



Development of ManiTo: A Manipuri Tonal Contrast Dataset

Thiyam Susma Devi^(✉)  and Pradip K. Das 

Indian Institute of Technology Guwahati, North Guwahati, Assam, India
{thiyam.devi, pkdas}@iitg.ac.in

Abstract. Tone recognition plays a vital role in understanding speech for tonal languages. Integrating tonal information from a robust tone recognition system can improve the performance of Automatic Speech Recognition (ASR) for such languages. The tonal recognition approaches adopted so far have focused on Asian, African and Indo-European languages. In India, there are very few works on tonal languages, especially those spoken in its North-Eastern part, from which the Manipuri language is largely unexplored. This paper presents the development of a Tonal Contrast dataset for Manipuri, a low resource language. It also presents an initial analysis of the recorded data.

Keywords: Tonal contrast · Speech recognition · Low resource languages

1 Introduction

Humans use a system of communication called speech using a language, which consists of sounds, words and grammar. English, Hindi, French and most European languages' words comprise of a sequence of distinctive units known as phonemes. However, several languages in the world are tonal, as Yip [21] points out. Tonal languages use tones to determine the meaning of the speech units.

As these languages are spoken by limited people, many languages have become extinct. Here, technology plays a crucial role in stopping the extinction by providing the techniques of Natural Language Processing (NLP) and the Automatic Speech Recognition (ASR) for these languages of the world. ASR is also needed for fostering economic growth and prosperity. For more flexibility and to have human-machine interaction, ASR system for tonal languages is required. Several tone recognition techniques have been developed for various tonal languages [20] such as Mandarin, Thai, Vietnamese, Punjabi, Yoruba, etc. However, it is found that no work has been done on automatic tone recognition of the Manipuri tonal language.

1.1 Manipuri

Manipuri, also known as Meiteilon/Meiteiron, is one of the scheduled Indian Tibeto-Burman language spoken predominantly in Manipur, a northeastern state

of India. Some people of Indian states, such as Assam, Mizoram and Tripura, and other countries like Bangladesh and Myanmar, also speak Manipuri. It is the official language of Manipur, which is spoken by over 1.5 million speakers. In Manipur, among 29 different ethnic groups, Manipuri is the only medium of communication [13]. Manipuri is a tonal language in which the tone distinguishes the meaning of words. For speech recognition and pronunciation evaluation, the identification of tone in Manipuri is essential.

Manipuri has its own script, which is known as Meitei/Meetei Mayek script. The Meitei Mayek Script has 27 Mapung Mayek (main alphabets). There are 8 Lonsum Mayek (unreleased characters), 8 Cheitap Mayek (vowel signs), 3 Khudam Mayek (punctuation marks including diacritics) and Cheising Mayek for the numerals [3, 11].

1.2 Tones in Manipuri

Using pitch in a language to distinguish lexical or grammatical meaning is known as tone [2]. As mentioned before, Manipuri is a tone language [17]. It has a lexically significant, contrastive, but relative pitch on each syllable. There are two tones in Manipuri [6, 9, 11, 15, 19]:

1. A level tone: unmarked
2. A falling tone: marked by lum mayek, “.”

Every syllable in Manipuri carries one of the two tones. The pitch (frequency) of level tone is lower than the pitch of the falling tone; thus, some authors (e.g., Chelliah 1997 [5]) have termed the level and falling tones low and high, respectively [19]. The level tone is unmarked while the falling tone marked as /˘/ in English representation. Furthermore, the lum mayek or the falling tone mark, “.” is represented in Manipuri script just after the syllable, which carries the falling tone.

2 Related Works

In the international scenario, intensive research is done in tonal language speech recognition in the last three decades. Peng et al. (2021) [14] proposed a Multi-Scale model that gathers the information at multiple resolutions capturing the attributes of tone variation. The experiment is performed on the dataset, Chinese National Hi-Tech Project 863. Their model achieves tone error rate (TER) of 10.5%. Hao et al. (2019) proposed a framework based on deep neural networks for Mandarin tone recognition. The model uses both the prosodic and the articulatory features as the raw input data. A 5-layer-deep belief network is employed to generate high-level tone feature. The 863-data corpus is used for the experiment and achieved an average tone recognition rate of 83.03% accuracy. Nguyen et al. (2016) [16] investigated the effect of tone in the Vietnamese Large Vocabulary Continuous Speech Recognition System and built an acoustic model using the

tonal feature. The experimental result obtained 19.25% improvement over the non-tonal phoneme system.

