Concordance of Sanskrit Synonyms

Dhaval Patel
Department of Sanskrit Studies,
University of Hyderabad
drdhaval2785@gmail.com

Abstract

The present paper tries to explore the possibility of creating a concordance of Sanskrit synonyms based on Samānārthaka Sanskrit kośas.

1 Introduction

Sanskrit kośas are mainly of two types - Samānārthakakośas and Anekārthakakośas. Samānārthakakośas enlist the synonyms depicting the same meaning in one place, whereas Anekārthakakośas provide multiple meanings of the same word. Various lexicographers enumerate different words to depict the same meaning in samānārthakakośas. In the present paper, the author has tried to create a concordance of Sanskrit synonyms based on five Samānārthaka Sanskrit kośas.

2 Selection of Samānārthakakośas

2.1 Amarakośa

Amarakośa (AMAR) is a well renowned Sanskrit kośa which covers both synonymic and homonymic components. In the synonymic component, the words denoting the same meaning are presented together. In the homonymic component, different meanings of the same word are presented together. Amarakośa's formal name is Nāmaliṅgānuśāsanam. It also presents the genders of the words explicitly. Amarakośa it the most widely used Sanskrit lexicon. Its popularity can be measured by the wide literature of commentaries on it in Sanskrit and many other languages. Usually, Amarakośa provides the gender information explicitly, but sometimes the commentaries thereon have been consulted to check for the gender of the given word.

2.2 Abhidhānacintāmaṇi of Hemacandra

Hemacandra has authored many lexicographic works. The most notable among them are Abhidhānacintāmaṇi - a samānārthaka kośa and Anekārthasaṁgraha - an anekārthaka kośa. Unlike Amarasiṃha, Hemacandra thought it fit to compose two separate works for synonymic and homonymic lexica. Therefore, in the present comparision, only the synonymic part of Amarakośa is compared with synonymic dictionary Abhidhānacintāmaṇi (ABCH) of Hemacandra. Hemacandra has also authored a separate work on gender named Liṅgānuśāsanam, and hence has not given gender information in his ABCH explicitly. This gap is easily filled through autocommentary on this work named Abhidhānacintāmaṇisvopajñavṛtti (ACVH) and a commentary on Abhidhānacintāmaṇi namely Vyutpattiratnākara by Devasāgaragaṇi (AVRD). Both these commentaries provide gender information explicitly.

2.3 Abhidhānacintāmaṇipariśiṣṭa and Abhidhānacintāmaṇiśiloñcha

Abhidhānacintāmaṇi also has two supplementary treatises Abhidhānacintāmaṇipariśiṣṭa by Hemacandra (ACPH) himself (also known as Śeṣasaṃgrahanāmamālā), and Abhidhānacintāmaṇiśiloñcha by Jinadeva (ACSJ). These two works provide words which are missing in Abhidhānacintāmaṇi. ACPH is a compilation of extra words which Hemacandra has provided in his

autocommentary ACVH. The only drawback in this dictionary is that the gender information is not explicitly provided by the author, and is not available in ACVH. ACSJ has a commentary on it by Śrīvallabhagaṇi named Haimanāmamālāśiloñchadīpikā (HSDS), which has gender details explicitly marked. As per the introduction to HSDS, ACPH also has a commentary by Śrīvallabhagaṇi, but it is not yet published. Therefore, the gender information regarding ACPH may have some errors.

2.4 Abhidhānacintāmaṇisvopajñavṛtti of Hemacandra

Abhidhānacintāmaṇisvopajñavṛtti sometimes lists additional words in its commentary over and above those mentioned in ABCH or ACPH.

2.5 Reason for selection of these kośas

Amarakośa is the most widely used Sanskrit lexicon. ABCH, ACPH and ACSJ put together provide a near complete coverage of Sanskrit synonyms. Amarakośa and commentaries thereon usually follow Paṇinian grammar, whereas ABCH, ACPH and ACSJ and commentaries thereon (with the exception of AVRD) follow a different system of grammar known as Haimavyākaraṇa propounded by Hemcandra himself in his work Haimaśabdānuśāsanam. Grammar and lexica of Hemcandra are widely used in the Jaina literature as a standard text. therefore, a concordance and comparision between AMAR and ABCH would provide a nice way to synthesize these two systems of grammar and lexicography.

