

Word Sense Alignment of Sanskrit Lexica

A thesis submitted during 2026 to the
University of Hyderabad
in partial fulfilment of the award of a Ph.D. degree
in

Sanskrit Studies

by

Dhavalkumar Kiritkumar Patel

21HSPH01



Department of Sanskrit Studies

School of Humanities

University of Hyderabad

(P.O.) Central University

Gachibowli, Hyderabad – 500046

Telangana

India



CERTIFICATE

This is to certify that the thesis entitled **Word Sense Alignment of Sanskrit Lexica** submitted by **Dhaval Kuamr Kiritkumar Patel**, bearing Registration Number **21HSPH01** in partial fulfilment of the requirements for award of **Doctor of Philosophy** in the **School of Humanities** is a bonafide work carried out by him/her under my supervision and guidance.

This thesis is free from Plagiarism and has not been submitted previously in part or in full to this or any other University or Institution for award of any degree or diploma.

Further, the student has the following publication(s) before submission of the thesis/ monograph for adjudication and has produced evidence for the same in the form of acceptance letter or the reprint in the relevant area of his research: (Note: At least one publication in referred journal is required)

1. 'Sāṅkhyatattvapradīpa of Tulasīrāma – an unpublished Sāṅkhya work'. शोध-प्रभा, ४७ वर्षे प्रथमोऽङ्कः (जनवरी-मार्च २०२२ ई.), pp. 69–104. (ISSN: 0974-8946)
2. 'Word Sense Alignment of Sanskrit Lexica'. *Proceedings of the 7th International Sanskrit Computational Linguistics Symposium*, edited by Arnab Bhattacharya, Association for Computational Linguistics, 2024, pp. 1–13. aclanthology.org/2024.iscls-1.1/. (ISBN: 979-8-89176-027-1)
3. 'Concordance of Sanskrit Synonyms'. *Computational Sanskrit and Digital Humanities - World Sanskrit Conference 2025*, edited by Amba Kulkarni and Oliver Hellwig, Association for Computational Linguistics, 2025, pp. 181–188, aclanthology.org/2025.wsc-csdh.12/. (ISBN: 979-8-89176-250-3)
4. 'साङ्ख्यकारिकायां सङ्गतितन्त्रयुक्तयोः प्रायोगिकविचारः (Usage of Saṅgati and Tantrayukti in Sāṅkhyakārikā)'. प्राची प्रज्ञा, accepted to be published in 22nd issue scheduled to be published in June 2026. (ISSN: 2348-8417)

and

has made presentations in the following conferences:

1. 7th International Sanskrit Computational Linguistics Symposium, Auroville, February 2024 (International)
2. 19th World Sanskrit Conference, Kathmandu, June 2025 (International)

Further, the student has passed the following courses towards fulfilment of coursework requirement for Ph.D.:

Course	Code Name	Credits	Pass/Fail
1.	Topics in Natural Language Processing (SK812)	4	Pass
2.	Dissertation Related Readings (SK830)	4	Pass
3.	Research and Publication Ethics (EN875)	2	Pass
4.	Research Methodology (SK801)	4	Pass

Ambep
22.1.2026
Supervisor
Department of Sanskrit Studies
School of Humanities
University of Hyderabad
P.O. Central University
Hyderabad-500 046.TS.

Ambep
22.1.2026
Head of the Department
HEAD
Department of Sanskrit Studies
School of Humanities
University of Hyderabad
Central University P.O.
HYDERABAD-500 046

(Ansari)
22/01/26
Dean of School
DEAN
SCHOOL OF HUMANITIES
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046. T.S. INDIA

DECLARATION

I, **Dhaval Kumar Kirit Kumar Patel**, hereby

Declare that this thesis entitled "**Word Sense Alignment of Sanskrit Lexica**"

Submitted by me under the guidance and supervision of **Dr. Amba Kulkarni** is a bonafide research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma. I hereby agree that my thesis can be deposited in Shodhganga/INFLIBNET. A report on plagiarism statistics from the University Librarian is enclosed.

Date: 06 February 2026

Place: Hyderabad



Dhaval Kumar Kirit Kumar Patel

21HSPH01

Relevance of the dissertation to the United Nations Sustainable Development Goals (SDGs)

Quality education of any language requires access to lexical resources. The present thesis explores the possibility of alignment of more than one lexical resources at the word sense level. This enriches the lexical resources and is especially helpful in resource-starved languages like Sanskrit. With availability of algorithm for semi-automatic linking of lexical resources at word sense level, the word sense disambiguated corpus in any language can be extended to any other language. Even though the research was primarily focussed on Sanskrit dictionaries, the algorithm is language agnostic, which makes it possible to link any two language pairs irrespective of origin and destination languages.

The work also generated a concordance of synonyms in Sanskrit language. It is extensible in future to multiple synonymic dictionaries. This concordance will also provide the users with multiple words to express the same concept. This may be extremely useful in case of poetry, where metrical exigencies constrain the words that can be used. It also helps researchers to understand the addition or deletion of some words from the vocabulary over a period of time. It also provides a useful comparison about genders of specific words.

The present thesis provides the tools for improvement and interlinking of lexical resources of Sanskrit language particularly and of any language

generally. The researcher hopes that the availability of better lexical resources will help improve the quality of education. The findings presented in the thesis are generic in nature and can be extended to any language.

Acknowledgements

I am thankful to everyone who has made this research journey possible.

I am deeply indebted to my supervisor, Prof. Amba Kulkarni, for her guidance throughout my research activity. She has been a very sharp teacher who always asked the right questions and prodded the research activity further. It has been my good fortune to be able to study under her able guidance. She was always there on every step of this research journey. I must also thank her for her timely, insistent mails which ensured that I submitted the research papers in time and prepare my thesis in a timely manner. She has been a great mentor.

I also extend my gratitude to Prof. J. S. R. Prasad for guiding me towards Indian systems of research methodology namely tantrayuktis and saṅgatis.

I am also thankful to the office staff for their help.

I dedicate this thesis to my wife Hiral who has been a constant source of encouragement in this endeavour. There were days when I would avoid the research work, but she would bring me back on the track. Her perseverance is the reason that this thesis is seeing the light of the day.

I owe an apology to my kids Dhrishat and Sugat, who were deprived of the time they deserved from me and I spent that in this research. My family has been a constant source of happiness, a spring on which I can count.

Table of Contents

1. DECLARATION.....	3
2. Relevance of the dissertation to the United Nations Sustainable Development Goals (SDGs).....	4
3. Acknowledgements.....	6
4. Articles Published During the Course.....	15
5. Chapter 1 - Overview.....	16
1.1. Goal of thesis.....	16
1.2. Types of Sanskrit Lexica.....	16
1.2.1. Samānāṛthaka kośas.....	17
1.2.2. Anekāṛthaka kośas.....	20
1.2.3. Sanskrit-Sanskrit dictionaries structured on Western principles.....	21
1.2.4. Digital Sanskrit Lexical Resources.....	22
1.3. Motivation.....	25
6. Chapter 2 Word Sense Alignment Problem.....	27
2.1. Lexical Ambiguity.....	27
2.2. Homonymy.....	27
2.3. Polysemy.....	28
2.4. Word senses and granularity of dictionaries.....	28
2.5. Word Sense Disambiguation.....	31
2.6. Word Sense Alignment.....	35
2.7 Use case for Word Sense Alignment.....	37
2.8. Challenges in Word Sense Alignment.....	38

2.9. Problem Statement.....	39
7. Chapter 3 Methodology.....	40
3.1. Difficulty for resource-starved languages.....	40
3.2. Existing algorithms.....	41
3.2.1. Content Similarity Measures.....	42
3.2.1.1. Text Similarity Measures.....	42
3.2.1.1.2. Longest Common Subsequence.....	43
3.2.1.1.3. Greedy String Tiling.....	43
3.2.1.1.4. Jaro.....	44
3.2.1.1.5. Jaro–Winkler.....	44
3.2.1.1.6. Monge–Elkan.....	45
3.2.1.1.7. Levenshtein.....	45
3.2.1.1.8. Term Frequency–Inverse Document Frequency (TF–IDF)....	46
3.2.1.1.9. Word n-grams.....	46
3.2.1.1.10. Jaccard Coefficient.....	47
3.2.1.1.11. Character n-gram Profiles.....	47
3.2.1.2. Semantic Similarity Measures.....	47
3.2.1.2.1. Jiang and Conrath.....	48
3.2.1.2.2. Lin.....	49
3.2.1.2.3. Resnik.....	50
3.2.1.3. Statistical Techniques.....	50
3.2.1.3.1. Latent Semantic Analysis (LSA).....	51
3.2.1.3.2. Explicit Semantic Analysis (ESA).....	51
3.2.2. Structural Similarity Measures.....	52

3.2.2.1. Stopword n-grams.....	52
3.2.2.2. Part-of-speech n-grams.....	53
3.2.2.3. Pair order.....	53
3.2.3. Stylistic Similarity Measures.....	54
3.2.3.1. Type-Token Ratio (TTR).....	54
3.2.3.2. Sequential TTR.....	54
3.2.3.3. Sentence length.....	55
3.2.3.4. Token length.....	55
3.2.4. Discussion about existing algorithms.....	56
3.3. Selection of dictionary pairs.....	57
3.4. Word Sense Alignment (WSA) of Sanskrit-English and Sanskrit-Hindi dictionaries of Apte.....	58
3.5. Gold Standard Data.....	59
3.6. Algorithms.....	63
3.6.1. Unit of comparison.....	63
3.6.2. Measure of comparison.....	64
3.6.3. Measures of similarity.....	66
3.6.4. Thresholds.....	66
3.7. Evaluation on Gold Standard Data.....	67
3.7.1. Identification percentage.....	67
3.7.2. Precision, Recall and F1-score.....	69
3.7.3. Cochran's Q Test.....	71
3.7.4. Post-hoc Pairwise Test: McNemar's Test.....	73
3.8. Merits of the usage of two languages as target languages.....	77

