



# Digital Sanskrit to Meiteilon Dhātupāṭha Translation

**K. Kabi Khanganba**

Independent Researcher, Nambol, Maibam, Manipur - 795134  
New Delhi, 110067  
E-mail: khanganba.360@gmail.com

## Abstract

The paper targets to prepare a parallel digital lexicon of Panini's Dhātupāṭha (DP) for Sanskrit and Meiteilon/Manipuri Language. It has adopted three different editions of DP as the standard and primary source of preparing the DP lexicon. They are namely Dhāturūpanandini of Janardana Hegade, Prakriyānusārī Pāṇinīyadhātupāṭha of Puspa Dixita and Mādhavīyadhātuvṛtti edited by Dvarikadas. Basically, the presented lexicon is a manual translation of DP into Meiteilon from Sanskrit for two primary proposes: 1) For future applications in research like verb ontology, and Machine Translations etc. 2) For both online and hardcopy of Meiteilon translations of DP for learners. The data are stored in Unicode as its encoding standard for further usage and computational processing. Meiteilon is one of the scheduled Indian less resource languages. Translations from Sanskrit texts may be one of the major resources for creations of Meiteilon language resource. This work is one of the initiatives of facing the challenges of the machine translations of Sanskrit texts into Meiteilon.

**Keywords:** Dhātupāṭha, Dhātu, Meiteilon, Puya, Pratyaya, Anubandha

## 1. Introduction

Sanskrit represents knowledge store not only for other Indian languages and cultures but also for most of the major cultures, societies and communities over the globe. The huge amount of translations from the very beginning of ancient time in different languages including Tibetan, Chinese, Indonesian, English, French and German etc. enriches their literature and culture. That even brought new knowledge systems and disciplines. In that sense so far Sanskrit became a donor language and has proved itself the source of huge knowledge. Thus, translation has a great role in interchanging intellectual traditions among different cultures. It is an age-old practice of the human civilization. The history witnesses how the Indian intellectual traditions, went beyond through translation and inspired other cultures and traditions. The instances of Buddhist Sanskrit translations in China in early 2nd century and the journey of Veda, Upaniṣads, and Kāvya etc. in the West as later as from the 17th century are enough to support the role of translation in knowledge exchange in the history of human civilization. Now, translation has been a well-established discipline. It discusses the real challenges in translating different texts having different subjects, literary qualities, linguistic behaviors, social and historical contexts, etc.

Transferring the language resource and knowledge system through digital technology may be a very useful, an ambitious, and a new challenging task (Sudhir, 2007). Sanskrit still has the potency to be a donor language and source language in the contemporary context of less resource Indian languages like Meiteilon or Manipuri language.

The native speakers call Manipuri or the Manipuri language as Meiteilon (Meitei + lon /ma:ɪ̯ta:ilon/) or Meiteiron (Meitei + lon < ron /ma:ɪ̯ta:iron/). Meitei (/ma:ɪ̯ta:i/) is the name of the community and lon or lol (/lon/ or /lol/) means ‘language’. Some prefer Meetei (/mi:ɪ̯ta:i/) instead of Meitei. In their case it may be Meeteilon (/mi:ɪ̯ta:ilon/) or Meeteirol (/mi:ɪ̯ta:ilol/). The Manipuri is the official name documented in the constitution (71st amendment, 1992). This may be etymologized as derived after the name of the North Eastern state Manipur where, rather than the Meiteis, the language is widely used as a lingua franca by non-Meitei speakers like other linguistic communities, Nepalese, and businessmen from different parts of the country, too. Presently Meiteilon is the first tongue of the Meiteis and Meitei-Pangal (/pa:ŋəl/ means Muslims). Meithei is linguistic term for Meiteilon coined by the British (Grierson: 1967), preferably used often by non-native or foreign linguists. It is presently written (officially) in two scripts namely Bangla and Meitei Mayek or Meitei scripts (Khangamba & Jha, 2014). However, the present generation, especially the youngsters widely use the Roman scripts for writing on social networks. Meiteilon occupies the 7th rank in terms of number of speakers among the roughly 250 Tibeto-Burman languages (Matisoff: 2001). In a linguistic perspective, Meiteilon can be divided into Puyalon and Puya Theithalon. The primary basis of this distinction between them is their phonetics and prosody/accent/tone. The Puyalon is archaic Meiteilon which literature is mainly preserved in the Puyas and it has lesser number of sounds. Another unique feature of Puyalon is its musical rhythm of speech. Such feature of speech is still practiced in rituals of Lai Ikoukhatpa (/la:i ika:uk<sup>h</sup>ɪ̯tə/ = Invocation of God from water), Yakairol (/jəka:irol/ = sung in awakening of God), prediction of Maibi Laitongba (/ma:ibi la:ɪ̯tə<sup>h</sup>bə/ = the stage of a priest when the Godly power reveals in her/him). There are radio drama and plays which are based on the Puyalon. The Puya Theithalon is modern Meiteilon which literature starts from the post-Puya literature.

