

Sanskrit Sentence, Dhatus, Lakara and Verb Identification for Karaka

Mr. Uday Narayan Singh,

Dr. N. Shankaranarayana Sastry

Introduction

The present research study presents about Sanskrit sentence, Sanskrit verb and its identification, classification of dhatus and concept of Lakara. Sanskrit is an inflectional language and words in sentence carry information about entities in terms of stem, endings, gender, case, number and case relation, while verbs denote activity, function reaction mode, voice, tense, person, number etc. Extracting and annotating this information is the first step towards understanding the language.

Sanskrit Sentence

What is the definition of *vakya*? How can computers identify it? If we look at the definition of *vakya* in historical perspective, Yaska's fourfold division of language indicates that he is analyzing sentence through these four (*nama*, *akhyata*, *upasarga* and *nipata*) structural components. Though Panini does not define *vakya* explicitly, his notion of sentence is inferred mainly through *karaka prakarana* and some other relevant *.suras*. The notion of *samarthya* is central to understanding the concept of *vakya* accepted by Panini.

Definition of Dhatu

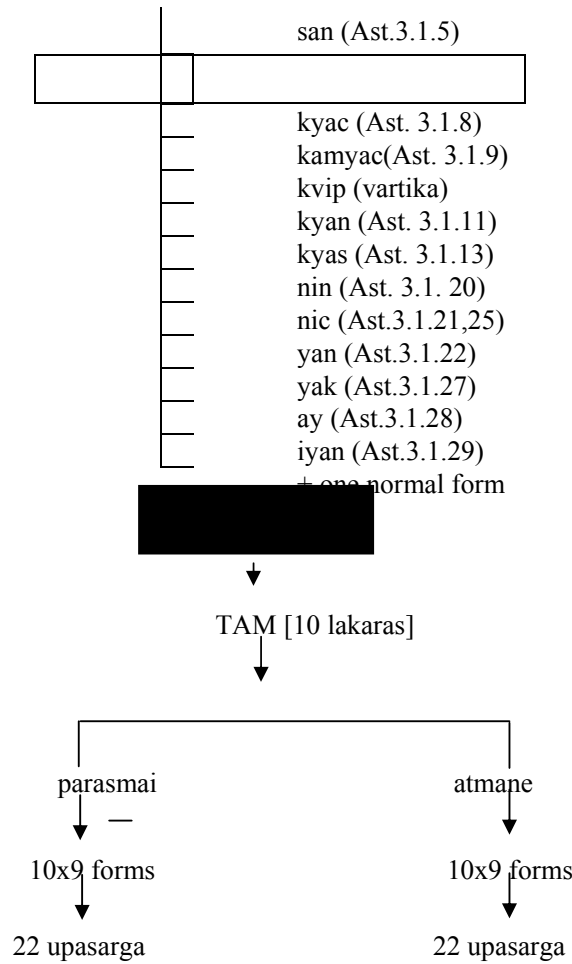
The verb is the most crucial and significant part of the sentence. It determines the nature of sentence on semantic as well as syntactic levels. According to Wikipedia¹ a verb is a part of speech that usually denotes action ("bring", "read"), occurrence ("to decompose" (itself), "to glitter"), or a state of being ("exist", "live", "soak", "stand"). Depending on a language, a verb may vary in form according to many factors, possibly including its tense, aspect, mood and voice. It may also agree with the person, gender, and/or number of some of its arguments (what we usually call subject, object, etc.).

The verb has been defined by Panini as that which belongs to the class of \sqrt{bhu} (become), \sqrt{va} (blow) etc. Panini does not define *dhatu samjna*. According to Kasika² this *samjna* is given by ancient grammarians. Panini says *bhuvadayo dhatavah* but does not explain *bhu* and *va*. Patanjali says- '*kuto ayam vakarah ? yadi tavat samhitaya nirdesh kriyate 'bhvadaya' eti bhavitavyam* Generally grammarians have explained it in different ways. According to Kasika, Panini has inserted *va* for auspicious reasons because in the tradition of Sanskrit, generally all authors like to do *mangalacarana* (benediction) which can be done at start, middle or end of the text³. According to this interpretation the *bahuvrihi* compound with its internal *dvandva*, will have analyzed form *bhusca vasca adisca bhuvadi adir yesam te bhuvadayah*.