3.8.1. Cases where one language fares better than another.....	79
3.8.2. Confidence levels with two target languages.....	80
3.9. Analysis of gold standard data.....	82
3.10. Analysis of ASH and AP90 full dictionary mapping.....	83
3.11. Mapping other dictionaries.....	85
3.11.1. Apte Sanskrit-English and Monnier Williams Sanskrit English dictionaries.....	86
3.11.2. Comparison between Wilson and Yates dictionaries.....	87
3.11.3. Comparison between Wordnet and Webster dictionaries.....	87
3.11.4. Comparative analysis of different dictionary pairs.....	89
3.12. Experiments with graph-based methods for finding similarity.....	90
3.13. Latent Semantic Analysis.....	98
3.14. Frontend for correction submission.....	100
8. Chapter 4 Synset concordance of Samānāṛthaka Sanskrit kośas.....	103
4.1. Amarakośa.....	103
4.2. Abhidhānacintāmaṇi of Hemacandra.....	104
4.3. Abhidhānacintāmaṇipariśiṣṭa and Abhidhānacintāmaṇiśiloṇcha.....	105
4.4. Reason for selection of these kośas.....	105
4.5. Limitation of existing digital resources.....	106
4.6. Development of e-resource.....	107
4.6.1. Generation of lexicon marked up with synset information.....	107
4.6.2. Alignment of synsets across lexica.....	110
4.7. Structure of information in Amarakośa.....	120
4.8. Structure of information in Abhidhānacintāmaṇi.....	121

4.9. Comparison between the knowledge representation of Amarakośa and Abhidhānacintāmaṇi.....	122
4.10. Statistics of the concordance.....	125
4.11. Differences between AMAR, ABCH, ACPH, ACSJ and ACVH dictionaries	133
4.12. The difference between lexicographers.....	133
9. Chapter 5 Conclusion and Way Ahead.....	135
10. Annexure 1 Gender differences between AMAR and ABCH.....	138
11. Annexure 2 The difference between lexica is only due to 'ka' at the end	164
12. Annexure 3 Analysis of concordance between AMAR, ABCH, ACPH, ACSJ and AVCH.....	169
13. Annexure 4 Additional information provided by lexicographers on specific words.....	188
14. Bibliography.....	236

Index of Tables

Table 1: Sample ASH entries.....	60
Table 2: Sample AP90 entries.....	60
Table 3: Gold Standard Data for Alignment of ASH and AP90.....	61
Table 4: Sample AP90 entries along with their Hindi translations.....	62
Table 5: Sample ASH entries with their English translations.....	62
Table 6: Metrics used for evaluation.....	66
Table 7: Percentage of word sense pairs correctly identified from gold standard data by already existing algorithms.....	67
Table 8: Percentage of word sense pairs correctly identified from gold standard data by various algorithms.....	68
Table 9: Precision, Recall and F1-score of various algorithms on Gold Standard data.....	71
Table 10: Cochran's Q Test results on Gold Standard data.....	73
Table 11: McNemar Test Results of CR6 algorithm versus all other algorithms.	76
Table 12: Entry of the word 'अहार्य' in the ASH (D1) dictionary.....	78
Table 13: Entry of the word 'अहार्य' in the AP90(D2) dictionary.....	78
Table 14: Entry of ASH translated to English via Google Translate (E).....	79
Table 15: Entry of AP90 translated to Hindi via Google Translate (H).....	79
Table 16: Categorisation of various mappings along with their confidence level	81
Table 17: Categorisation of gold standard mapping created via algorithm CR6	82

Table 18: Categorisation of word sense mappings generated by algorithm CR6 for Apte Sanskrit-Hindi and Apte Sanskrit-English dictionaries.....	84
Table 19: False positives from randomly selected mappings.....	85
Table 20: Categorisation of mapping between Apte Sanskrit-English and Monnier Williams dictionaries.....	86
Table 21: Categorisation of mapping between Wilson and Yates dictionaries....	87
Table 22: Categorisation of mapping between Wordnet and Webster dictionaries	88
Table 23: Categorisation of mappings for various dictionary pairs.....	89
Table 24: Precision, Recall and F1-Score for Latent Semantic Analysis and BoW, Doc2Vec, Dense Embeddings.....	99
Table 25: Precision, Recall and F1-score for synonymic dictionaries.....	112
Table 26: Sample entry of the samānāṛthaka concordance.....	119
Table 27: Mapping of knowledge representation between Amarakośa and Abhidhānacintāmaṇi.....	125
Table 28: Synset co-occurrence across samānāṛthaka kośas.....	128
Table 29: Word-gender pair co-occurrence across samānāṛthaka kośas.....	131

Index of Figures

Figure 1: Graph for the word “kamala”	94
Figure 2: graph for the word “amBoja”	95
Figure 3: Combined graph generated by amalgamating the graphs for “kamala” and “amBoja” with English words as a bridge between graphs.....	96
Figure 4: Visual aid to the annotator alongwith the similarity scores and deltas	100
Figure 5: Frontend for correction submission.....	101
Figure 6: Edwards-Venn diagram showing synset co-occurrence across dictionaries.....	129
Figure 7: Edwards-Venn diagram showing (word, gender) pair co-occurrence across dictionaries.....	132

Articles Published During the Course

1. 'Sāṅkhyatattvapradīpa of Tulasīrāma – an unpublished Sāṅkhya work'. शोध-प्रभा, ४७ वर्षे प्रथमोऽङ्कः (जनवरी-मार्च २०२२ ई.), pp. 69–104. (ISSN: 0974-8946)
2. 'Word Sense Alignment of Sanskrit Lexica'. *Proceedings of the 7th International Sanskrit Computational Linguistics Symposium*, edited by Arnab Bhattacharya, Association for Computational Linguistics, 2024, pp. 1–13. aclanthology.org/2024.iscls-1.1/. (ISBN: 979-8-89176-027-1)
3. 'Concordance of Sanskrit Synonyms'. *Computational Sanskrit and Digital Humanities - World Sanskrit Conference 2025*, edited by Amba Kulkarni and Oliver Hellwig, Association for Computational Linguistics, 2025, pp. 181–188, aclanthology.org/2025.wsc-csdh.12/. (ISBN: 979-8-89176-250-3)
4. 'साङ्ख्यकारिकायां सङ्गतितन्त्रयुक्त्योः प्रायोगिकविचारः (Usage of Saṅgati and Tantrayukti in Sāṅkhyakārikā)'. प्राची प्रज्ञा, accepted to be published in 22nd issue scheduled to be published in June 2026. (ISSN: 2348-8417)

Chapter 1 - Overview

1.1. Goal of thesis

Word sense alignment is a field of study in which lexical resources or texts are aligned at the level of word sense rather than the word. The present thesis tries to evaluate the possibility of mechanically aligning conventional monolingual/bilingual Sanskrit lexica at the level of word sense computationally. Sanskrit also has dictionaries of synonymous words (more analogous to Thesauri). As the knowledge representation in this type of dictionary is different from conventional dictionaries, a different methodology is required to align such dictionaries. The present thesis also aims to create a synonym set concordance for such dictionaries, namely Amarakośa (AMAR), Abhidhānacintāmaṇi (ABCH), Abhidhānacintāmaṇipariśiṣṭa (ACPH), Abhidhānacintāmaṇiśiloṇcha (ACSJ) and Abhidhānacintāmaṇisvopajñavṛtti (ACVH).

1.2. Types of Sanskrit Lexica

There are two types of lexical resources available in the majority of Western languages, including English.

The first type is known as a dictionary, which has words of a given language arranged in alphabetical order and enlists different meanings for the same word. Merriam-Webster (2025) defines a 'dictionary' as 'a reference source in print or electronic form containing words usually alphabetically arranged along with information about their forms, pronunciations, functions, etymologies, meanings, and syntactic and idiomatic uses'.

The second lexical resource is known as a thesaurus. Merriam-Webster (2025) defines a 'thesaurus' as 'a book of words or of information about a particular field or set of concepts. Especially: a book of words and their synonyms'. A thesaurus has a collection of words based on some semantic relationship, mostly a synonymic relationship. In a thesaurus, different words connoting the same meaning are clubbed together at a place.

Sanskrit, an ancient Indian language, has been a medium of transmission of knowledge in various fields of study for centuries. Compilation of word lists in Sanskrit commenced at an early date, as it was found necessary to access the old literature, such as Vedic literature, while the language was undergoing some transformations with meaning shifts. The lexical resources known as kośas were developed. They are of two types: (1) Samānāṛthaka kośas and (2) Anekāṛthaka kośas.