A Sanskrit verb paradigm has two primary blocks of Dhātu and Pratyaya i.e. the verbal root and the suffix respectively. By the time of Panini the concept of Dhātu had already been mature and widely used as it can be inferred from the Panini’s texts. Panini uses a well-structured verbal roots lexicon called Dhātupāṭha where the Dhātus are arranged as metadata with their phonetic tags called Anubandhas which are meant to be interpreted by rules in paradigm generations.

As Panini’s school is considered to be the most popular among the several grammar schools, the Dhātupāṭha (DP) of Paninian school is also widely used among the available Dhātu lexicons.

### **1.1.The Dhātus**

The roots in Paninian DP are listed in a syntactic form with their meanings i.e. the root followed by a phrase stating the meaning of the root. Thus the sentence is sometimes referred to as



Dhātusūtra. Most of the Dhātus in DP, thus has one meaning and appears to be a noun in locative. For instance, Bhū Sattāyām, edha vṛddhau etc. In some cases the whole sentences appears to be a Bahuvrīhi compound with the word – artha as gatyartha, śabdārtha etc.

From the generative perspective (not semantic) two types of Dhātus are there in the system of Panini. First is a set of primary stems listed in DP, i.e. Bhūvādayo dhātavah (01.03.001). These Dhātus in DP are divided in 10 classes. The second type of Dhātus are generated by the suffixation of 12 types of suffixes (Dixita, 2007). They are “san, kyac, kāmyac, kyaṣ, kyañ, kvip, ñiñ, īyañ, ñic, yak, āya, yañ”. The basic semantics of so-called postulation of these suffixes is that the suffixes modify the meaning of the primary verbal stem. The second point is that after its modification it is again inflected for tenses, aspects and moods. Structurally it seems that suffixes occur between the primary terminal and a primary stem which is actually not. In terms of morphological semantics, the suffixes attached to the primary meaning of the primary stem depending on the speaker’s intension. Then the terminal takes place. Besides, there are many other suffixes that occur between the primary terminals and stem. But they must not be intermingled with them.

### 1.2. Anubandha

In Paninian system, Anubandha (Anb) is the technique of tagging data, way or marking a particular object etc. An Anb is a mark, a key or an ID attached to an object. Panini defines Anubandhas between 1.3.2 to 1.3.8 of the Aṣṭādhyāyī. The Anubandhas are the keys that interact with the Sutras/rules in the generative system of Aṣṭādhyāyī. The rules process the roots data through keys of Anb. In this way both of the Anb. and rules are interdependent. The rules are generalized according to the data in DP and data in DP also requires to be extended/updated according to the prevailing rules.

### 1.3. Listing of Roots and Classes

One of the reasons for dividing roots in classes is phonetic variances of the paradigms of Dhātus. Technically a syllabic modifier (SM or a Vikaraṇa) of a particular class of a stem occurs when the Dhatu undergoes primary terminal process. SM modifies the primary syllabic feature, mostly mono-syllabic, of a stem. Just as Bhū is modified as Bhava while it undergoes of primary terminal process of Tiñ. Thus it becomes Bhavati > Bhū + ti. SM has no role of semantic modification. Thus Bhava still means Bhū. There are 7 SM of the 10 classes. They are listed below:

Class	Root Syllabic Modifier
bhvādī (C1)	a > śap
adādī (C2)	---
juhotyādī (C3)	---
divādī (C4)	a > śyan
svādī (C5)	nu > śnu

tudādī (C6)	a > śa
rudhādī (C7)	na > śnam
tanādī (C8)	U
krayādī (C9)	nā > śnā
curādī (C10)	aya > ṇic