3 Survey of existing digital resources

Currently, the dictionaries which are digitized and available online on Cologne Digital Sanskrit Dictionaries (CDSD) are mostly compiled on Western style of lexicography which enlist a word and its multiple meanings. Some Sanskrit-Sanskrit dictionaries like Śabdakalpadruma and Vācaspatyam are also made on that pattern.

There exist two digital resources which have dealt with Sanskrit synonym sets. The first is Mahakosha[1] of Karnataka Sanskrit University, which has tried to note down synonym sets and relationships between those synonym sets for two lexica namely Amarakośa and Vaijayantīkośa. The work still seems to be incomplete in its coverage, as the synsets in Vaijayantīkośa are still 1966 as on 30th January 2025. The work does not provide a concordance of synonyms across both dictionaries. The second resource is Sanskrit Wordnet[2] of Computation For Indian Languages Technology of IIT Bombay. Kulkarni et al. (2010) have taken Sanskrit words from Hindi Wordnet and three different word lists. It also includes some information from Śabdakalpadruma and Vācaspatyam definitions, but they are not directly taken from traditional Sanskrit kośas.

The present work is an attempt to provide a framework for alignment of 5 Sanskrit samānārthaka kośas. The framework is easily extensible to add new lexical resources as and when one wants to add.

4 Development of e-resource

The development of e-resource has been carried out in two pronged way - (1) Generation of lexicon marked up with synset information and (2) Alignment of synsets across lexica.

4.1 Generation of lexicon marked up with synset information

The lexica AMAR, ABCH, ACPH and ACSJ were typed in, proofread and then they were provided with the synset information. A synset is denoted with a tag "<syns>" and it is followed by word-gender pairs separated by comma. Every synset is also given a unique id for a given lexicon, which can help in identifying and aligning this synset across lexica. It is possible to use TEI standards to tag the data, but it was decided to keep it in plain text format so that

[1] https://sambhasha.ksu.ac.in/CompLing/index.html
[2] https://www.cfilt.iitb.ac.in/wordnet/webswn/

it is more human readable. It is possible to generate desired XML or HTML format from the plain text format with help of some scripts.

4.2 Alignment of synsets across lexica

Methodology adopted for alignment was a semi-automated process.

The automated part of the process is as follows. The synsets of AMAR and ABCH were checked for words which are common in prātipadika form and also in gender. (word, gender) tuple was taken as a unit for comparision. The synsets which have maximum number of common entities were aligned with each other. For example, AMAR.eid.28 would have a lot of common (word, gender) pairs with ABCH.eid.692 e.g. (viṣṇu, m), (janārdana, m), (hṛṣīkeśa, m) etc. Therefore, they will be automatically aligned with each other computationally. It may be cursorily checked by annotator manually if one wishes. In case a synset has same number of common entities with more than one synset from other dictionary, it was attached with the first occurrence. Rest of the synsets were treated to have no corresponding entries across dictionaries.

Thereafter, a display for frontend was created and it was manually examined. Necessary corrections were made in the lexical resource and mapping file wherever needed. This process also brought out a lot of cases in which AMAR and ABCH have marked genders of the word differently. The commentaries of AMAR and ABCH were consulted where necessary and gender differences were resolved as far as possible.

4.3 Concordance of Sanskrit synonyms

At the end of the activity a concordance of AMAR and ABCH was created. This worked as a scaffolding on which ACPH and ACSJ were also joined by following the same procedure. ACPH and ACSJ linking was relatively easy, because they provide the reference to synset in ABCH explicitly. Therefore, it was possible to align them with the corresponding synsets of ABCH without much hassle. There are only a few synsets which are unique to ACPH or ACSJ which are absent from ABCH, their numbers being 142 and 8 respectively. Abhidhānacintāmaṇisvopajñavṛtti (ACVH) of Hemacandra provided a great help in locating these unique synsets to appropriate place in the concordance.