1.2.1. Samānāṛthaka kośas

Samānāṛthaka kośas (synonymic dictionaries) list the synonyms together. Synonymic dictionaries are systematic catalogues of words with the same meaning (ekārtha, samānārtha) (Vogel 2015). The synonyms are arranged following some theme, semantic criterion, or ontological classification scheme. For example, in the most famous samānāṛthaka kośa viz. Amarakośa, the words are arranged in three kāṇḍas and further within the kāṇḍas, the headwords are arranged based on either semantic or ontological properties.

Important Samānāṛthaka kośas in Sanskrit literature are as follows (Vogel 2015).

1. Vācaspati's Śabdārṇava (Unavailable)
2. Vyāḍi's Utpalinī (Unavailable)
3. Vikramāditya's Saṃsārāvarta (Unavailable)
4. Amarasiṃha's Nāmaliṅgānuśāsana
5. Halāyudha's Abhidhānaratnamālā
6. Yādavaprakāśa's Vaijayantī (Both synonymic and homonymic)
7. Bhoja's Nāmamālikā
8. Rabhasapāla's Rabhasakośa (Both synonymic and homonymic)
9. Dhanañjaya's Nāmamālā
10. Dhanañjaya's Paryāyenduśekhara
11. Puruśottamadeva's Trikāṇḍaśeṣa
12. Puruśottamadeva's Hārāvalī
13. Hemacandra's Abhidhānacintāmaṇi
14. Hemacandra's Nighaṇṭuśeṣa

15. Śrīdharasena's Viśvalocana or Muktāvalī (Both synonymic and homonymic)
16. Padmanābhadatta's Bhūriprayoga (Both synonymic and homonymic)
17. Vāmanabhaṭṭa Bāṇa's Śabdaratnākara (Both synonymic and homonymic)
18. Vāmanabhaṭṭa Bāṇa's Śabdacandrikā
19. Makarandadāsa's Paramānandīyanāmamālā (Both synonymic and homonymic)
20. Harṣakīrti's Śārādīyākhyānāmamālā
21. Rūpacandra's Rūpamañjarīnāmamālā (Both synonymic and homonymic)
22. Puṇḍarīka Viṭṭhala's Śīghrabodhinīnāmamālā (Both synonymic and homonymic)
23. Appaya Dīkṣita's Nāmasaṃgrahamālā
24. Sahajakīrti's Siddhaśabdārṇava
25. Viśvanātha's Koṣakalpataru (Both synonymic and homonymic)
26. Veṇidatta's Pañcatattvaparakāśa
27. Mīramīrāsuta's Asālatiprakāśa
28. Keśava's Kalpadrukośa
29. Mathureśa's Śabdaratnāvalī (Both synonymic and homonymic)

1.2.2. Anekārthaka kośas

Anekārthaka kośas list different meanings of a given word. The words may or may not have any alphabetic arrangement. Both kinds of kośas were meant to be memorised, applied to texts, and cited as and when the usage of the said word in the literature was to be justified in a commentary. Therefore, the kośas were almost invariably in a verse form. Vogel (2015) has given a comprehensive coverage of these Sanskrit kośas and commentarial literature thereon.

Important Anekārthaka kośas in Sanskrit literature are as follows (Vogel 2015)

1. Śāśvata's Anekārthasamuccaya
2. Mahākṣapaṇaka's Anekārthadhvanimañjarī
3. Yādavaprakāśa's Vaijayantī (Both synonymic and homonymic)
4. Rabhasapāla's Rabhasakośa (Both synonymic and homonymic)
5. Ajayapāla's Nānārthasaṃgraha
6. Dharaṇidāsa's Anekārthasāra
7. Maheśvara's Viśvaprakāśa
8. Puruśottamadeva's Ekākṣarakośa
9. Puruśottamadeva's Dvirūpakośa
10. Maṇkha's Anekārthakośa
11. Hemacandra's Anekārthasaṃgraha
12. Keśavasvāmin's Nānārthārṇavasamkṣepa
13. Medinikara's Nānārthaśabdakośa

14. Śrīdharasena's Viśvalocana or Muktāvalī (Both synonymic and homonymic)
15. Padmanābhadatta's Bhūriprayoga (Both synonymic and homonymic)
16. Irugapa's Nānārtharatnamālā
17. Rāghava's Nānārthamañjarī
18. Vāmanabhaṭṭa Bāṇa's Śabdaratnākara (Both synonymic and homonymic)
19. Makarandadāsa's Paramānandīyanāmamālā (Both synonymic and homonymic)
20. Harṣakīrti's Anekārthanāmamālā
21. Harṣakīrti's Śabdānekārtha
22. Rūpacandra's Rūpamañjarīnāmamālā (Both synonymic and homonymic)
23. Puṇḍarīka Viṭṭhala's Śīghrabodhinīnāmamālā (Both synonymic and homonymic)
24. Viśvanātha's Koṣakalpataru (Both synonymic and homonymic)
25. Mathureśa's Śabdaratnāvalī (Both synonymic and homonymic)
26. Śāhajī's Śabdaratnasamanvaya
27. Śāhajī's Śabdārthasaṃgraha

1.2.3. Sanskrit-Sanskrit dictionaries structured on Western principles

Because of the influence of Western lexicography, a few Sanskrit-Sanskrit dictionaries like Vācaspatyam and Śabdakalpadruma were also compiled on the lines with the Western methodology of arranging headwords alphabetically and in prose form. Both these dictionaries cite various previous kośas which had been available to the editors of these dictionaries at the relevant time. Several bilingual Sanskrit dictionaries, such as Sanskrit-English, Sanskrit-French, and Sanskrit-German, were compiled starting from the early 19th century.

1.2.4. Digital Sanskrit Lexical Resources

Almost all of the major dictionaries that are free from copyright are available on the Cologne Digital Sanskrit Dictionaries website (CDSD, 2023). This digitised data has various levels of markup. Some dictionaries are extensively marked up to identify various senses of a given word mechanically. Whereas some of the dictionaries, like Vācaspatyam and Śabdakalpadruma, are not marked up properly. Therefore, it is not possible to mechanically use these two dictionaries without spending a huge time improving the markup.

In recent years, Huet (2019) has developed a digital Sanskrit-French dictionary where the lexical items are directly linked to the inflectional and derivational morphology.

A list of Sanskrit lexical resources available at CDS is placed below, along with the year of their publication.

Sanskrit-English Dictionaries		
ID	date	Dictionary
WIL	1832	Wilson Sanskrit-English Dictionary
YAT	1846	Yates Sanskrit-English Dictionary
GST	1856	Goldstücker Sanskrit-English Dictionary
BEN	1866	Benfey Sanskrit-English Dictionary
MW72	1872	Monier-Williams Sanskrit-English Dictionary
LAN	1884	Lanman's Sanskrit Reader Vocabulary
LRV	1889	Vaidya Sanskrit-English Dictionary
AP90	1890	Apte Practical Sanskrit-English Dictionary
CAE	1891	Cappeller Sanskrit-English Dictionary
MD	1893	Macdonell Sanskrit-English Dictionary
MW	1899	Monier-Williams Sanskrit-English Dictionary
SHS	1900	Shabda-Sagara Sanskrit-English Dictionary
AP	1957	Practical Sanskrit-English Dictionary, revised edition
PD	1976	An Encyclopaedic Dictionary of Sanskrit on

		Historical Principles
English-Sanskrit Dictionaries		
MWE	1851	Monier-Williams English-Sanskrit Dictionary
BOR	1877	Borooah English-Sanskrit Dictionary
AE	1920	Apte Student's English-Sanskrit Dictionary
Sanskrit-French Dictionaries		
BUR	1866	Burnouf Dictionnaire Sanscrit-Français
STC	1932	Stchoupak Dictionnaire Sanscrit-Français
Sanskrit-German Dictionaries		
PWG	1855	Böhtlingk and Roth Großes Petersburger Wörterbuch
GRA	1873	Grassmann Wörterbuch zum Rig Veda
PW	1879	Böhtlingk Sanskrit-Wörterbuch in kürzerer Fassung
CCS	1887	Cappeller Sanskrit Wörterbuch
SCH	1928	Schmidt Nachträge zum Sanskrit-Wörterbuch
Sanskrit-Latin Dictionaries		
BOP	1847	Bopp Glossarium Sanscritum
Sanskrit-Sanskrit Dictionaries		
VCP	1873	Vācaspatyam

SKD	1886	Śabdakalpadruma
ARMH	1861	Abhidhānaratnamālā of Halāyudha
ABCH	1896	Abhidhānacintāmaṇi of Hemacandra
ACPH	1896	Abhidhānacintāmaṇipariśiṣṭa of Hemacandra
ACSJ	1896	Abhidhānacintāmaṇiśiloṇcha of Jinadeva

Out of these dictionaries, PD is digitised, but not available publicly because of copyright restrictions.

1.3. Motivation

Some of these dictionaries, in addition to providing the meaning of Sanskrit words in the target language, also provide citations from Sanskrit texts. The citations in different dictionaries vary. These citations play an important role in understanding the context in which the sense is being used. Aligning the senses of different dictionaries would provide us with more than one example sentence for each sense to understand the context and the semantic criterion that decides the sense of the word in a given usage.

Further, with the availability of word embeddings for words in several languages, such as Hindi, English, French, German, etc., if the senses in

Sanskrit bilingual dictionaries are aligned, one can take advantage of the existing modelling of the world knowledge and the domain knowledge of other languages to disambiguate Sanskrit words. Such sense mapping would be useful in the Machine Translation system, for Information Retrieval, and even for a casual learner of the Sanskrit language. This motivated us to look at the problem of aligning various Sanskrit bilingual dictionaries according to the senses.