Classification of dhatus

Sanskrit verb forms carry tense, aspect, person, number information all in the inflection forms. Besides, they can also contain derivations containing semantic information like causation, desire, repetition, negation etc. Therefore it becomes very difficult to split out the verb and separate the verb root and complex information units encoded in it. Sanskrit has about 2000 verb roots classified in 10 morphological and semantic classes called *ganas*, and can also be further sub-classified as normal forms (without any of the 12 derivational affixes - 11 listed by Panini [P 3.1.32], one more '*kvip*' added by Katyayana), and the derived forms with *nijanta* (causative - *nic*), *sannata* (expressing desire - *san*), *yananta* (duplicated - *yan* and *yanlunanta*), *namadhatu* (nominalized -*kyac*, *kamyac*, *kvip*, *kyan*, *kyas*, *nin*, *yak*, *ay* and *iyam*). Further, these can have atmane and parasmai forms in 10 lakaras and 3x3 person and number combinations, and can also be potentially prefixed with 22 prefixes. Finally there could be innumerable *namadhatus* (nominalized verbs). According to a rough calculation, all potential verb forms in Sanskrit may be around 10,29,60,000 plus *namadhatus*. The distribution of Sanskrit verbs can be understood as follows-

VR (2000)/ND (namadhatus)



Therefore the approach followed by many to store Sanskrit verb forms is not going to work. Hence a reversed Paninian approach is proposed for parsing the complex verb forms in Sanskrit in the following sequence:-

The verb inflection (parasmai / atmane) is identified from database and a rough guess is made about the verb,

The lakara information based on inflection is obtained.

Each of the 12 derivational affixes is evaluated.

Each of the 22 upasargas (prefixes) is searched.

The verb root is determined by weeding out other elements and database matching.

The verb root thus identified with all the other potential components and the grammatical information gathered in the form of verb tags can be potentially used in a machine translation system performing translation from Sanskrit.

Concept of *lakara*

There are eleven *lakaras* in Sanskrit grammar in which ten are used in *laukika Sanskrit* (for example, *lat*, *lit*, *lut*, *It*, *lot*, *lan*, *vidhilin*, *asirlin lun* and *lrm*) and one namely *let lakara* is used only in the Vedic text. For this, the term used is 'the Vedic Subjunctive'. The *lat*, *lot*, *Ian* etc. is the suffix which is counted by Panini in third chapter of Astadhyayi. In these suffixes, after *anubandha lopa*, the only remaining letter is *-l*. Therefore, the combined names of these are *lakara* and are counted in the sequence of *pratyahara*. These *lakaras* can classify in two groups as *tit* (t-ending) and *nit* (n-

ending). Panini says the purpose of *tit*⁴ of that in these tenses 'ti' (from the beginning of a word, last vowel is called '*ti samjna*'⁵ of the *atmanepada* affixes change into 'e' and the purpose of *nit*⁶ is in *lun* (Aorist), *vidhilin* (Potential). *Ian* '.Imperfect) and *In lakara* (Conditional).

The 's' is elided in the *uttama purusa* (first person). According to Panini in the derivation of verb form after dhatu comes 'l' and in the place of 'l' will be substituted⁷ the 18 affixes⁸ which are counted in another *sutra* as -

tiptasjhisipthasthamibvasmastatamjthasathamdhvamidvahirahin.