In India for the language Manipuri, Thoudam (1980) [15] doctoral thesis has devoted a chapter on Manipuri phonology. He suggested that there were only two distinctive tones in Manipuri, namely, falling tone and level tone. Mahabir (1982) [4] argued for two tones, falling and level in his master’s thesis. Chelliah (1990) [18] studied the level ordered morphology and phonology in Manipuri and presented several phonological rules. Chelliah (1997) [5] explained the tone system in Manipuri. She presented a framework that correctly described that Manipuri exhibits a two-way tonal contrast, low tone and a default high tone. The fundamental frequency contours were used as the phonetic representations of the underlying tone pattern in the experiment. Meiraba (2014) [12] claimed that the tone bearing unit in Manipuri is the Rhyme of the syllable. The relative simplicity of the tone system of Manipuri is due to its rich consonantal inventory which can occur at the Coda position and that the realisation of tonal contrast can be affected by the Coda consonants.

3 Motivation

After exhaustive search it is found that there are limited tonal languages (Mizo, Punjabi, Singpho, Manipuri, etc.) in India and virtually no datasets are available for tonal analysis. It is also evident that there is a critical need to develop speech dataset of tonal contrast pairs to study the characteristic of the tonal variation leading to the understanding of distinct words for the language. This motivates us to develop a tonal contrast word pair for the Manipuri language and study the tone information present in it for developing robust ASR systems for Manipuri.

4 Creation of Tonal Contrast Word Pair Corpus

Fifty pairs of Manipuri tonal contrast words are collected from different sources [6, 8, 12, 15, 19]. The words are listed below in Fig. 2 with their respective meanings.

The data is collected from six people: three males and three females, age range of the speakers is from 21 to 45. All of them are native speakers, out of which three of them (two male and one female) are working in the Linguistic Department of Manipur University, Imphal and their recording is done in the Audio, Visual, Language and Phonetic Laboratory Complex of Manipur University. The remaining three native local speakers’ recording is done in a quite office environment. A total of 50 tonal contrast words, five instances of each pair with some pause between the speech sounds, are recorded separately for each person. The steps of creating the dataset is shown in Fig. 1. The Cool Edit 2000 tool is used for recording the utterances. While recording, the following three parameters have been set in Cool Edit 2000.

Sampling Rate: It is the number of samples per second to be captured by the microphone into the system. Sampling rate is set to 44,100 Hz.

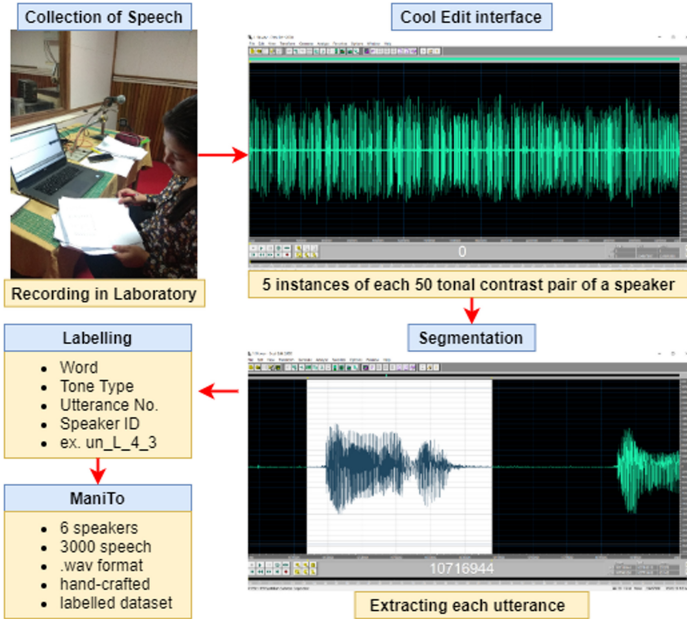


Fig. 1. Creation of ManiTo dataset.

Channel: Mono channel is selected. In mono, all audio signals are routed through a single audio channel.

Resolution: Each sample is represented using 16 bits.

4.1 Preprocessing

The recorded speech sounds are further analyzed and segmented manually, with about 1000 samples of silence at the beginning and end of each word and saved in a .wav format, where each wav file has been named by using word name, tone detail ‘f’ for falling and ‘L’ for level, instance number and speaker ID.

For example, un_f_2_1.wav

Word: un Tone: falling Instance: 2 Speaker ID: 1

The corpus, ManiTo consists of 3,000 hand-crafted labeled speech data of size 273 MB. The recordings are carefully double checked and stored.