In what follows, we first explain the word sense alignment problem and the challenges therein. This is followed by the discussion on the methodology followed for automatic sense alignment. Thereafter, we discuss the sense alignment between the Sanskrit-English and Sanskrit-Hindi dictionaries by Apte. The results of the alignment algorithms are further extended to other pairs of dictionaries. Finally, we discuss other possible ways of alignment.

As far as Samānāṛthaka kośas are concerned, their alignment would allow us to find out an extensive list of words used to depict a concept. It would be particularly useful to choose a word from a given synset that satisfies given constraints, for example, fitting a metrical exigency. The acquisition of new words in the synset over a period of time would provide a fertile ground for research. The differences of gender and spelling across lexica would also be an interesting research topic.

Chapter 2

Word Sense Alignment Problem

2.1. Lexical Ambiguity

Lexical ambiguity refers to words that share the same orthography but have different meanings (word senses) (Stokoe 2005). It can be subdivided into two distinct types, homonymy and polysemy.

2.2. Homonymy

Homonymy describes when two senses of a given word (or derivation) are distinct. Typically, they are separated by etymology and are therefore entirely unrelated in meaning. One classic example (Kilgariff, 1992) is 'bat' as in an airborne mammal (from the Middle English word 'bakke' meaning flying rodent) vs. 'bat' as in an instrument used in the game of cricket (from the Celtic for stick or cudgel). There is no underlying relationship between these two meanings, which have come about independently from different root languages (Stokoe 2005). In Sanskrit language, an example is the headword

“aka (अक)”. It has two meanings (1) अक from the suffix अकच् and (2) अ-क; unhappiness, pain, trouble.

2.3. Polysemy

Polysemy describes where two senses of a word are related in that they share membership of a subsuming semantic classification (Stokoe 2005). Consider the word ‘mouth’ as in a part of the body vs. ‘mouth’ as in the outlet of a river. Both meanings are subsumed by a higher concept (in this case, they both describe an opening).

2.4. Word senses and granularity of dictionaries

Homonymy and polysemy are differentiated in most dictionaries by the major (homonyms) and minor (polysemes) entries for a given word. Where a lexical resource is described in terms of granularity, a coarse-grained approach only differentiates between homonymy, whereas a fine-grained approach also considers polysemy. (Stokoe 2005)

Invariably, in every language, some words may have more than one meaning. Zipf (1949) even put forth a hypothesis that a word’s polysemy is

proportional to the square root of its frequency (Kuiper 2016). Dictionaries vary widely in their choices of meanings. Words have a wide variety of shades of meaning. It is not at all clear when a sense of the word should be treated as a separate meaning in a dictionary and when it should be subsumed within an already existing meaning. Kilgarriff (2006) mentions that there are no decisive ways of identifying where one sense of a word ends and the next begins. Therefore, this work is mostly left to the discretion of the lexicographer.

Modern dictionaries are compiled on the basis of evidence from language corpora (Kilgarriff 2006), based on Key Word In Context (KWIC). But it is not possible to apply this thought process in retrospect.

The goal of the lexicographer is to present a full account of the words of a language, in all their meanings and patterns of use (Kilgarriff 2006). In commercial life, the goal is always compromised by practical considerations such as the market that the dictionary is aimed at, its size, its editorial stance, and the speed with which it must be prepared.

Even in the case of dictionaries from the same lexicographer, the intended audience, printing, or economic considerations may force the lexicographer to deal with granularity in different ways across different dictionaries.

Thus, the word sense mapping between dictionaries will not always be a one-to-one mapping. Kouvara et al. (2021) suggested that there are four possibilities for two senses in different dictionaries. One-one, One-many,

many-one and many-many, which they term as perfect, wider-than, narrower-than and partial, respectively.

Some lexical resources have a special purpose, and do not cover all the headwords in the given language, e.g. English WordNet is designed to enumerate various senses of common English words. Its focus is on delineating word senses and not much on coverage of all the words.

Difference in the granularity of word senses across lexical resources can be seen from the following dictionary entries for the word “akūpāra” from Monier-Williams (MW) dictionary and Apte’s dictionary (AP). The entries are stripped of the gender and literary resource citations.

MW ID	MW Definition	AP ID	AP Definition
MW1	Unbounded.	AP1	Resulting in good, having a good issue.
MW2	The sea.	AP2	Unlimited, unbounded.
MW3	Tortoise, the mythical tortoise that upholds the world.	AP3	The sea, the receptacle of waters.
MW4	Name of a man.	AP4	The sun.

MW5	Name of an Āditya.	AP5	A tortoise in general.
		AP6	King of tortoises sustaining the world.
		AP7	A stone or rock.

It can be seen that AP1 and AP7 do not have any matching sense in MW. MW3 has been split into two senses, AP5 and AP6. MW4 does not find a matching entry in AP. MW5 thinks that this word means only the name of a specific Āditya, whereas AP4 thinks that this word means “The sun” in general. If we want to align these two lexical resources at the word sense level, we may need to take care of cases of one-to-one, one-to-many, many-to-one and many-to-many mappings.

2.5. Word Sense Disambiguation

Word Sense Disambiguation (WSD) is a field of Natural Language Processing (NLP) that is devoted to disambiguating the meaning of polysemous or homonymous words in a given specific context. Human beings use a lot of real-world knowledge to achieve this feat, without even realising that they are doing it. A human being will be able to easily identify the meaning of the word “bank” as a financial institution in the sentence “Bank refused to give him a loan” and as a geographical feature in the sentence “My

village is on the Northern bank of the river Ganga.” Whereas, the computational systems require explicit real-world information or other features like the occurrence of the word ‘loan’ with a financial institution and ‘river’ with the geographical feature.

Bevilacqua et al. (2021) have provided a summary of various recent trends in the field of WSD. Resources for WSD include (1) Sense Inventories and (2) Sense Annotated Data.

Sense inventories enumerate the set of possible senses for a given lexeme (<lemma, part of speech> pair). Most popular sense inventories are Princeton Wordnet (Miller et al., 1990), BabelNet (Navigli and Ponzetto, 2012) and Wiktionary.¹ WordNets have been extended to various Indian languages like IndoWordNet (Bhattacharya, 2010), Hindi WordNet (Narayan and Bhattacharyya, 2002), Sanskrit WordNet (Kulkarni et al., 2010), etc.

Sense Annotated Data include SemCor (Miller et al., 1993), English Princeton Wordnet Gloss Corpus (WNG)², and some improvements thereon.

The approaches to WSD have been classified as follows by Bevilacqua et al. (2021).

1 Knowledge-Based WSD

They use computational lexicons such as WordNet or BabelNet, etc.

2 Supervised WSD

They are neural network-based models. They are of the following types.

¹ <https://en.wiktionary.org/>

² <https://wordnetcode.princeton.edu/glosstag.shtml>

- 2.1 Purely Data-Driven WSD
- 2.2 Supervised models exploiting glosses
- 2.3 Supervised models exploiting relations in a knowledge graph
- 2.4 Supervised approaches using different sources of knowledge.

Currently, the Supervised WSD have outperformed Knowledge-based WSD, but knowledge bases remain useful because many supervised methods are effectively hybridised with knowledge-based methods (Bevilacqua et al., 2021). Bevilacqua et al. (2021) also provide an evaluation of the performance of various approaches to the WSD problem.

It is quite possible that the machine translation may not be able to correctly disambiguate the meaning, especially if the context is not clear or the given information is not associated with any specific words, or the words are used metaphorically. Even then, WSD continues to improve day by day. Knowledge-based methods such as SyntagRank (Scozzafava et al., 2020) give around 71.7 F-score, whereas the best performance is obtained by ESCHER (Barba et al., 2021), which gives an F-score of around 80.7.

Currently freely available translators like Google Translate can disambiguate common sentences and provide correct translations with reasonable accuracy from a layman's routine requirements. Sanskrit is a low-resource language as far as NLP resources are concerned. Pradeep et al. (2024) evaluated the performance of Google Translate on Bhagavadgītā from a WSD perspective. Out of erroneous outputs, errors due to polysemous words

were 139, errors due to knowledge-sensitive words were 40, errors due to compounding were 12 and 25 were incorrect outputs. A total of 700 verses were put under analysis in this experiment. Even in a complex language like Sanskrit, Google Translate has been able to translate 484 verses (69.14%) fully well. With the improvement in machine translation algorithms and the availability of training data, the accuracy will keep on improving further.

WSD is an important task for several NLP applications such as Machine Translation, Information Retrieval, Question Answering, Summarisation, and so on. At the same time, it is one of the most difficult problems in the field of Natural Language Processing (NLP). It is considered as being an AI-complete problem (Agirre and Edmonds, 2007). The difficulties arise due to a poor understanding of the process involved. Various factors, such as linguistic, contextual, domain-specific, cultural, and world knowledge, contribute to the process of manual word sense disambiguation. In the case of resource-rich languages such as English, there are several lexical resources with varied granularity, such as WordNet (Miller, 1995), FrameNet (Baker et al., 1998), ConceptNet (Speer et al., 2017), VerbNet (Schuler, 2005), etc., and sense-tagged corpora available in digital media. This resulted in several efforts aiming at the alignment of such resources known as Word Sense Aligned (WSA) resources. The development of Euro-WordNet (Vossen, 1998) and Indo-WordNet (Bhattacharyya, 2010) is also a step towards generating Word sense-aligned lexical resources so that the sense-tagged corpus in one language can become available in another with minimum effort.