**Figure 1: Table Showing the Syllabic Modifiers of Ganas.**

The table indicates that Class 2 and Class 3 have no syllabic modifier which means stems/Dhātus of these two class undergoes no syllabic modification in paradigm generation. Other classes also have limited syllabic modification according to the Lakāra (La). Technically the limitation of syllabic modifier occurrence is handled by classifying the 10/11 Las into two groups as Sārvadhātuka (SLa) and Ārdhādhātuka (ALa) (Dixita, 2007). A syllabic modifier occurs only in the SLa of the non-Class 2 and non-Class 3, that also while the lakāra expresses the agent, i.e. in the active voice. Laṭ, Loṭ, Laṅ, Sārvadhātukaḷeṭ are SLa and Liṭ, Luṭ, Lṛṭ, Āśirliṅ, Luṅ, Lṛṅ are ALa.

#### 1.4.Literature Review

There are seven traditional DPs. Each of them has around 2000 verbal roots. These all are examined by Palsule (1961). He has considered every one of these DPs fundamentally; including the compilation of the verbal roots, other specialized mechanisms that shape a piece of every section in DP and so forth. He has likewise examined the issue of implications given in the DP in detail. There is one more article which is imperative in such manner. Whitney (1885) lessens the quantity of verbal roots in Sanskrit as given in DP (around 2000) to only 845 and gives a point by point record of the use of the verbs and records these roots as accessible in tremendous writing of Sanskrit, running from Vedic period upto eighteenth Century A.D. This article is vital as it contains an authentic record of the usages of the verbal roots. Edgern (1885) endeavours to decrease the quantity of the roots in Sanskrit applying the diachronic approach. He breaks down all the 10 classes of verbal roots in Sanskrit as given in tradition. He diminishes the quantity of the verbal roots based on Proto Indo-European historical underpinnings. Every one of the examinations said above and those of Rocher (1967), Nooten (1969) have concentrated from one perspective the verifiable position of the verbal roots and then again the hypothesis of importance as discussed about in Paninian tradition. Be that as it may be, we find scarcely any investigation which talks about the verb lexical semantics and how to lexicalize their mixes with pre-verbs and so forth. Confert.

#### 2. The Standard of the Lexicon

Panini's DP is an example of a well-structured verbal root lexicon. This database is designed with phonetic keys, called Anubandhas for making the roots easily interpretable to the rules. The phonetic keys are directly attached to a root. For instance, gam is entered as gamlṛ in DP. An Anb either precedes or succeeds a root entry. In some cases an Anb can be attached to both sides of a root. Anb is a Paninian hierarchy which tags data as their lexical ID that handles many operations in the derivation process. For example the sound ñ, Svarita and Anudātta



accents (SAC & AAc) are some of the Anb which are attached to roots. These Anb are tagged as terminal ID of roots that decide which roots will be terminated in AP or PP (Ātmanepadam & Parasmaipadam). The rule Anudāttañitāttamanepadam (03.01.021, Aṣṭādhyāyī) interprets the roots tagged with Anb or ID of AAc and ñ will be terminated in AP. And the rule Svaritañitah Kartrabhiprāye Kriyāphale (03.01.072, Aṣṭādhyāyī) terminates roots tagged with ñ and SAC in AP in the semantic environment of the result (fruit) of a root goes to the agent.

### **2.1. Selected Source Texts for Digitisation**

Three editions of DP of different authors have been adopted to develop this standard of DP. Dhātūrūpanandini of Hegade is published in 2013 by Samskrita Bharati, New Delhi. This edition has semantic entries of Sakarmaka and Akarmaka and terminal entries of Parasmai, Aatmane and Ubhaya. Apart from this it discusses the meaning entries of different Ācharyas like Madhava etc. This is the latest up to date edition of the DPs. The importance of Prakriyānusārī Pāṇinīyadhātupāṭha of Puspa Dixita is its Hindi translation of the meaning entries. The Mādhavīyadhātuvṛtti edited by Dvarikadasa has extensive commentary on meaning entries.

