms* (pp. 99-114). Routledge.

Chen, X., Jin H. & Yu H. (2010). “Acoustic Research on Long and Short Vowels in Tibetan Lhasa Dialect.” Key Lab of National Linguistic Information Technology, Northwest University for Nationalities.

Dockum, R. & Gehrmann, R. (2021, January). “The East Asian Voicing Shift and its Role in the Origins of Tone and Register.” In Proceedings of the 95th Annual Meeting of the Linguistic Society of America, San Francisco, CA, USA.

Dorjee, K. (2014). “Linguistic landscape of Bhutan: An overview of number of languages, language policy, language education, and language use in Bhutan.” *Bhutan Journal of Research & Development*, *3*(1), 79-101.

Duanmu, S. (2007). *The Phonology of Standard Chinese.* OUP Oxford.

Everett, C., Blasi, D. E., & Roberts, S. G. (2015). Climate, vocal folds, and tonal languages: Connecting the physiological and geographic dots. *Proceedings of the National Academy of Sciences, 112(5),* 1322-1327.

Fynn, C. J. (2019). “"Dzongkha" / རྫོང་ཁ་ The name for the national language of Bhutan, Dzongkha, written in Bhutanese (joyig) script.” The Wikimedia Foundation.

Gutman, A. & Avanzati B. (2014). *Languages and ethnic Groups of Bhutan.* The Language Gulper.

Hill, N. W. (2010). An overview of Old Tibetan synchronic phonology. *Transactions of the philological society, 108*(2), 110-125.

Hyslop, G. (2011). A grammar of Kurtöp. University of Oregon.

Kirby, J. & Hyslop, G. (2012). “Acoustic Analysis of Onset Voicing in Dzongkha Obstruents.” University of Edinburgh and University of Sydney.

Michailovsky, B. & Mazaudon, M. (1988). Lost syllables and tone contour in Dzongkha (Bhutan).

Moseley, C. (Ed.). (2010). *Atlas of the World's Languages in Danger.* Unesco.

Norman, J. (1988). *Chinese.* Cambridge University Press.

NSB, Bhutan (2022). *Bhutan Living Standards Survey Report.*

Tournadre, N. (2005). “L'aire linguistique tibétaine et ses divers dialectes.” *Lalies,* 2005, n°25, p. 7–56.

Tournadre, N. (2014). “The Tibetic languages and their classification.” *Trans-Himalayan Linguistics.*

van Driem, G. (1991). *The Grammar of Dzongkha.* Thimphu, Bhutan: RGoB, Dzongkha Development Commission (DDC).

Watters, S. A. (1996). “A preliminary study of prosody in Dzongkha.” University of Texas at Arlington. ProQuest Dissertations Publishing. 1383506.

Watters, S. A. (2018). “A grammar of Dzongkha.” Rice University.

Wylie, T. (1959). “A Standard System of Tibetan Transcription.” *Harvard Journal of Asiatic Studies,* p. 261-267.

9. Bibliography

Boersma, P. & Weenink, D. “Praat: doing phonetics by computer.” https://www.fon.hum.uva.nl/praat/

Bommarito, N. “The Online English-Tibetan Dictionary.” http://eng-tib.com

Dharma Dictionary, The. “Tibetan-English Dictionary, Dharma Glossaries and Resources.” https://rywiki.tsadra.org

Lotsawa House. “Tibetan transliteration: convert between Wylie and Unicode.” https://www.lotsawahouse.org/Cgi/wylie.pl

Nègre, X. “Tibetan Dictionary.” *Lexilogos.* https://www.lexilogos.com/english/tibetan_dictionary.htm

Steinert, C. “Tibetan-English Dictionary.” https://dictionary.christian-steinert.de

Szynalski, T. P. “Type IPA phonetic symbols for all languages.” https://ipa.typeit.org/full

Tibetan & Himalayan Library, The. “Tibetan Historical Dictionary.” https://www.thlib.org/reference/dictionaries/tibetan-dictionary

Visit Bhutan. “Dzongkha English dictionary online.” https://www.bhutan.me/online/dzongkha-english-online