In the present thesis, we are not involved in the problem of Word Sense Disambiguation per se. Therefore, we will not delve further into this topic.

2.6. Word Sense Alignment

Word sense alignment, also known as sense linking, sense alignment or sense mapping, matches two senses of the same word from two lexical resources. Manual compilation of dictionaries is nevertheless a costly and time-consuming activity, which has led to efforts towards developing methods for (semi-)automating the process (Saurí et al., 2019). There are two ways that this goal can be achieved.

1. Lexical translation - Automatically creating new bilingual dictionaries, a task that touches upon the area known as lexical translation, concerning systems able to return translations of words or phrases.
2. Enriching existing bilingual dictionary information with additional data available from other lexical resources (e.g., sense definitions, grammatical notes, domain information, etc.). This second task has to do with the area referred to as word sense linking (aka sense alignment, sense mapping or sense matching).

In our work, we are primarily interested in the second method. In recent years, Word Sense Alignment has gained importance. Languages with low resources would like to take advantage of the resources available in resource-rich languages by aligning their resources to those of the rich languages. For example, Salgado et al. (2020) describe the challenges of word sense alignment of Portuguese Language Resources. Joshi et al. (2012) present a heuristic approach to link English and Hindi WordNets by linking their senses. Two closely related Czech lexical resources VALLEX³ and PDT-VALLEX⁴ were aligned fully automatically (Bejcek et al., 2014). Johansson and Pina (2015) used word sense embeddings to automatically link the Swedish language banks.

During his post-doctoral fellowship in 2012 at INRIA, Pawan Goyal aligned the Sanskrit Heritage dictionary with an XML version of Monier-Williams available at CDSO (Goyal et al., 2012). Goyal used the online Google Translate to translate the French entries into English and then aligned them with the entries in Monnier Williams’ Sanskrit-English dictionary, by manually aligning the entries wherever there were ambiguities/multiple choices available. The alignment process is incremental and thus may be iterated on successive versions of the Sanskrit Heritage dictionary.

³ <http://ufal.mff.cuni.cz/vallex/2.6/>

⁴ <http://lindat.mff.cuni.cz/services/PDT-Vallex/>

2.7 Use case for Word Sense Alignment

With an increasing amount of word-sense disambiguated data being available, there is another aspect that would require closer scrutiny. Once a word's meaning has been disambiguated in the given context, only the said meaning from the myriads of meanings of that word can be shown to the user to help him understand the text. Once we have tagged those meaning numbers in a lexical resource, let us say WordNet, we have a text that has been tagged with word meanings from a lexical resource. If we can find a way to link various lexical resources at the word sense level, it would be a great help to the user. It will also help us to extend the user base to new languages. If we can link the Sanskrit-English dictionary with the Sanskrit-French dictionary at a word sense level rather than word level, we can take advantage of the WSD done with the help of the Sanskrit-English dictionary, and port it to the Sanskrit-French dictionary too. This process can be extended ad infinitum and hence improve the user experience. For example, I would like to see only the meaning 'financial institution' when I click on the word 'bank' in 'Bank refused to give him a loan'. I would not like to be cluttered by 18 meanings of the word bank. If we can find out a more or less language-agnostic way of automatically or semi-automatically aligning lexical resources at the word sense level, this advantage can be extended to many language pairs.

Saurí et al. (2019) also suggest an interesting use case. By taking the monolingual dictionary as the pivot to which several bilinguals are linked at

the sense level, we expect to be able to automate the creation of bilingual dictionaries involving language pairs not covered by the original bilinguals.

2.8. Challenges in Word Sense Alignment

The conceptual space is a continuum that is divided into discrete units by the lexicon of a language. Since the lexicon is denumerably finite, a word represents a piece of continuous conceptual space and not a discrete point.

This sometimes leads to one word representing a spectrum of meanings. Such words are termed polysemous words. Sometimes, more than one lexical unit produces the same word form. Such word forms are called homonyms.

Among the homonymous and polysemous meanings, typically, the homonyms are provided with different headword entries, while the polysemous meanings are clubbed under a single headword.

Within polysemous meanings, the granularity is decided by the lexicographer. Deciding the granularity of the meaning is not trivial. It is not at all clear when a sense of the word should be treated as a separate meaning and when it should be subsumed within an already existing meaning.

Further deciding between polysemy and homonymy is subjective due to the fuzzy boundary between them.

Another factor worth noting is the inclusion of metaphoric meanings in the dictionary. Indian tradition discusses three types of meanings, viz. *abhidhā* (literal), *lakṣaṇā* (metaphoric or secondary) and *vyañjanā* (suggestive). While it is impossible to provide the suggestive meanings, which are subjective in nature and also depend on the context, the lexicographers do consider the secondary or metaphoric meanings for inclusion in the dictionaries.

Even in the case of dictionaries from the same lexicographer, the intended audience, printing, or economic considerations may force the lexicographer to deal with sense granularity in different ways across different dictionaries. Therefore, the choice of sense granularity is mostly left to the discretion of the lexicographer, as has been observed through various lexical resources. Because of these reasons, the word sense mapping between different lexical resources is not trivial.

2.9. Problem Statement

“To find out a methodology by which lexical resources can be aligned at word sense level computationally with some scoring mechanism so that the allocation of scarce human resources can be done more efficiently”.

In the next section, we shall discuss the methodology adopted by us in this research.

Chapter 3

Methodology

3.1. Difficulty for resource-starved languages

Manually aligning lexical resources at the word sense level is a very laborious task. It would also require the person to be well-versed in two languages, e.g. aligning a Sanskrit-English and a Sanskrit-Hindi dictionary would require the person to know at least English and Hindi, and preferably Sanskrit too. For a resource-starved language like Sanskrit, this may be very costly and time-consuming.

The present work focuses on finding out the similarity between different meanings of a given word and presents the human annotator with a similarity score or a confidence score, so that the annotator may devote more time to the places where the machine performs with a low confidence level. We also aim at finding a more or less language-agnostic way of automatically or semi-automatically aligning lexical resources at the word sense level so that it can be extended to other language pairs.

For any mechanical mapping between entries from two dictionaries to be successful, either both target languages need to be the same or a model trained on both languages to identify similar concepts across both languages is needed. The first approach is simpler. Because of the advancement in machine translation technologies and publicly available resources such as Google Translate⁵, it has become possible to translate various texts from one language to another. Thus, in the absence of models trained in two different languages, one can still use Google Translate to identify similar concepts across languages.

3.2. Existing algorithms

The task of finding out the similarity between two documents (in our case, the meaning of the word for a given sense) is a common theme in information retrieval (IR), topic modelling, ontology matching, etc. There are various algorithms that have already been tested for the same.

Bär et al. (2012) have enumerated various similarity measures in their paper. A brief description of these similarity measures is as follows.

⁵ <https://translate.google.com/> accessed on 20 September 2023

3.2.1. Content Similarity Measures

These similarity measures depend on the content of the texts to identify the similarity between two sets of texts. They are primarily of three types: (1) Pure Text Similarity Measures, (2) Semantic Similarity Measures and (3) Statistical Similarity Measures.

3.2.1.1. Text Similarity Measures

Text similarity measures are computational techniques designed to quantify the degree of likeness between two text segments. They operate primarily on surface-level features such as exact character matches, token overlaps, and structural alignments, etc. These measures are particularly useful in identifying verbatim copying, minor paraphrasing, or typographical variations. They often serve as the foundational layer in plagiarism detection, duplicate content identification, and near-duplicate retrieval systems.

3.2.1.1.1. Longest Common Substring

The Longest Common Substring (LCS) method identifies the longest contiguous sequence of characters present in both texts. For example, in the sentences “The quick brown fox jumps” and “A quick brown fox ran,” the LCS is “quick brown fox.” This method is highly effective in detecting direct, unaltered copying, as it is sensitive to continuous sequences of matching text.

However, it performs poorly when substantial reordering or insertion of unrelated text breaks the contiguous match.

3.2.1.1.2. Longest Common Subsequence

The Longest Common Subsequence approach measures similarity based on the longest ordered sequence of tokens or characters that appears in both texts, allowing for gaps. Unlike the Longest Common Substring, it does not require the matches to be contiguous, making it more tolerant of insertions and deletions. For instance, in “The brown fox jumps” and “The fox quickly jumps,” the subsequence “The fox jumps” would be identified, even though an extra word appears in one version.

3.2.1.1.3. Greedy String Tiling

Greedy String Tiling (Wise, 1996) is an algorithm that matches identical substrings above a certain minimum length, ensuring that no part of the text is matched more than once. It works by repeatedly finding the largest possible match and “tiling” it, then continuing with the remaining unmatched segments. This approach is effective for detecting scattered matching phrases across texts, making it particularly suitable for plagiarism cases where copied material is fragmented or reordered. For example, given sequences $A = \text{“abcde123fghij”}$ and $B = \text{“xyzabc123klm”}$, with a minimum match length of 3, the Greedy String Tiling algorithm first identifies the longest common

substrings "abc" and "123", tiles them, and then terminates when no further matches above the threshold remain.