### **2.2. The Entries of the Lexicon**

Three Primary entries of the DP are created for a digital DP. They are: E1 for the Dhātus with their Anubanthas, E2 for the Dhātus without their Anubandha, E3 for the Dhātus with their meanings as it is as given in the original DP. Meiteilon Translations of E1 and E3 are provided. Other constituents and attributes of these word heads are their transliterations and voice recordings. IAST transliterations of Devanagari are provided while IPA for Meiteilon is provided against this. The voice pronunciation recording of both Sanskrit and Meiteilon entries were manually done. Meiteilon voice recording were collected from native speakers. The Sanskrit voice recordings were recorded by the author and were commented and verified by speakers of two different North and South Indian native speakers.

Sanskrit and Meiteilon shares primary verb morphology. It made easier in one to one translation mapping in certain cases. The roots of the both languages are bound morphemes. They are mono-syllabic composed of a maximum of four phones. All the first letters (sounds) of the roots are similar in both cases except begun by a vowel. Most of the vowels at the second position of the roots take changes. The cluster consonant at the beginning of a root is absent in Meiteilon. Most of the last consonants in Sanskrit almost disappear in Meiteilon. Major semantic variances occur in shifting of transitive to intransitive to transitive verbs. Meiteilon is strictly following the syntactic rule Suptiñantam Padam. This rule defines a syntactic word must have a Sup i.e., nominal affix or a Tiñ. In generative process of Panini system, a lexicon word cannot occur in a sentence directly without anyone of these affixes.

### **2.3. The encoding standard**

The lexicon data of both Sanskrit and Meiteilon are stored in Unicode using Unicode font faces. Meitei Script fonts namely RATHA, RATHA99, rathayek, Meetei Mayek, Eeyek Unicode are available online free. Eeyek Unicode has been use for Meiteilon.

## 2.4. The Meaning Entries

First of all, meaning entry is important to trace the roots in mapping the rules for their paradigm derivations or generations. The root entries are very technical and hard to understand even for a Grammarian. The roots do not represent the actual practical data. It may cause difficulties while constructing the rules of generation. For instance, Edh cannot be found in spoken data. And there may be such homophones in the same DP also which may create more problems. Secondly, meaning entry provides a chronological account also. Many other families of Sanskrit can be comparatively established on the basis of meaning entries of the roots. By the time, the meaning of a root might be narrowed or broadened and changed. Other attributes and additional update can also be added. Entry of very genuine examples on the basis of contemporary spoken data could be a very wise task. Pictorial presentation of actions is very possible and simple task in this age of technology.

## 3. Sanskrit to Meiteilon Translation

Technically, from the view point of morphology, Sanskrit to Meiteilon translation does not seem to be as complex as Sanskrit to English, though both Sanskrit and Meiteilon belong to two different language families according to presently accepted researches. Here are some facts that would make Sanskrit to Meiteilon Translations easier and not time consuming. Both of them follow SOV syntax. Verbs (syntactic not lexicon) are agglutinative and bound morphemes and inflected from the Dhātus. Both follow the inflectional hierarchy of  $V \square R + T$  where verb or V is a production of a verbal Root/Dhātu called R and primary terminal called T. Modifications in this primary rule produces other paradigms of other tenses and moods in both of the languages. Meiteilon exhibits clear cut Vibhaktiyartha system of semantics though the verb is not inflected for persons and numbers primarily, though it can be seen in imperative and optative moods. Thus the semantic net of Kāraka is also easily applicable and must have a vital role in translation.

Major challenges may occur in vice versa Translation of Meiteilon to Sanskrit. Divergence of SOV where S is omitted in many cases in Meiteilon. It is also found that the objective case assumes zero morphemes in Meiteilon. Meiteilon verb does not have voice where subject or object is expressed by the constituents of the morphemes of a syntactic verb. Instead the primary terminal of a Meiteilon verb expresses only the action/verb itself which corresponds to Bhāvavācyā in Sanskrit.

Literally Vācyā (voice) means: “to be spoken”. Then the following argument can be built up—  
*What to be spoken?*

Either subject or object or the verb in a particular time. (All of them cannot be spoken in a time).

*By what* subject or object to be spoken?

By the Lakāra or primary terminal.

It can be drawn that Vācyā is an inherent element of a verb that primarily expresses the number and person of the subject or the object.