5 Experimental Analysis

Praat [1] is a tool that can analyze, synthesize, and manipulate speech data. Praat version 6.1.51 is used for the experiment. From the developed dataset, the speech sample are analyzed using Praat. In tone analysis, features that reflect the pitch contour are lexically significant. The fundamental frequency, F0, acts as an indication for tone. For the preliminary study on ManiTo dataset, the

Sl. No.	Falling Tone	Meaning	Level Tone	Meaning	Sl. No.	Falling Tone	Meaning	Level Tone	Meaning
1	ꯁꯩꯃ /un/	skin	ꯁꯩꯃ /un/	ice, snow	26	ꯩꯃꯩ /lúbə/	clear	ꯩꯃꯩ /lúbə/	difficult
2	ꯩꯃꯩ /in/	push	ꯩꯃꯩ /in/	follow	27	ꯩꯃꯩ /sábə/	to make	ꯩꯃꯩ /sábə/	hot
3	ꯩꯃꯩꯃ /t ^h oŋ/	door	ꯩꯃꯩꯃ /t ^h oŋ/	bridge	28	ꯩꯃꯩꯃ /laubə/	to shout	ꯩꯃꯩꯃ /laubə/	large hole
4	ꯩꯃꯩ /mi/	man	ꯩꯃꯩ /mi/	spider	29	ꯩꯃꯩꯃ /kawbə/	to kick	ꯩꯃꯩꯃ /kawbə/	to forget
5	ꯩꯃꯩꯃ /púbə/	to borrow	ꯩꯃꯩꯃ /pubə/	to bring	30	ꯩꯃꯩꯃ /haibə/	to say	ꯩꯃꯩꯃ /haibə/	to swing
6	ꯩꯃꯩ /cá/	wax	ꯩꯃꯩ /ca/	tea	31	ꯩꯃꯩꯃ /kəubə/	short	ꯩꯃꯩꯃ /kəubə/	to call
7	ꯩꯃꯩ /i/	blood	ꯩꯃꯩ /i/	thatch	32	ꯩꯃꯩꯃ /lɔŋ/	spear	ꯩꯃꯩꯃ /lɔŋ/	fishing net
8	ꯩꯃꯩꯃ /k ^h oi/	navel	ꯩꯃꯩꯃ /k ^h oi/	bee/ fishing hook	33	ꯩꯃꯩꯃ /pábə/	thin	ꯩꯃꯩꯃ /pábə/	to read
9	ꯩꯃꯩ /lá/	wide basket	ꯩꯃꯩ /la/	banana leaf	34	ꯩꯃꯩꯃ /túbə/	tailoring	ꯩꯃꯩꯃ /túbə/	fall
10	ꯩꯃꯩꯃ /siŋ/	firewood	ꯩꯃꯩꯃ /siŋ/	ginger	35	ꯩꯃꯩꯃ /sibə/	to break down	ꯩꯃꯩꯃ /sibə/	death
11	ꯩꯃꯩꯃ /səŋ/	hair	ꯩꯃꯩꯃ /səŋ/	basket	36	ꯩꯃꯩꯃ /soibə/	mistake	ꯩꯃꯩꯃ /soibə/	cut
12	ꯩꯃꯩꯃ /məŋ/	dream	ꯩꯃꯩꯃ /məŋ/	grave	37	ꯩꯃꯩꯃ /lábə/	roam	ꯩꯃꯩꯃ /lábə/	male
13	ꯩꯃꯩꯃ /kəŋ/	mosquito	ꯩꯃꯩꯃ /kəŋ/	manipuri game	38	ꯩꯃꯩꯃ /lanbə/	to cross	ꯩꯃꯩꯃ /lanbə/	wrong
14	ꯩꯃꯩꯃ /tɪn/	worm	ꯩꯃꯩꯃ /tɪn/	saliva	39	ꯩꯃꯩꯃ /koibə/	to travel	ꯩꯃꯩꯃ /koibə/	not shortcut
15	ꯩꯃꯩꯃ /pi/	give	ꯩꯃꯩꯃ /pi/	tear	40	ꯩꯃꯩꯃ /lumbə/	heavy	ꯩꯃꯩꯃ /lumbə/	warm
16	ꯩꯃꯩꯃ /k ^h oŋ/	leg	ꯩꯃꯩꯃ /k ^h oŋ/	canal	41	ꯩꯃꯩꯃ /sɔnbə/	weak	ꯩꯃꯩꯃ /sɔnbə/	to chant
17	ꯩꯃꯩꯃ /kənbə/	hard	ꯩꯃꯩꯃ /kənbə/	to protect	42	ꯩꯃꯩꯃ /c ^h əniŋ/	weight measuring unit	ꯩꯃꯩꯃ /c ^h əniŋ/	lower abdomen
18	ꯩꯃꯩꯃ /tábə/	to hear	ꯩꯃꯩꯃ /tabə/	to fall	43	ꯩꯃꯩꯃ /subə/	wash	ꯩꯃꯩꯃ /subə/	hammering
19	ꯩꯃꯩꯃ /cábə/	eating	ꯩꯃꯩꯃ /cabə/	to hang around	44	ꯩꯃꯩꯃ /múbə/	roast	ꯩꯃꯩꯃ /múbə/	black
20	ꯩꯃꯩꯃ /kəmbə/	killing	ꯩꯃꯩꯃ /kəmbə/	over ripen	45	ꯩꯃꯩꯃ /paibə/	touch	ꯩꯃꯩꯃ /paibə/	fly
21	ꯩꯃꯩꯃ /leŋbə/	movement	ꯩꯃꯩꯃ /leŋbə/	making rows	46	ꯩꯃꯩꯃ /puŋ/	drum	ꯩꯃꯩꯃ /puŋ/	hour
22	ꯩꯃꯩꯃ /asəŋbə/	to dye	ꯩꯃꯩꯃ /asəŋbə/	green	47	ꯩꯃꯩꯃ /kəi/	granany	ꯩꯃꯩꯃ /kəi/	tiger
23	ꯩꯃꯩꯃ /kábə/	to climb	ꯩꯃꯩꯃ /kabə/	burnt	48	ꯩꯃꯩꯃ /hɪi/	dog	ꯩꯃꯩꯃ /hɪi/	rivet
24	ꯩꯃꯩꯃ /tɔŋbə/	raised platform	ꯩꯃꯩꯃ /tɔŋbə/	riding	49	ꯩꯃꯩꯃ /ləi/	flower	ꯩꯃꯩꯃ /ləi/	tongue
25	ꯩꯃꯩꯃ /tumbə/	to sleep	ꯩꯃꯩꯃ /tumbə/	diluted	50	ꯩꯃꯩꯃ /pə/	grimace	ꯩꯃꯩꯃ /pə/	umbrella