3.2.1.1.4. Jaro

The Jaro similarity metric (Jaro, 1989) measures the similarity between two strings by considering the number of matching characters and the number of transpositions required to align them. It is well-suited for short strings, such as names or titles, where minor typographical errors are common. Jaro metric is calculated as

$$\frac{1}{3} \left(\frac{m}{|S_1|} + \frac{m}{|S_2|} + \frac{m-t}{m} \right)$$

Where m is the number of matching characters. t is the number of transpositions. For example, in strings "KARTHIK" and "KATRHIK", the number of matching characters $m = 7$ and the number of transpositions $t = 1$, the length of $S1$ and $S2$ are 7, giving a similarity score of 0.952.

3.2.1.1.5. Jaro-Winkler

The Jaro-Winkler metric (Winkler, 1990) extends the Jaro similarity by giving additional weight to matches that occur at the beginning of the strings. This makes it especially useful in situations where common prefixes are meaningful, such as in proper names or technical terms. **Jaro-Winkler similarity** is defined as:

$$JW = J + (\ell \times p \times (1 - J))$$

where:

- J is the Jaro similarity,
- ℓ is the length of the common prefix (maximum of 4),
- p is a constant scaling factor (typically 0.1).

Example:

For the strings "KARTHIK" and "KATRHIK", the Jaro similarity $J=0.952$ and the common prefix length $\ell=2$. Using $p=0.1$, the Jaro–Winkler similarity is:

$$JW = 0.952 + (2 \times 0.1 \times (1 - 0.952)) = 0.962$$

3.2.1.1.6. Monge–Elkan

The Monge–Elkan method (Monge and Elkan, 1997) compares two multi-token strings by finding, for each token in the first string, the most similar token in the second string and averaging these best-match scores. It is particularly effective in matching records or phrases where word order may vary, as in “City of New York” and “New York City.” This flexibility allows it to handle partial reordering without a significant loss of similarity score.

3.2.1.1.7. Levenshtein

The Levenshtein distance (Levenshtein, 1966) calculates the minimum number of single-character insertions, deletions, or substitutions required to

transform one string into another. For example, transforming “kitten” into “sitting” requires three edits: substitute “k” with “s,” substitute “e” with “i,” and insert “g” at the end. The measure is widely used for detecting typographical errors, approximate matches, and simple paraphrasing.

3.2.1.1.8. Term Frequency-Inverse Document Frequency (TF-IDF)

The TF-IDF approach represents each document as a weighted vector of term frequencies, where rare terms are assigned higher weights than common terms. The similarity between two documents is typically computed using cosine similarity between these vectors. For instance, in a corpus where “tree” is rare but “money” is common, a match on “tree” contributes more to the similarity score. This method is effective for topic-based similarity detection.

3.2.1.1.9. Word n-grams

Word n-grams are contiguous sequences of n words extracted from a text. Similarity is computed based on the proportion of overlapping n-grams between two documents. For example, using trigrams ($n = 3$), the phrase “The quick brown fox jumps” produces “The quick brown”, “quick brown fox”, “brown fox jumps” as tri-grams. Overlap in n-grams can indicate shared phrasing, making this method sensitive to word order changes.

3.2.1.1.10. Jaccard Coefficient

The Jaccard Coefficient measures similarity between two sets as the size of their intersection divided by the size of their union. In text analysis, the sets often consist of tokens, n-grams, or shingles. For example, the sets {cat, sat} and {cat, sat, mat, hat} have a Jaccard coefficient of 0.5. This metric is simple but ignores token order and frequency.

3.2.1.1.11. Character n-gram Profiles

Character n-gram profiles represent a text as a sequence of n -character units, often including boundary markers. For example, for $n = 3$, “superb” yields “sup”, “upe”, “per” and “erb”. This method is robust to minor spelling variations and typographical errors, as small changes affect only a subset of n-grams.

3.2.1.2. Semantic Similarity Measures

Semantic similarity measures evaluate the likeness between texts based on meaning rather than surface form, often leveraging lexical databases or semantic networks. They are capable of detecting paraphrases, synonyms, and conceptually related terms that share little or no lexical overlap. These methods are particularly valuable in contexts where different wording is used to convey the same or related concepts.

3.2.1.2.1. Jiang and Conrath

The Jiang and Conrath measure (1997) combines information content with path length in a semantic taxonomy such as WordNet. This approach balances hierarchical relationships and corpus-based frequency statistics.

$$\text{dist}_{JCN}(c_1, c_2) = \text{IC}(c_1) + \text{IC}(c_2) - 2 \times \text{IC}(\text{LCS}(c_1, c_2))$$

$$\text{sim}_{JCN}(c_1, c_2) = \frac{1}{\text{dist}_{JCN}(c_1, c_2)}$$

where,

- $\text{IC}(x) = -\log(p(x))$ is the **information content** of concept x ,
- $p(x)$ is the probability of encountering x in a corpus,
- $\text{LCS}(c_1, c_2)$ is their **lowest common subsumer** in the taxonomy.

For explanation of the concept, we will take the following input data. “eagle” and “sparrow” have a common ancestor “bird”, and based on occurrence of these words in the corpus, the $\text{IC}(\text{sparrow}) = 9.12$, $\text{IC}(\text{eagle}) = 7.35$ and $\text{IC}(\text{bird}) = 5.23$.

In this case,

$$\begin{aligned} \text{dist}_{JCN}(\text{sparrow}, \text{eagle}) &= \text{IC}(\text{sparrow}) + \text{IC}(\text{eagle}) - 2 \times \text{IC}(\text{LCS}(\text{sparrow}, \text{eagle})) \\ &= \text{dist}_{JCN}(\text{sparrow}, \text{eagle}) = \text{IC}(\text{sparrow}) + \text{IC}(\text{eagle}) - 2 \times \text{IC}(\text{bird}) \\ &= 9.12 + 7.35 - 2 \times 5.23 \\ &= 6.01 \end{aligned}$$

The semantic distance between a sparrow and an eagle is **6.01**. If you were to run this same math comparing a *sparrow* and a *robin* (which has a lower IC closer to sparrow), the total distance would drop significantly, showcasing a tighter semantic cluster.

3.2.1.2.2. Lin

Lin's similarity measure (1998) is an information-theoretic approach that normalises the shared information between two concepts by the total information content of the concepts. It yields a value between 0 and 1, with higher values indicating greater similarity. This normalisation allows for more intuitive comparisons across different concept pairs.

The Lin similarity between two concepts c_1 and c_2 in a taxonomy is defined as:

$$\text{sim}_{\text{Lin}}(c_1, c_2) = \frac{2 \times IC(LCS(c_1, c_2))}{IC(c_1) + IC(c_2)}$$

where:

- $IC(x) = -\log(p(x))$ is the **information content** of concept x ,
- $p(x)$ is the probability of encountering x in a corpus,
- $LCS(c_1, c_2)$ is the **lowest common subsumer** of c_1 and c_2 .

For the same eagle, sparrow and bird example,

$$\text{sim}_{Lin}(\text{sparrow}, \text{eagle}) = \frac{2 \times IC(\text{bird})}{IC(\text{sparrow}) + IC(\text{eagle})}$$

$$= \{2 \times 5.23\} / \{9.12 + 7.35\} = 0.6351 = 63.51\%$$

The concepts share a similarity of **63.51%**. The score is high because a major portion of their total meaning is bound up in their shared "bird-ness."

3.2.1.2.3. Resnik

Resnik's measure (1995) defines the similarity of two concepts as the information content of their lowest common subsumer in a taxonomy.

$$\text{sim}_{Resnik}(c_1, c_2) = IC(LCS(c_1, c_2))$$

where:

- $IC(x) = -\log(p(x))$ is the **information content** of concept x ,
- $p(x)$ is the probability of encountering x in a corpus,
- $LCS(c_1, c_2)$ is the **lowest common subsumer** of c_1 and c_2 .

For our example,

$$\text{sim}_{Resnik}(\text{sparrow}, \text{eagle}) = IC(LCS(\text{sparrow}, \text{eagle}))$$

$$= IC(\text{bird}) = 5.23$$

3.2.1.3. Statistical Techniques

3.2.1.3.1. Latent Semantic Analysis (LSA)

Latent Semantic Analysis (Landauer et al., 1998) models texts in a reduced latent space derived from a term–document matrix via singular value decomposition (SVD), capturing higher-order co-occurrence structure that correlates with conceptual similarity beyond surface word overlap. Concretely, after weighting (e.g., tf-idf) and decomposing the matrix, documents and terms are compared in the truncated k -dimensional space, where cosine similarity between projected vectors reflects their proximity along latent topics. For example, a document about “interest rates” and another about “monetary policy” can be judged similar even with minimal shared vocabulary, because both load heavily on the same latent financial/economic dimensions. LSA is data-efficient for moderate corpora and robust to synonymy and polysemy, but its linearity and global factors can blur domain-specific nuances; careful choice of k and weighting is therefore essential.

3.2.1.3.2. Explicit Semantic Analysis (ESA).

Explicit Semantic Analysis (Gabrilovich and Markovitch, 2007) represents a text as a high-dimensional sparse vector over interpretable knowledge concepts (typically Wikipedia pages), where each dimension’s weight reflects the text’s association with that concept based on term–concept indexing. Similarity is then computed—usually via cosine—between two concept vectors, enabling comparisons that leverage world knowledge rather than raw lexical overlap. For instance, a snippet on “immune cells targeting

pathogens” and another on “T-lymphocytes combating infections” align strongly on concepts like “Adaptive immune system,” yielding high similarity despite paraphrasing. ESA’s interpretability and coverage of encyclopaedic knowledge make it effective for cross-domain comparisons, though performance depends on the quality and recency of the underlying knowledge base and the disambiguation of ambiguous terms.

