Adverbs in the Sanskrit wordnet

Tanuja Ajotikar

Malhar Kulkarni

Roadmap

▸ Introduction

▸ Adverbs in Sanskrit

▸ Treatment of adverbs in lexical resources of Sanskrit

▸ Adverbs in the Sanskrit wordnet

▸ Conclusion

Introduction

- Adverb: an open-class lexical category
  - Manner
  - Time
  - Place
  - Cause
  - Answers to the questions as "how," "where," "when" and "how much"

- English WordNet
  - Fellbaum describes adverbs as a *heterogeneous group* in which not only *adverbs derived from adjectives* are included but *phrases used adverbially* are also included. Some of these phrases are included in the WordNet. These phrases are mainly *frozen phrases* that are *used widely.*

(Fellbaum (1998): p.

Adverbs in Sanskrit

- Traditionally only two classes in Sanskrit
  - *Subanta* Words that end in nominal affixes
  - Tiṅanta Words that end in verbal affixes, i.e. verbal roots
  - Verbs and Non-verbs

- *Tiṅanta* -s: verbs

- *Subanta*-s: nouns, adjectives, adverbs, particles
  - No formal distinction between these word classes
- Adjectives or adverbs not formally distinguished in traditional grammar of Sanskrit

- sa śīghraṁ gacchati.
  - He goes quickly.
  - śīghraṁ is the nominative and accusative singular form of the nominal base śīghra ending in the vowel *a*.
  - It is a *subanta* according to the traditional grammar.
  - Traditional analysis does not categorize this form further.
  - This is a same form as any other nominal base that ends in the vowel *a*.
- Mere form does not help categorize words in Sanskrit.

- Modern attempts to analyze Sanskrit in word classes
  - Monier-Williams(1846), Wilson (1841), Speijer (1886), Whitney (1879) and Macdonell (1927)
- Words in Sanskrit can be classified under adjective and adverb
- Summary of analysis of adverbs in Sanskrit by modern Sanskrit scholars
  - Some of the particles
  - Words end in *vat, tra, tas, śas dā* and *thā*
  - Accusative of a noun or an adjective
  - Instrumental and ablatives derived from substantives
  - Some *Bahuvrīhi* (exocentric) compounds
  - *Avayībhāva* compounds

- Gombrich (1979) discusses six cases of adverbs
  - *atra* 'here'
  - *ciram* 'for a long time'
  - *tūṣṇīm* 'silently'
  - *javena* 'speedily'
  - *śīghram* 'quickly'
  - *vividhaprakāram* 'variedly'
- First three are analyzed and categorized by the traditional grammarians.
  - They are categorized as particles.
- Remaining are not differentiated by their use.

The importance of part-of-speech categories in lexical entries

- Lexicon- inventory of lexemes
  - Lexemes – an abstract unit emerged from morphological analysis of language
  - Does not occur in isolation
  - Determine syntactic structure of a sentence
  - Morphological and syntactic behavior determines its class
- Contains word-class of lexeme
- Lexicon='structuring sentences'
- Nature of language decides which word-classes occur in lexicon

The treatment of adverbs in Sanskrit dictionaries

- 18 dictionaries of Sanskrit : 1819–1981
  - Radhakanatdeva, (Monolingual), 1819-1858
  - Wilson H. H., Sanskrit-English, 1832
  - Yates W., Sanskrit-English, 1846
  - Bopp F., Sanskrit-French, 1847.
  - Böhtlingk, O. and Roth R., Sanskrit-German, 1855-1875
  - Goldstükar T., Sanskrit-English, 1856
  - Benfey, T., Sanskrit-English, 1866
  - Burnouf , É., Sanskrit-French, 1866
  - Böhtlingk, O., Sanskrit-German, 1879-1889
  - Monier-Williams M., Sanskrit-English, 1872
  - Bhattacharya T., (Monolingual), 1873
  - Cappeller, C., Sanskrit-German, 1887
  - Apte V. S., Sanskrit-English, 1890
  - Cappeller, C., Sanskrit-English, 1891
  - Macdonell A. A. Sanskrit-English 1893
  - Monier-Williams M., Leumann, and Cappeller, Sanskrit-English, 1899
  - Stchoupak, N., Nitti, L. and Renou L.,Sanskrit-French, 1932
  - Ghatge, A. M., Sanskrit-English (Encyclopedic dictionary on historical principles), 1981

Atra 'here'