The following example will show a sample regarding construction of this concordance. The synsets in all five dictionaries regarding the meaning “kṣamā” (pardon) are shown below.
AMAR - <eid>621<syns>kṣānti-f,titikṣā-f “kṣāntistitikṣā”....
ABCH - <eid>1225<syns>titikṣā-f,sahana-n,kṣamā-f “titikṣā sahanaṃ kṣamā.”
ACPH - Nil
ACSJ - <eid>1225<syns>kṣānti-f “kṣamā syātkṣāntirityapi.”
ACVH - <eid>1225<syns>kṣānti-f “kṣāntirityapi.”

The concordance generated from the above data is given in Table 1.

The concordance is easily extendable to other dictionaries as and when someone wants to add new dictionaries. This is borne out by the fact that Abhidhānaratnamālā of Halāyudha (ARMH) is currently in process of integration.

There are two versions of the concordance, one being in tab separated value format (.tsv) and one in HTML format. HTML format also has links to specific verse under consideration which mentions the given word. It will facilitate the user to verify the occurrence of word in the given lexicon.

Currently, only the samānārthaka kośas are taken up. Anekārthaka kośas have a different type of knowledge representation and may not be suitable to the same treatment. It will require a different type of visual representation in which multiple meanings of a single word would be enumerated in the concordance.

Sr. No.	HEADWORD	GENDER	AMAR	ABCH	ACPH	ACSJ	ACVH
1380	kṣānti	f	621	x	x	1225	1225
1380	titikṣā	f	621	1225	x	x	x
1380	sahana	n	x	1225	x	x	x
1380	kṣamā	f	x	1225	x	x	x

Table 1: Sample entry of the samānārthaka concordance

5 Comparision between knowledge representation of Amarakośa and Abhidhānacintāmaṇi

During the investigation for finding alignment between synsets of AMAR and ABCH, the knowledge representation structure of both dictionaries came to the fore. Nair (2011) has already studied the knowledge structure of Amarakośa. The arrangement of words into various kāṇḍas or vargas in AMAR and ABCH do not match one to one. Based on the empirical observations, the rough correspondence between these two lexica is as mentioned in Table 2.

Over and above this broad categorisation of mapping, there are many entries where the corresponding entries are not found in the corresponding part of the other dictionary as shown in the above chart. Therefore, some mapping had to be done based on manual examination.

6 Statistics

The concordance was analysed at synset level and (word,gender) tuple level to identify co-occurrence across dictionaries and also to identify unique contributions of each dictionary.

6.1 Statistics of the synsets

1. Total number of synsets : 5489

2. Total synsets in individual dictionaries: AMAR – 3692, ABCH – 4791, ACPH – 347, ACSJ – 534, ACVH – 591

3. Synset co-occurrence across dictionaries - See Table 3

6.2 Statistics about words

1. Total number of Word-gender pairs in concordance: 20975

2. Word-gender pairs in individual dictionaries: AMAR – 10157, ABCH – 14704, ACPH – 1506, ACSJ – 693, ACVH - 949

3. Word-gender pair co-occurrence across dictionaries - See Table 4

7 Way forward

The present concordance needs to be extended with the help of various Samānārthakakośas so that all words attested to be having the same meaning can be seen at a single place.

A systematic analysis of Anekārthakakośas may aslo be undertaken to see if there are cases where there are knowledge gaps between Samānārthakakośas and Anekārthakakośas. There is a possibility that the Samānārthakakośas may only enlist meanings which have two or more synonyms. Anekārthakakośas may be more exhaustive in coverage in such cases.

An analysis of dictionaries compiled in different era may throw some light on the changes in the meaning of words over a period of time or addition of new words over a period of time.