3.2.2. Structural Similarity Measures

Structural similarity measures evaluate the degree of likeness between texts by comparing their structural or syntactic organisation rather than solely their lexical or semantic content. This approach is particularly effective when authors preserve the overall framework of a source while changing specific words or phrases, such as in disguised plagiarism.

3.2.2.1. Stopword n-grams.

Stopword n-gram similarity examines the sequence patterns of function words—such as “the,” “and,” “of”—that are often ignored in content-based similarity but can reveal shared structural and syntactic habits between texts. In this approach, contiguous sequences of n stopwords are extracted from both texts, and overlap is computed using measures such as Jaccard similarity. Because stopwords are less domain-specific and more stable across paraphrasing, this method is particularly useful in plagiarism detection or authorial style analysis. For example, two sentences— “She went to the

market to buy vegetables” and “They went to the store to purchase fruits”—share the stopword bigram “to the,” which contributes to their structural resemblance despite differing in lexical content.

3.2.2.2. Part-of-speech n-grams.

Part-of-speech (POS) n-gram similarity captures recurring grammatical patterns by replacing each token with its POS tag (e.g., NN for noun, VB for verb, JJ for adjective) and then extracting sequences of n tags. This abstracts away from the specific words, focusing instead on the syntactic skeleton of the text. The method can reveal stylistic or genre-specific patterns, such as frequent “adjective–noun” bigrams in descriptive writing or “noun–verb” bigrams in instructional texts. For instance, the sentences “The quick fox jumped” and “A clever boy ran” share the same POS structure “JJ NN VB”, which signals structural similarity despite no lexical overlap.

3.2.2.3. Pair order.

Pair order similarity assesses whether the relative ordering of particular word pairs is preserved across two texts, regardless of the words appearing consecutively or with intervening terms. By examining ordered pairs of content words, it detects structural alignment in the sequencing of ideas. For example, if both a source and a suspicious document mention “Paris” before “France” and “France” before “Europe,” the preserved ordering of these pairs increases their structural similarity score. This measure is less sensitive to

small insertions or deletions, making it useful for identifying paraphrased passages where the narrative sequence remains intact.

3.2.3. Stylistic Similarity Measures

3.2.3.1. Type–Token Ratio (TTR).

The Type–Token Ratio is a measure of lexical diversity, calculated as the number of distinct word types divided by the total number of word tokens in a text. A higher TTR indicates richer vocabulary use, whereas a lower value suggests repetition or limited lexical variation. For example, in “The cat chased the dog,” there are 5 distinct words out of 6 total tokens, giving a TTR of $5/6 \approx 0.83$. TTR is sensitive to text length, as shorter texts tend to yield artificially high values; nonetheless, it is useful for characterising stylistic tendencies, such as differentiating between academic prose (often high lexical diversity) and procedural instructions (often lower diversity).

3.2.3.2. Sequential TTR.

Sequential TTR mitigates the text-length bias of traditional TTR by computing the ratio incrementally over sequential segments of equal length (e.g., 50 words) and then averaging the results. This allows for more meaningful comparisons between texts of varying lengths. For instance, when comparing two novels, segment-wise TTR can reveal whether lexical diversity

is consistently maintained throughout the narrative or if it fluctuates between descriptive and dialogue-heavy sections. Sequential TTR is particularly useful for authorship attribution, where sustained patterns in lexical variation can serve as stylistic fingerprints.

3.2.3.3. Sentence length.

Sentence length, typically measured in words or characters per sentence, captures structural aspects of an author's writing style. Longer sentences often indicate complex syntactic constructions, while shorter sentences may suggest a direct, terse style. For example, academic journal articles tend to have higher average sentence lengths than online news briefs. Comparing sentence length distributions between texts can reveal stylistic imitation, simplification, or elaboration in paraphrased material. This measure is simple yet effective in distinguishing writing genres and authors.

3.2.3.4. Token length.

Token length refers to the average number of characters per word token, reflecting lexical complexity and preference for shorter or longer words. Texts in scientific and legal domains often contain longer average token lengths due to technical vocabulary, whereas casual or spoken language tends toward shorter words. For example, "photosynthesis" (14 characters)

contributes more to the average token length than “cat” (3 characters). Token length patterns can help identify domain-specific writing.

Similarity in function word profiles between two texts can suggest common authorship or deliberate stylistic imitation.

3.2.4. Discussion about existing algorithms

As the dictionary meanings are relatively small chunks of text, with sizes ranging up to two or three sentences at maximum, structural and stylistic similarity measures discussed above are not of much relevance to the task at hand.

Semantic similarity measures presume some graph-like structure and use the structure of those graphs to find out the similarity or nearness between two nodes. These work best when there is some ontological representation of the world knowledge or some hierarchy of the word/word senses and their relationship is explicitly coded. In an alphabetically arranged dictionary, such a relationship is almost nonexistent. We have suggested how a graph-based structure can be generated in such cases.

Latent Semantic Analysis (LSA) and explicit semantic analysis (ESA) require a lot of computational resources. Training and running these

algorithms on a large corpus like two full-fledged dictionaries will be computationally too heavy. LSA may be able to identify similarities between ‘child’ and ‘offspring’, which a normal text-based scorer may miss. As LSA requires very high computational powers, we ran the LSA on Gold Standard data alone and analysed the results.

Other than these measures, there are the following string-based measures implemented by the rapidfuzz library⁶, namely Damerau (1964), Hamming (1950), Indel, OSA, Prefix and Postfix.

3.3. Selection of dictionary pairs

Our focus is on the usability of string-based similarity measures for finding out the mapping of word senses. We present here our efforts towards the word sense alignment of Sanskrit lexical resources. We present three experiments. (a) The first one is with two different target languages, but the same compiler. Here we have chosen the Sanskrit-English⁷ and the Sanskrit-Hindi dictionaries of Apte⁸. Since the second dictionary is based on the first

⁶ <https://pypi.org/project/rapidfuzz/>

⁷ <https://www.sanskrit-lexicon.uni-koeln.de/scans/AP90Scan/2020/web/webtc/download.html>

⁸ Developed by the SHMT (Sanskrit-Hindi Machine Translation) consortium during 2008-2011, now a part of

one, the assumption is that there would be a good chance of getting a one-to-one mapping. (b) The second pair is with the same target language but different compilers. Here we have chosen the Sanskrit-English dictionaries by two different compilers – Wilson (1832)⁹ and Yates (1846)¹⁰. (c) The third one was a pair of monolingual English lexical resources, viz. Webster’s Unabridged Dictionary of the English Language (Webster, 1900)¹¹ and English Wordnet (Miller et al., 1990)¹².

3.4. Word Sense Alignment (WSA) of Sanskrit-English and Sanskrit-Hindi dictionaries of Apte

The Apte Sanskrit-English (AP90) dictionary used in this experiment is the Apte Sanskrit-English dictionary published in 1890. The later 1957 version of the dictionary (AP) was still under copyright. Therefore, CDSO did not have its data for open usage at the time of the study. AP90 is also not fully marked up to show different word senses separately. It had some rudimentary markup or regular expression patterns by means of which a crude parsing was done, and word meanings were separated.

Samsaadhanii Platform at <https://sanskrit.uohyd.ac.in/scl/>

⁹ <https://www.sanskrit-lexicon.uni-koeln.de/scans/WILScan/2020/web/webtc/download.html>

¹⁰ <https://www.sanskrit-lexicon.uni-koeln.de/scans/YATScan/2020/web/webtc/download.html>

¹¹ <https://www.gutenberg.org/files/29765/29765-0.txt>

¹² <https://github.com/fluhus/wordnet-to-json/releases/download/v1.0/wordnet.json.gz>

Apte Sanskrit-Hindi (ASH) dictionary (Apte, 2007) is a Hindi translation of Apte’s Sanskrit-English dictionary. It is not an exact translation. Many of the words have been omitted, and many meanings have been merged, deleted or separated. It seems that ASH had the advantage of using the data from the 1957 edition, too. Therefore, the new words or meanings added in that edition are also used in ASH. At the same time, ASH has been made more concise. Therefore, multiple meanings have been combined. Rare meanings have been dropped altogether, too. Therefore, it was not trivial to align the word senses in these two dictionaries, and hence, these were taken up to attempt word sense level alignment between them.

3.5. Gold Standard Data

As there is no previously existing gold standard data regarding word sense alignment of unstructured lexical resources like a dictionary pair in the Sanskrit language, a manual gold standard data set was created by selecting a random starting point and taking roughly 1000 ASH entries starting therefrom (See Table 1). Corresponding entries of AP90 were also taken up (See Table 2).

Head word	Hindi sense_id (ASH)	Hindi Meaning (ASH)
आकल्पः (Ākalpaḥ)	247	आभूषण, अलंकार
आकल्पः (Ākalpaḥ)	248	वेशभूषा
आकल्पः (Ākalpaḥ)	249	रोग, बीमारी

Table 1: Sample ASH entries

Head word	English sense_id (AP90)	English Meaning (AP90)
आकल्पः (Ākalpaḥ)	440	An ornament, decoration
आकल्पः (Ākalpaḥ)	441	Dress (in general), accoutrement
आकल्पः (Ākalpaḥ)	442	Sickness, disease
आकल्पः (Ākalpaḥ)	443	Adding to, increasing

Table 2: Sample AP90 entries

Every word sense had been given a unique identifier for each of the dictionaries separately. A manual examination of the data was done, and a manual mapping of the data was created. As and when a parsing error was detected in the data, the same was manually corrected. Sample entries of the sense alignment of entries from ASH and AP90 are shown in Table 3.