- Tradition : derivable, an indeclinable and a *subanta*
- Lexicographers' treatment :
  - Category
    - Indeclinable : 6 dictionaries
      - Agree with the tradition
      - Monier–Williams (1872), Monier–Williams, Leumann, and Cappeller (1899), Apte (1890), Goldstücker (1856), Radhakantadeva (1819–1858) and Bhattacharya (1873–1884)
    - Adverb : 12 dictionaries
    - Lack of any category : 1 dictionary
      - Cappeller (1887)

Ciram 'for a long time'

▸ Tradition : non-derivable, an indeclinable and a *subanta*

▸ Lexicographers' treatment :
▸ Category
- Indeclinable : 5 dictionaries
  - Agree with the traditon
  - Radhakantadeva (1819-1858), Wilson (1832), Monier-Williams (1872), Bhattacharya (1873-1884), Monier-Williams, Leumann, Cappeller (1899)
- Adverb : 5 dictionaries
  - Macdonell (1893), Yates (1846), Bopp (1847), Cappeller (1887), Cappeller (1891)
- Lack of any category : 5 dictionaries
  - Apte (1890), Böhtlingk and Roth (1855-1875), Benfey (1866), Burnouf (1866) Böhtlingk (1879-1889)

▸ Listed under the lexeme *cira* : 6 dictionaries
- Macdonell (1893),, Böhtlingk (1879-1889) Monier-Williams (1872), Monier-Williams, Leumann, and Cappeller (1899), Benfey (1866), Burnouf (1866)
▸ Listed separately : 2 dictionaries
- Radhakantadeva (1819-1858), Bhattacharya (1873-1884)
- They do not consider *ciram* derivable from *cira*

Tūṣṇīm 'silently'

- Tradition : non-derivable, an indeclinable and a *subanta*

- Lexicographers' treatment :
- Category
  - Indeclinable : 6 dictionaries
  - Agree with the tradition
    - Radhakantadeva (1819–1858), Wilson (1832), Monier-Williams (1872), Monier-Williams, Leumann, and Cappeller (1899), Bhattacharya (1873–1884), Apte (1890)

  - Adverb : 10 dictionaries

Javena 'speedily'

- Tradition : the instrumental singular of the nominal base *java* 'speed'
- Lexicographers' treatment :
- No mention
- Ablative singular form of *java* : adverb
  - 1 dictionary
  - Cappeller (1887)
- No clue from the available dictionaries to analyze this form.

Śīghram 'quickly'

- Tradition : the nominative and accusative singular of the nominal base *śīghra* 'quick'
- Lexicographers' treatment :
  - Derived from the nominal base *śīghra*
    - Do not consider an independent lexeme
  - Indeclinable: 3 dictionaries
    - Monier-Williams (1872), Monier-Williams, Leumann, and Cappeller (1899), Apte (1890)
  - Adverb: 11 dictioanries
  - No category: 2 dictionaries
    - Stchoupak, Nitti, and Renou (1932), Burnouf (1866)
  - No mention of the adverbial use: 2 dictionaries
    - Radhakantadeva (1819-1858), Bhattacharya (1873-1884)

- POS of the nominal base *śīghra* in the dictionaries
  - Adjective : 10 dictionaries
  - Noun : 2 dictionaries
    - Wilson (1832), Cappeller (1887)
  - A word used in all genders i.e. *mfn* : 2 dictionaries
    - Monier-Williams (1872), Monier-Williams, Leumann, and Cappeller (1899)
  - Noun and adjectival : 2 dictionaries
    - Radhakantadeva (1819-1858), Bhattacharya (1873-1884)

Vividhaprakāram 'variedly'

- No mention of this word in any of the dictionary.
- This is the accusative singular form of the compound *vividhaprakāra* 'varied'.

Yathāśakti 'according to one's power and ability'

- An *avyayībhāva* compound
- Tradition :
  - derived from *yathā* 'according to' and *śakti* 'power, ability, capacity'
  - Indeclinable and *subanta*
- Lexicographers' treatment
  - Listed as a separate lexeme
  - Adverb 8 dictionaries
  - Indeclinable : 5 dictionaries
    - Radhakantadeva (1819–1858), Bhattacharya (1873–1884), Monier-Williams (1872), Monier-Williams, Leumann, and Cappeller (1899), Apte (1890)
  - Not listed: 3 dictionaries
    - Benfey (1866), Bopp (1847), Macdonell (1893)