Amarakośa	Abhidhānacinitāmaṇi
No corresponding entries, as devādhidevakāṇḍa is dedicated to Jaina deities	1. devādhidevakāṇḍa (ślokas 1-86)
1.1. svargavargaḥ	2. devakāṇḍaḥ (ślokas 87-126, 171-240)
1.2. vyomavargaḥ	2. devakāṇḍaḥ (śloka 163)
1.3. digvargaḥ	2. devakāṇḍaḥ (ślokas 164-170)
1.4. kālavargaḥ	2. devakāṇḍaḥ (ślokas 127-162)
1.5. dhīvargaḥ	2. devakāṇḍaḥ (ślokas 308-324)
1.6. śabdādivargaḥ	2. devakāṇḍaḥ (ślokas 241-278)
1.7. nāṭyavargaḥ	2. devakāṇḍaḥ (ślokas 278-307, 325-336)
1.8. pātālabhogivargaḥ	4.5.1. sthalacarapañcendriyāḥ (ślokas 1302-1315)
1.9. narakavargaḥ	5. nārakakāṇḍaḥ (ślokas 1358-1364)
1.10. vārivargaḥ	4.1.2. ekendriyeṣu apkāyaḥ (ślokas 1069-1096)
2.1. bhūmivargaḥ	4.1.1. ekendriyeṣu pṛthvīkāyaḥ (ślokas 935-970)
2.2. puravargaḥ	4.1.1. ekendriyeṣu pṛthvīkāyaḥ (ślokas 971-1026)
2.3. śailavargaḥ	4.1.1. ekendriyeṣu pṛthvīkāyaḥ (ślokas 1027-1036)
2.4. vanauṣadhivargaḥ	4.1.5. ekendriyeṣu vanaspatikāyaḥ (ślokas 1110-1201)
2.5. siṃhādivargaḥ (ślokas 1-12)	4.5.1. sthalacarapañcendriyāḥ (ślokas 1283-1301)
2.5. siṃhādivargaḥ (ślokas 13)	4.2. dvīndriyāḥ (ślokas 1202-1206)
-	4.3. trīndriyāḥ (ślokas 1206-1209)
2.5. siṃhādivargaḥ (ślokas 13-14)	4.4. caturindriyāḥ (ślokas 1210-1216)
2.5. siṃhādivargaḥ (ślokas 14-43)	4.5.2. khacarapañcendriyāḥ (ślokas 1316-1343)
1.1. svargavargaḥ (ślokas 53-57)	4.1.3. ekendriyeṣu tejaskāyaḥ (ślokas 1097-1105)
1.1. svargavargaḥ (ślokas 61-63)	4.1.4. ekendriyeṣu vāyukāyaḥ (ślokas 1106-1109)
1.10. vārivargaḥ (ślokas 17-25)	4.5.2. jalacarapañcendriyāḥ (ślokas 1343-1357)
2.6. manuṣyavargaḥ	3. martyakāṇḍaḥ (ślokas 337-688)
2.7. brahmavargaḥ	3. martyakāṇḍaḥ (ślokas 807-863)
2.8. kṣattriyavargaḥ	3. martyakāṇḍaḥ (ślokas 689-806)
2.9. vaiśyavargaḥ	3. martyakāṇḍaḥ (ślokas 864-894)
2.10. śūdravargaḥ	3. martyakāṇḍaḥ (ślokas 894-934)
3.1. viśeṣyanighnavargaḥ	Spread throughout the work
3.2. saṅkīrṇavargaḥ	Spread throughout the work
3.3. nānārthavargaḥ	Absent, as there is a separate work authored by Hemacandra for Nānārtha.
3.4. avyayavargaḥ	7. avyayāni (ślokas 1525-1542)
3.5. liṅgādisaṃgrahavargaḥ	Absent, as there is a separate work authored by Hemacandra for Liṅgānuśāsanam.
Spread throughout the work	6. sāmānyakāṇḍaḥ (ślokas 1365-1524)
2.9. vaiśyavargaḥ (ślokas 91-111)	4.1.1. ekendriyeṣu pṛthvīkāyaḥ (ślokas 1036-1068) - metals and minerals
2.8. kṣattriyavargaḥ (ślokas 34-50)	4.1.5. sthalacarapañcendriyāḥ (ślokas 1217-1252) - entries related to elephant and horse
2.9. vaiśyavargaḥ (ślokas 50-77)	4.1.5. sthalacarapañcendriyāḥ (ślokas 1253-1283) - domesticated animals

Table 2: Mapping of knowledge representation between Amarakośa and Abhidhānacintāmaṇi