Fig. 2. List of Manipuri tonal contrast word Pairs with their respective meanings.

pitch or F0 is extracted using Praat. Praat use the most accurate pitch analysis algorithm [7]. Figure 3 shows the analysis of falling tone “un” sound. The blue line is the pitch listing of the speech. Similarly, Fig. 4 shows the analysis of level tone “un” sound. From the two figures we can notice that the pitch of the level tone is lower than that of falling tone. Figure 5 shows the graph comparing the five utterances tonal contrast Pair1, “un” spoken by Speaker1. Figure 5a plots the pitch listing of falling tone, Fig. 5b plots the pitch listing of level tone, Fig. 5c is the normalisation of falling tone, Fig. 5d is the normalisation of level tone and Fig. 5e shows the comparison of average pitch listing of falling versus level tone. From the graph we can initially infer that the pitch of the falling tone is higher than that of level tone. Using parselmouth [10], a python library for the Praat software, mean F0, harmonics-to-noise-ratio(HNR), jitter, shimmer information are extracted and analysis is being conducted on this features to distinguish the tones accurately.

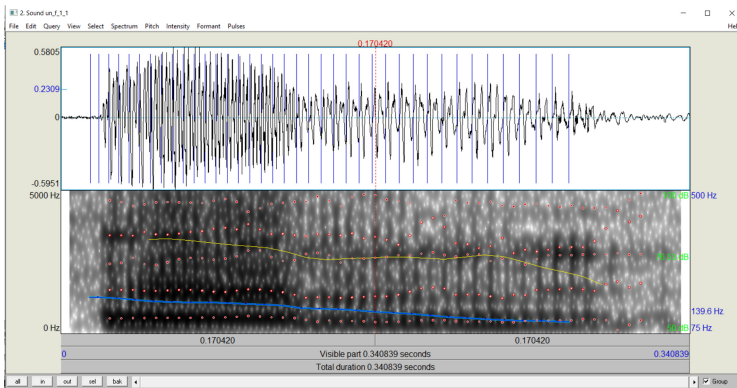


Fig. 3. Waveform and spectrogram with overlaid pitch contour of falling tone “un” sound.