Observations

- *tūṣṇīm* 'silently'
  - Not derived from any word
  - Either indeclinable or adverb
- *atra* 'here', *yathāśakti* 'according to one's ability'
  - Listed separately though derivable
  - Either indeclinable or adverb
- *ciram* 'for a long time'
  - Varied analysis
    - Independent lexeme
    - Derived from *cira*
- *Śīghram* 'quickly'
  - A form the nominal base *śīghra*
  - Ideally should not be a part of the consulted dictionaries
  - Adverbial use mentioned but the form not listed independently
- *Javena* 'speedily'
  - A form of the nominal base *java*
  - Ideally should not be a part of the dictionaries consulted
  - Does not find any place
- *Vividhaprakāram* 'variedly'
  - Not listed

- Lexicographers :
  - follow the traditional analysis
  - Adopt modern categories
  - Multiple functions stored under one lexeme

- Form vs function
- Formal adverbs
  - *Tūṣṇīm, atra, yathāśakti*
  - *Ciram:* if considered non-derivable
- Functional adverbs
  - *Javena, Śīghram, Vividhaprakāram*

Adverbs in the Sanskrit wordnet

- Sanskrit wordnet
  - Expansion approach
  - Source wordnet: the Hindi wordnet
  - Two options for the category adverb
    - Create the indeclinable category and then link with adverbs in the Hindi wordnet
    - Shortcomings:
      - Broad category consisting function as well as content words
      - Not acceptable to Wordnet principle
  - Adverb as a separate category
    - Strategy about adverbials
      - Store the base form
      - Store the declined form
    - SWN follows the second option

- Two synsets for those words adverbially used
  - ◦ Base form under noun or adjective category
  - ◦ Declined form under the adverb category
- Id 1922 (Adverb) *śīgrham śīghreṇa javena javāt*
- Id 5118 (Noun) *śīgrha, java*
- Making explicit the implicit strategy of dictionaries
- An automatic parser would be benefitted

Issues in linking synsets of adverbs

▸ It is difficult to link a synset in the source language if it uses an adverb to express what the target language conveys by using pre-verbs that are bound morphemes.

▸ According to the policy of the expansion approach, we cannot link a synset whose part-of-speech category in the source language differs from that in the target language.

Adverbs in Hindi and preverbs in Sanskrit

- Synset Id 10819:
- Gloss: *lauṭakara phira apane sthāna par/* 'Returning to his own place again.'

Example: *mohana kala hi veśase vāpasa āyā/*

'Mohana came back yesterday from abroad.'

Synset: *vāpasa, vāpisa* 'Back'

- Sanskrit uses the preverb and verb combination to convey the meaning 'back. The pre-verb *prati* and a verb of motion convey the concept expressed by Id. 10819
- Pre-verbs: bound morphemes
- Not possible to form a separate synset

Cross part-of-speech category

- Id 11374

Gloss: *āṁkhoṁ ke sāmanevālā* / 'One who is in fornt of eyes'

Example: *śikṣaka ne chātroṁ ko pratyakṣa ghaṭanā para ādhārita nibaṁdha likhane ko kahā* / The teacher askes students to write an essay based on actual incident.

Synset: *pratyakṣa, sākṣāata, anvakṣa, aparokṣa, Samakṣa, nayanagocara* 'evident'

- *Pratyakṣa* : *avyayībhāva* compound
  ◦ An adverb
  ◦ Not an adjective
  ◦ Not linkable

Adverbs and their relations

▸ Two relations
  ◦ 'Derived from'
  ◦ 'Modifies a verb'
  ◦ Cross POS category relations
▸ Derived from
  ◦ Links a noun or adjective and adverb
  ◦ Easy for forms like *javena* 'speedily', *śīghram* 'quickly'
    • *Javena* linked to java 'speed'
    • *Śīghram* linked to *śīghra* 'quick'
  ◦ *Ciram* 'for a long time'
    • If considered derived from *cira*: such word does not exist either as a noun or an adjective
    • Non-derived: cannot be linked to any other synset
  ◦ *Tūṣṇīm* 'silently'
    • Non-derived: cannot be linked to any other synset
  ◦ *Yathāśakti* 'according to one's power or ability'
    • Derived from *yathā* and *śakti*
    • Should be linked to both?
    • Currently linked to *śakti*

Conclusion

- Adverbs in Sanskrit formal as well as functional
- Not received any uniform treatment in the hands of lexicographers
- Formal adverbs easy to store under the adverb category
- Real challenge:
  - nominal forms, adverbially used
  - to collect all of the possible cases
- Contribution of the SWN: to lexicalize the adverbials
  - especially the declined forms of nouns and adjectives
- Adverbs in Sanskrit complex in nature
- A lexicon developed for a machine use need to adopt strategies suitable for its system.