AMAR	ABCH	ACPH	ACSJ	ACVH	Total
Y					609
	Y				1556
		Y			66
			Y		5
				Y	0
Y	Y				2382
Y		Y			17
Y			Y		0
Y				Y	1
	Y	Y			29
	Y		Y		22
	Y			Y	46
		Y	Y		0
Y	Y	Y			134
Y		Y	Y		0
Y	Y		Y		64
Y	Y			Y	75
	Y	Y	Y		0
	Y		Y	Y	66
	Y	Y		Y	3
Y	Y	Y	Y		14
Y	Y		Y	Y	316
Y	Y	Y		Y	37
Y	Y	Y	Y	Y	43
	Y	Y	Y	Y	4
Total					5489

Table 3: Synset co-occurrence across samānārthaka kośas

AMAR	ABCH	ACPH	ACSJ	ACVH	Total
Y					3527
	Y				8239
		Y			1459
			Y		299
				Y	500
Y	Y				6412
Y		Y			32
Y			Y		26
Y				Y	80
	Y	Y			1
	Y		Y		17
	Y			Y	11
		Y	Y		1
		Y		Y	3
			Y	Y	287
Y	Y	Y			6
Y		Y	Y		1
Y		Y		Y	2
Y	Y		Y		6
Y	Y			Y	10
Y			Y	Y	53
		Y	Y	Y	1
Y	Y		Y	Y	2
Total					20975

Table 4: Word-gender pair co-occurrence across samānārthaka kośas

8 Conclusion

The concordance of samānārthaka kośa provides a useful tool to see all the synonyms of a given concept (synset) at a single point. This can be expanded with inclusion of more and more samānārthaka kośas.

An analysis of the concordance provides a fertile ground to research the differences of gender, word form or addition of new words to the synset by different lexicographers and also in different time periods.

References

CDSD. 2024. *Cologne Digital Sanskrit Dictionaries.* Cologne University, Cologne. version 2.7.91, accessed on 30 January 2024, at https://www.sanskrit-lexicon.uni-koeln.de.

Devasāgaragaṇi. V.S. 2054. *Kalikālasarvajñaśrīhemacandrācāryaviracitā Vācakaśrīdevasāgaragaṇikṛtavyutpattiratnākarakalitā Abhidhānacintāmaṇināmamālā.* Śrī Rāndera Roḍa Jaina Saṅghaḥ.

Halāyudha. V.S. 2014. *Halāyudhakośaḥ (Abhidhānaratnamālā).* Sarasvatī Bhavana, Vārāṇasī.

Hemacandra. 1896. *Abhidhānacintāmaṇipariśiṣṭa. The Abhidhana-saṅgraha or A Collection of Sanskrit Ancient Lexicons. Book No. 7.* Nirṇayasāgara Press, Bombay.

Hemacandra. V.S. 2032a. *Abhidhānacintāmaṇi. Abhidhānacintāmaṇināmamālā (svopajñasārthaśabdānukramaṇikābhyāṃ samalaṅkṛtā), Śrī vṛddhi-nemi-amṛta granthamālā-granthāṅka 72.* Śrī Jaina Sāhitya Vardhaka Sabhā, Ahmedabad (Gujarat).

Hemacandra. V.S. 2032b. *Abhidhānacintāmaṇisvopajñavṛtti. Abhidhānacintāmaṇināmamālā (svopajñasārthaśabdānukramaṇikābhyāṃ samalaṅkṛtā), Śrī vṛddhi-nemi-amṛta granthamālā-granthāṅka 72.* Śrī Jaina Sāhitya Vardhaka Sabhā, Ahmedabad (Gujarat).

Jinadevasūri. 1896. *Abhidhānacintāmaṇiśiloñcha. The Abhidhana-saṅgraha or A Collection of Sanskrit Ancient Lexicons. Book No. 11.* Nirṇayasāgara Press, Bombay.

Malhar Kulkarni, Chaitali Dangarikar, Irawati Kulkarni, Abhishek Nanda, and Pushpak Bhattacharyya. 2010. Introducing sanskrit wordnet. In *Proceedings on the 5th global wordnet conference (GWC 2010), Narosa, Mumbai*, pages 287–294.

Sivaja S. Nair. 2011. *The knowledge structure in Amarakośa.* University of Hyderabad, Hyderabad.

Śrīvallabha. 1974. *Jinadevasūri's Haimanāmamālāśiloñcha with a commentary by Śrivallabha. L. D. Series 46.* L. D. Institute of Indology, Ahmedabad.