Structure of Sentence

These *lakaras*¹⁹ always come with *sakarmaka* roots in the sense of *kartr* and *karma*, and with *akarmaka* roots in the sense of *kartr* and *bhava*. Panini accepts in this *sutra* three meanings of *lakaras*.¹⁰ Therefore there are three types of sentences as follows

- *kartr vacya* (active voice) as *ramah grham gacchati*. In *kartr vacya*, *karta* plays an important role in the sentence. According to *karta*, *lin*, *vacana*, *vibhakti* or *purusa* are used in the verb. In *kartr vacya*, *dhatu* conjugates in *atmanepada* or *parasmaipada*. The basic structure of sentence in *kartr vacya* is- S = subject in first case + object in second case + verb according to subject.
- *karma vacya* (passive voice) as *ramena phalam khadyate*. In *karma vacya*, *karta* is completely free and according to *karma*, *in*, *vacana*, *vibhakti* or *purusa* are used in verb. In *karma vacya*, *dhatu* conjugates only in *atmanepada*. The basic structure of sentence in *kartr vacya* is - S = subject in third case + object in first case + verb according to object.
- *bhava vacya* (impersonal voice) as *tena gamyate*. In *bhava vacya* subject is used in third case and verb is always used in *prathama purusa eka vacana*. In *bhava vacya dhatu* conjugates only in *atmanepada*.

All *dhatu*s can be classified in three groups according to the sentence but in this division Panini neither discussed nature of *dhatu*s nor number of *dhatu*s. May be it is understood by name as *sakarmaka*, *akarmaka* and *dvikarmaka*. Later grammarians try to count the number of *dhatu*s in the above groups. Those are *sakarmaka dhatu*s which expect at least one object in the sentence for example, *sah grham gacchati*. Those *dhatu*s which do not expect any object are *akarmaka dhatu*s for example, *sah asti*.

TGL id module

This module will identify the verb inflections (*tin*), class (*gana*) and TAM (*lakara*) based on a database. Panini used 18 verb inflection suffixes called *tin* (P 3.4.78) (9: 3x3 for *parasmai* and 9: 3x3 for *atmane*) for verbs in different *ganas* (*bhvadi*, *adadi*, *juhotyadi*, *divadi*, *svadi*, *tudadi*, *rudhadi*, *tanadi*, *kryadi* and *curadi*) and tense (*lat*, *lot*, *lan*, *vidhilin*, *It*, *lut*, *asirlin*, *In*, *lit* and *lun*). Among these tenses only six are tenses and four are moods as following:-

Lakara	Tense/Mood	Kala/Bhava ¹¹
Lat	Present Tense	Vartamana Kala
Lit	Past Perfect Tense	Bhuta Kala
Lut	Ist Future Tense	Bhavisyan Kala
Lrt	2nd Future Tense	Bhavisyan Kala
Lot	Imperative Mood	Ajnartha
Lan	Past Imperfect	Anadyatana Bhuta Kala
Vidhilin	Potential Mood	Vidhi
Asirlin	Benedictive Mood	Asih
Lun	Aoris't Tense	Asanna Bhuta Kala
lrn	Conditional Mood	Kriyatipattiyartha

Table : 1(TAM)

Verb Identification for Karaka

Verb analysis through reverse Paninian approach is a very difficult task and along with karaka analysis, it can not be completed in a limited time of research Paninian grammar is generative so except primary or normal verb, forms all other secondary forms like *nijanta*, *sannata* and *yananta* etc. are making it more difficult. It has been decided to make a data-bases of frequently used verbs of the Sanskrit (approximately 438). This number may be debatable but verbs are searched from a Sanskrit corpus, current Sanskrit literature and those appearing in the *karaka prakarana* of SK and Ast.. In the first database (*dhatu-pdtha*) verb information are stored. The structure of Dhatu-patha data-base is as follows-

dhatu id	dhatu	gana	artha	pada	sct/vct/anit	karmak
212	डुकृञ्	तनादिगण	करणे	उभय	सेट्	सकर्मक
367	कुप	दिवादिगण	क्रोधे	परस्मै	सेट्	अकर्मक
39	क्रटि	भ्वादिगण	आह्वाने	परस्मै	सेट्	आह्वाने सकर्मक रोदने त्वकर्मकः
174	ष्ठा	भ्वादिगण	गतिनिवृत्तौ	परस्मै	अनिट्	अकर्मक
365	माङ्	जुहोत्यादिगण	माने	आत्मने	अनिट्	सकर्मक / अकर्मक
391	आप्त	स्वादिगण	व्याप्तौ	परस्मै	अनिट्	सकर्मक