The Semantic of Voice firstly lies on the expresser's or speaker's emphasis on Subject or Object. As a psychological fact that the positions of the subject and object take changes according to the emphasis of the expresser as in English. In the passive voice, the object takes the first position of order of the words in the sentence. For instance, the article was completed by the researcher. It shows the emphasis of the expresser on the article. He is intending to talk about the article more importantly than the researcher. He thinks the article first. Therefore he speaks the article first and he puts the article first in the sentence. As a result it changes the normal word order SVO into OVS in the passive voice.

Although the Sanskrit no doubt has more flexible word order and its SOV even in passive voice is not ignorable, there are enough literary data that favour the change of its normal SOV into OSV in passive voice. Let us observe the following examples:

- ajñāpito-smi pariśadā... (mudrārākṣasam, 1st Act.)
- etad-api nāma śrotavyam (mudrārākṣasam, 7th Act.)
- abhijñānaśakuntalākhyena nāṭakena-upasthātavyam asmābhiḥ (abhijñānaśākuntalam, 1st Act.)
- eṣā me manorathapriyatamā sakusumāstarāṇaam śilāpaṭṭamadhisayānā sakhībhyām anvāsyate (abhijñānaśākuntalam, 3rd Act.)
- bhuktojjhita eva viṣayotrābhavatyā (svapnavāsadattam, 1st Act.)

Meiteilon verb is primarily inflected for the tense, aspect, and mood only. However a concept like voice can be discussed on the basis of the expresser's emphasis on the subject or object which frees to change the normal Meiteilon SOV into OSV. Thus, instead of the voice, a Meiteilon sentence can be divided into two expressions:

- Subjective Expressions (SE), and
- Objective Expressions (SE)

Let us discuss the following examples:

- miḥātlè
- əŋɑ<sup>n</sup> linnəçike
- ɛənma:inə ci<sup>n</sup>ŋe çakkənanə<sup>h</sup>oŋge?
- no:e:gi:la:ùəənnəlo:icɑ:fa:urəbəni

SE follows the SOV while the OE follows OVS. The main difference between the Vācyā and Expression is the person-number agreement of the verb. Else, the rest of the syntactical rules of Karmavācyā (passive voice) is followed by the OE, i.e. a) the object at the first position, b) the subject in instrumental case (Kartṛkaraṇayostṛīyā) and c) the verb at last position of the sentence.

### **3.1. The Case of Dhātupāṭha Translation**

There are more than 600 frequently used Meiteilon roots which corresponds to Sanskrit Dhātus for one to one translation. The divergence of Vibhakti (Saptami) occurs in meaning entry

translation. However, the exact translation of Saptami into Meiteilon as it is also not a very poor and unintelligible one. It can be translated as it is. Lengthy translation of those Nāmadhātus appears to be more interpretation than a common translation. Because such feature of morphology does not prevail in Meiteilon.

#### **4. Conclusions**

Fortunately, by the time encoding standard and font faces for the parallel digital lexicon creation have not been a big issue due to Unicode and available free fonts of both Devanagari and Meiteilon. Overall, it is primarily a text based digital lexicon which would be used for other research proposes. The voice recordings of the texts would be applicable for development of TTS, AST and NLI tools. One of the important applications of the voice entries will be in online eLearning lexicons web interfaces.

#### **5. Future Works**

Rule based verb paradigm generation tools for both of the languages can be developed. The paradigm generation tools may further simplify and help in building verbal paradigm databases and lexicons. Sanskrit to Meiteilon translations of simple sentence (of one phrase) can be carried out. Verb has a vital role in syntax of the both of the languages. Translations of syntactic verb paradigms would be a very beneficial for further complex translations of Sanskrit to Meiteilon.

#### **6. Acknowledgements**

To reach some of the points discussed in this paper, a number of research articles, books, dissertation and thesis etc. have been gone through. They are the base and source of the ideas presented in the paper. The suggestions of those persons including friends, juniors, seniors, and teachers etc. have been a big contribution for this piece of research.