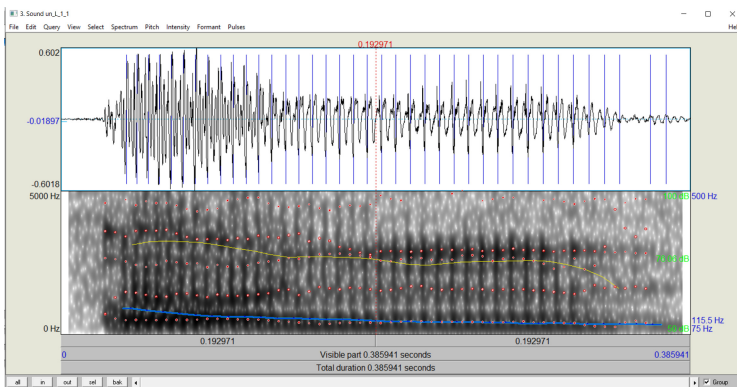


Fig. 4. Waveform and spectrogram with overlaid pitch contour of level tone “un” sound.

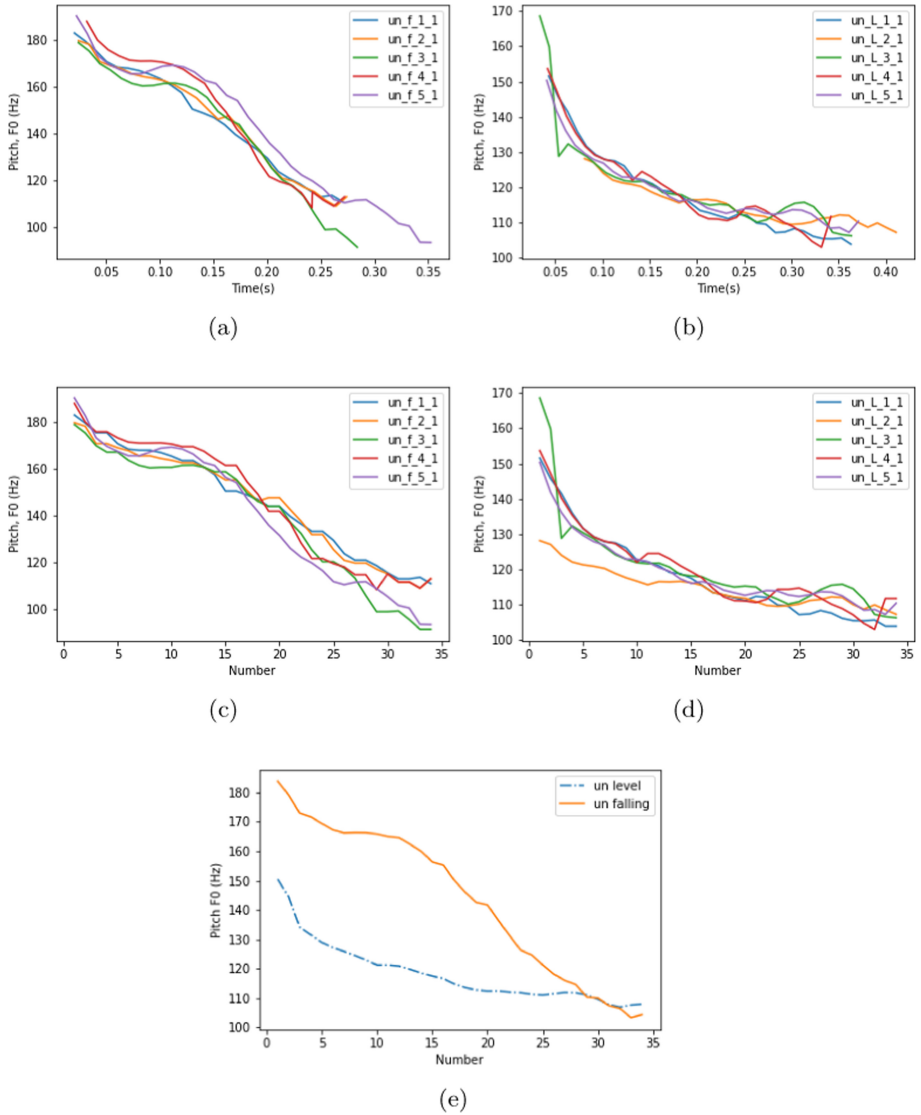


Fig. 5. Pitch contour of (a) “un” five utterances (falling tone) (b) “un” five utterances (level tone) (c) normalise pitch “un” (falling) (d) normalise pitch “un” (level) (e) average pitch comparison of level and falling tone.