In the second (dhatu roopa) data-base verb forms are stored with same unique dhatu id. The information of वाच्य, पद, लकार, पुरुष and वचन is also stored along with dhatu roopa. The sequences of the information are as follows-

वाच्य- 1 = कर्तृ, 2 = कर्म, 3 = भाय

पद- 1 परस्मै, परस्मै 2 = आत्मने

लकार- 1 = लट् 2 = लोट्, 3 = लङ् 4 = विधि, 5 = लिट्, 6 = लुट् 7 = लृट्, 8 = आशी, 9 = लुङ्, 10 = लृङ्

पुरुष - 1 = प्रथम, 2 = मध्यम, 3 = उत्तम

वचन - 1 = एकवचन, 2 = द्विवचन, 3 = बहुवचन

The karmani and bhava vacya verb forms are not stored till the time due to limited time of research work. The structure of the dhatu roopa data bases is as follows

Dhatu_id	Dhatu_roopa	Vacya	Pada	Lakara	Purusa	Vacana
311	त्यजति	1	1	1	1	1
291	चिकित्सताम्	1	1	2	1	2
99	अपठन्	1	1	3	1	3
29	मन्थे :	1	1	4	2	1
170	बुबुधथुः	1	1	5	2	2
177	हसितारथ	1	1	6	2	3
426	प्रेष्यामि	1	1	7	3	1
318	नम्यास्व	1	1	8	3	2
429	अयोष्म	1	1	9	3	3
267	अवर्णयिष्यामहि	1	2	10	3	3

Table (dhatu rupa identification)

Conclusion

Panini lists approximately 2000 *dhatu*s in the DP. Among these more than one *dhatu*s are used in the same meaning. For example, 182 *dhatu*s are used in the sense of *gati* (movement), whereas $\sqrt{\text{gam}}$ generally used in Sanskrit literature in this sense. Whitney¹² in his 'Review of recent studies in Hindu grammar' informs us that in the DP there are 'thousands or twelve hundred false roots and declares that the fact their voices being not less carefully defined by the *Dhatupa tha* those of the eight or nine hundred genuine ones casts a shade of unreality over the whole subject of voice-conjugation'. Georg Buhler objected to this remark by Whitney by saying that this couldn't have been Panini's or Whitney's invention.

REFERENCES

1. <http://en.wikipedia.org/wiki/verb>
2. *dhatuśabdah purvacaryasamijna-* Kasika p. 45.
3. *mangaladini mangalamadhyani mangalantani hi sastrani prathante virapurūśani ca bhavāntyaśmatpurūśani cadhyetarasca managalayukta* -Mahabhasya , p. 153.
4. Ast. 03.04.79.
5. Ast. 01.01.64.
6. Ast. 03.04.99
7. Ast. 03.04.77.
8. Ast. 03.04.78
9. Ast. 03.04.69.
10. Shastri, Bhimasena, 1998. 'Laghusiddanta-Kaumudi, Bhaimi Publication, Delhi, Vol-2, p. 4.
11. *ladvartamane ledvede bhute lunlanilatastatha vidhyasisostu lidlotau lutlrtrn ca bhavisyati.*
12. George Buhler, 'The Roots of the Dhatupatha not found in the Literature form J.F. Stall (ed.), 'A Reader on the Sanskrit Grammarians', MLBD, Delhi, pp. 194. 1985.