#### **7. Bibliographical References**

- [1] Apte, Vaman Shivram. (2000). The Students Sanskrit-English Dictionary. Delhi: New Bharatiya Book Corporation.
- [2] Benedict, Paul K. (1972). Sino-Tibetan. Cambridge: Cambridge Publication.
- [3] Bharati, Akshar et. al. (2010). Natural Language Processing: A Paninian Perspective. New Delhi: PHI Learning Private Limited.
- [4] Cardona, George (1998). Panini: A Survey of Research. New Delhi: Motilal Banarsidass Publishers.
- [5] Devi, Bidyarani Mayangbam, (2002). Manipuri Verbs. Unpublished Ph. D. Thesis, Imphal: Manipur University.
- [6] Dvivedi, Kapildev (2008). Sanskrit Vyākaraṇa Evam Laghusiddhāntakaumudī. Varanasi: Visvavidhyālay Prakāsan.
- [7] Gupta, Pushpa (2007). Aṣṭādhyāyī Sahajabodha. Delhi: Pratibha Prakashan.
- [8] Hemachandra, Chanam (2014). Poireitol Khunthok. Imphal: Chingtam Press.





- [9] Hemachandra, Chanam (2014). Numit Kappa. Imphal: Lamyamba Publishers.
- [10] Jha, Girish Nath (1993). Morphology of Sanskrit Case Affixes: A Computational Analysis (Unpublished M. Phil. Dissertation). Jawaharlal Nehru University: New Delhi.
- [11] Kale, M. R. (2002). A Higher Sanskrit Grammar: For the Use of School and College Students. New Delhi: Motilal Banarsidas Publishers.
- [12] Kapoor, Kapil (2005). Dimensions of Panini Grammar: The Indian Grammatical System. New Delhi: D.K. Print World Ltd.
- [13] Khangamba, S. Kangleinganba (2014). Wakoklol Hilel Thilel Salai Amailol Pukok Puya. Imphal: Leimahal Press.
- [14] Mishra, Sudhir Kumar (2007). Sanskrit Karaka Analyzer for Machine Translation (Unpublished Ph.D. Thesis). Jawaharlal Nehru University: New Delhi.
- [15] Mitkov, Ruslan, ed. (2004). The Oxford Handbook of Computational Linguistics. Oxford: Oxford University Press.
- [16] Parrat, Saroj Nalini Arambam (2005). The Court Chronicle of the Kings of Manipur: The Cheitharol Kumpapa. 3 vols. London: Routledge Publication.
- [17] Sharma, Rama Nath (2000). The Astadhyayi of Panini. Vol. I. Delhi: Munshiram Manoharlal.
- [18] Singh, N. Khelchandra (2004). Manipuri to Manipuri and English Dictionary. Imphal.
- [19] Singh, Chungkham Yashwanta (1984). "Some Aspects of Meiteilon (Manipuri) Syntax." Diss. Jawaharlal Nehru University.
- [20] Singh, Chungkham Yashwanta (2014). Will Manipuri Be a Classical Language. Retrieved from E-Pao: [http://e-pao.net/epPageExtractor.asp?src=MeiteiMayek.Will\\_Manipuri\\_be\\_a\\_Classical\\_Language\\_By\\_Chungkham\\_Yashawanta.html](http://e-pao.net/epPageExtractor.asp?src=MeiteiMayek.Will_Manipuri_be_a_Classical_Language_By_Chungkham_Yashawanta.html).
- [21] Singh, Chungkham Yashwanta (2000). Manipuri Grammar. New Delhi: Rajesh Publications.
- [22] अष्टाध्यायीसूत्रपाठ (सं. पण्डित ब्रह्मदत्त जिज्ञासु), पाणिनि, रामलाल कपूर ट्रस्ट, २००९
- [23] अष्टाध्यायी (श्रीललितावृत्तिसंवलित), पाणिनि, कैलास विद्या प्रकाशन, ऋषिकेश (उत्तराञ्चल) २००५
- [24] धातुरूपनन्दिनी, जनार्दन हेगडे, संस्कृतभारती नवदेहली, २०१३
- [25] प्रक्रियानुसारी पाणिनीयधातुपाठ, पुष्पादीक्षित, संस्कृतभारती, नवदेहली, २०१०
- [26] माधवीया धातुवृत्ति, (सं. स्वामी द्वारिकादास शास्त्री), सायणाचार्य, तार बुक एजेंसी वाराणसी, २०००
- [27] संस्कृतव्याकरण एवं लघुसिद्धान्त-कौमुदी, डा. कपिलदेव द्विवेदी, विश्वविद्यालय प्रकाशन, वाराणसी, २००८