6 Conclusion and Future Work

A speech dataset containing tonal contrast pair of the Manipuri language is being created. ManiTo containing 3,000 samples of Manipuri tonal contrast words is developed from data collected from 6 speakers. Fundamental analysis of the

dataset is currently being done. It is found that the pitch of the falling tone word is higher than the level tone word. The pitch value can be used to distinguish the tones in Manipuri. Further analysis on feature selection is currently being done to accurately differentiate the tones and develop a robust model for tone recognition for the Manipuri language.

Acknowledgments. We thank Prof. Chungkham Yashawanta Singh, Dr. Yumnam Aboy Singh, Nameirakpam Amit and Laishram Niranjana from Linguistics Department of Manipur University, Imphal for their advice and support in creating the ManiTo dataset.

References

1. Praat. <https://www.fon.hum.uva.nl/praat/>
2. Tone linguistics. [https://en.wikipedia.org/wiki/Tone_\(linguistics\)](https://en.wikipedia.org/wiki/Tone_(linguistics))
3. Meetei Mayek Tamnaba Mapi Lairik, Textbook. Global Publications (2017)
4. Mahabir L.: A contribution to the study of tone in Manipuri. In: Master's thesis, Deccan College Postgraduate and Research Institute, Pune (1982)
5. Chelliah, S.L.: Tone in Manipuri (1997)
6. Khan, A.G.: A Contrastive Study of Manipuri (Meiteilon) And English Phonology. In: Thesis of Doctor of Philosophy, Guwahati University (1987)
7. Boersma, P., Van Heuven, V.: Speak and UnSpeak with PRAAT. *Glott. Int.* **5**, 341–347 (2001)
8. Singh, C.Y.: Manipuri Grammar, 2nd edn. Textbook, Rajesh Publications, New Delhi (2019)
9. Devi, H.S.: Loanwords in Manipuri and their impact. *Linguist. Tibeto-Burman Area* **27**(1), 29–60 (2004)
10. Jadoul, Y., Thompson, B., de Boer, B.: Introducing Parselmouth: a python interface to Praat. *J. Phonetics* **71**, 1–15 (2018). <https://doi.org/10.1016/j.wocn.2018.07.001>
11. Singh, L.S., Thaoroijam, K., Das, P.K.: Written Manipuri (Meiteiron) from phoneme to grapheme. *Lang. India* **7**(6), 2–22 (2007)
12. Takhellambam, M.: Tones in meiteilon: a phonetic description. *Lang. India* **14**(7), 440–460 (2014)
13. Haokip, P.: The languages of Manipur: a case study of the kuki-chin languages. *Linguist. Tibeto-Burman Area* **34**(1), 85–118 (2011)
14. Peng, L., Dai, W., Ke, D., Zhang, J.: Multi-scale model for mandarin tone recognition. In: 2021 12th International Symposium on Chinese Spoken Language Processing (ISCSLP), pp. 1–5 (2021). <https://doi.org/10.1109/ISCSLP49672.2021.9362063>
15. Thoudam, P.C.: A grammatical sketch of Meiteiron. In: Thesis of Doctor of Philosophy, Jawaharlal Nehru University, New Delhi (1980)
16. Nguyen, Q.B., Vub, T.T., Luong, C.M.: The effect of tone modeling in Vietnamese LVCSR system. In: *Procedia Computer Science* (2016)
17. Rev. W. Pettigrew: Manipuri (Meitei) Grammar with Illustrative Sentences. The Pioneer Press, Allahabad (1912)
18. Chelliah, S.L.: Level ordered morphology and phonology in Manipuri. *Linguist. Tibeto-Burman Area* **13**(2), 27–72 (1990)

19. Chelliah, S.L.: A Grammar of Meitei. Mouton de Gruyter, New York (1997), ISBN: 3110143216, 9783110143218
20. Singh, A., Kadyan, V.: Automatic speech recognition system for tonal languages: state-of-the-art survey. Arch. Comput. Methods Eng. **28** (2020). <https://doi.org/10.1007/s11831-020-09414-4>
21. Yip, M.: Tone. Cambridge Textbooks in Linguistics, Cambridge University Press, Cambridge (2002). <https://books.google.co.in/books?id=KFv2lojXjpwC>, ISBN: 9780521774451