Head word	Hindi sense_id (ASH)	English sense_id (AP90)
आकल्पः (Ākalpaḥ)	247	440
आकल्पः (Ākalpaḥ)	248	441
आकल्पः (Ākalpaḥ)	249	442
आकल्पः (Ākalpaḥ)	-	443

Table 3: Gold Standard Data for Alignment of ASH and AP90

Since the two dictionaries selected had different target languages, for aligning the entries, we decided to use Google Translate to translate the meanings of AP90 into Hindi. In Table 4, column GSH shows the Google translation of the entries in AP90 into Hindi. The task at hand is to map the English sense_id to Hindi sense_id using GSH. Please note that sometimes Google Translate does not translate some difficult words, like 'accoutrement' and leaves them as they are when processed via bulk upload.

Head word	English sense_id (AP90)	English Meaning (AP90)	GSH
आकल्पः (Ākalpaḥ)	440	An ornament, decoration	एक आभूषण, सजावट
आकल्पः (Ākalpaḥ)	441	Dress (in general), accoutrement	पोशाक (सामान्य रूप से), accoutrement
आकल्पः (Ākalpaḥ)	442	Sickness, disease	रोग, रोग
आकल्पः (Ākalpaḥ)	443	Adding to, increasing	जोड़ना, बढ़ाना

Table 4: Sample AP90 entries along with their Hindi translations

Similarly, entries of ASH were translated into English with the help of Google Translate. Please see column GSE of Table 5.

Head word	Hindi sense_id (ASH)	Hindi Meaning (ASH)	GSE
आकल्पः (Ākalpaḥ)	247	आभूषण, अलंकार	jewellery, Ornament
आकल्पः (Ākalpaḥ)	248	वेशभूषा	Costumes
आकल्पः (Ākalpaḥ)	249	रोग, बीमारी	Disease, Disease

Table 5: Sample ASH entries with their English translations

In the next section, we present various algorithms and their performance on the gold standard data Algorithm.

3.6. Algorithms

Our algorithms are based on simple string-level similarity measures. We define four different units of comparison and four different units of measure for comparison, resulting in 16 different algorithms. We describe them below.

3.6.1. Unit of comparison

The three basic units we propose are words, shingles (n-grams at character levels), and syllables. While glancing at the ASH entries with GSH manually, we also realised that in the case of languages like Hindi, depending upon the presence of post-positions, the last character of the word is changed as in ‘baharā’ (बहरा) versus ‘bahare’ (बहरे). Hence, we decided to consider a word with the last character trimmed also as a unit of comparison.

3.6.2. Measure of comparison

We have identified four different measures for calculating the similarity. Suppose the meanings from two dictionaries are stored as a list of words $L1$ and $L2$. $l1$ and $l2$ are sets of unique words amongst $L1$ and $L2$ respectively. In the following notation, $|A|$ denotes the cardinality of set A. The four different measures are defined as

$$\begin{aligned} m_1 &= \frac{|L_1 \cap L_2|}{|L_1 \cup L_2|} \\ m_2 &= \frac{|l_1 \cap l_2|}{|l_1 \cup l_2|} \\ m_3 &= \frac{|L_1 \cap L_2|}{|L_1|} \\ m_4 &= \frac{|l_1 \cap l_2|}{|l_1|} \end{aligned}$$

The phrases describing the senses are tokenised, and stop-words are removed. In the case of English, all the words are converted to lowercase. Let us assume the two senses that need to be aligned are ‘space, place in general’ and ‘free space or vacuum’. As a first step, the phrases are tokenised and the stop words are removed, and the words are changed to lowercase. This results in two word lists.

$L1 = ["space", "place", "general"]$, and

$L2 = ["free", "space", "vacuum"]$

As there are no duplicate words in any of these two word lists, $I1 = L1$ and $I2 = L2$.

Thus, for the above lists,

$$m1 = 1/6; m2 = 1/5; m3 = 1/3; m4 = 1/3$$

The shingles for each word are all possible n-grams of characters.

e.g.

$\text{shingles}(\text{"space"}) = [\text{"s"}, \text{"sp"}, \text{"spa"}, \text{"spac"}, \text{"space"}, \text{"p"}, \text{"pa"}, \text{"pac"}, \text{"pace"}, \text{"a"}, \text{"ac"}, \text{"ace"}, \text{"c"}, \text{"ce"}, \text{"e"}]$

The trimmed words are obtained by trimming the last character of the word. So, the trimmed word list for L1 is $[\text{"spac"}, \text{"plac"}, \text{"genera"}]$. While we do not see this trimmed word list as of any advantage in the case of English, - for languages like Hindi, these are useful. For example, in the word mapping, the words 'baharā' (बहारा) and 'bahare' (बहारे) will not match, but after trimming the last phoneme, both words will match 'bahar'.

Words were also split at the vowel level. It was tried as an attempt to reduce the overgeneration of n-grams. e.g. $\text{split_at_vowel}(\text{'space'}) = [\text{'spa'}, \text{'ce'}]$

3.6.3. Measures of similarity

With the 4 units of comparison and 4 units of measures of similarity, there are 16 different measures for judging the similarity between the two senses. These 16 measures are shown in Table 6.

Unit	m1	m2	m3	m4
word	CR1	CR2	CR3	CR4
shingles	CR5	CR6	CR7	CR8
trimmed word	CR9	CR10	CR11	CR12
syllable	CR13	CR14	CR15	CR16

Table 6: Metrics used for evaluation.

3.6.4. Thresholds

A threshold similarity score of 0.2 was empirically selected to ignore the mappings with a low similarity score. Another measure, delta, was also calculated. It is the difference between the word sense pair across dictionaries with the highest similarity score and the second-best pair. If delta is high, it means that the pair at the first rank is ahead of the pair at the second rank comfortably. A threshold value of delta was kept at 0.1 empirically.

3.7. Evaluation on Gold Standard Data

3.7.1. Identification percentage

After setting these thresholds, the comparison of the results of all algorithms was made. The gold standard consists of 2022 word sense pairs manually validated. The results of some of the standard algorithms implemented by the rapidfuzz library are shown in Table 7.

Algorithm	Pairs identified	Percentage
Levenshtein	1793	88.67%
Damerau	1794	88.73%
Hamming	1688	83.48%
Indel	1822	90.11%
Jaro	1816	89.81%
Jaro-Winkler	1818	89.91%
OSA	1793	88.67%
Prefix	1749	86.50%
Postfix	1690	83.58%

Table 7: Percentage of word sense pairs correctly identified from gold standard data by already existing algorithms

With the same thresholds, the number of pairs identified out of total 2022 pairs of gold standard data by algorithms CR1 to CR16 are shown in Table 8.

Algorithm	Pairs identified	Percentage
CR1	1866	92.28%
CR2	1871	92.53%
CR3	1855	91.74%
CR4	1856	91.79%
CR5	1856	91.79%
CR6	1887	93.32%
CR7	1858	91.89%
CR8	1863	92.14%
CR9	1861	92.04%
CR10	1865	92.24%
CR11	1848	91.39%
CR12	1847	91.35%
CR13	1847	91.35%
CR14	1878	92.87%
CR15	1841	91.04%
CR16	1851	91.54%

Table 8: Percentage of word sense pairs correctly identified from gold standard data by various algorithms

3.7.2. Precision, Recall and F1-score

For each algorithm, the predicted set of (A, B) pairs is compared against the gold-standard set. Each pair is treated as a binary classification instance:

Symbol	Definition
TP	Pair appears in both the prediction and the gold standard
FP	Pair appears in the prediction but not in the gold standard
FN	Pair appears in the gold standard but not in the prediction

The following standard information-retrieval metrics are computed (Rijsbergen, 1979):

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

The results of various algorithms are as shown below:

Algorithm	Precision	Recall	F1-score
cr1	0.8814	0.9228	0.9017
cr2	0.8842	0.9253	0.9043
cr3	0.8762	0.9174	0.8964
cr4	0.8767	0.9179	0.8968
cr5	0.8834	0.9179	0.9003
cr6	0.9003	0.9332	0.9165
cr7	0.8814	0.9189	0.8998
cr8	0.8821	0.9214	0.9013
cr9	0.8782	0.9204	0.8988
cr10	0.8801	0.9224	0.9007
cr11	0.8729	0.9139	0.8930
cr12	0.8716	0.9135	0.8921
cr13	0.8791	0.9135	0.8959
cr14	0.8986	0.9288	0.9134
cr15	0.8713	0.9105	0.8904
cr16	0.8764	0.9154	0.8955
damerau	0.8490	0.8872	0.8677
hamming	0.7866	0.8348	0.8100
indel	0.8668	0.9011	0.8836
jaro	0.8611	0.8981	0.8792
jarowinkler	0.8620	0.8991	0.8802

Algorithm	Precision	Recall	F1-score
levenshtein	0.8482	0.8867	0.8670
osa	0.8482	0.8867	0.8670
postfix	0.7703	0.8358	0.8017
prefix	0.7972	0.8650	0.8297

Table 9: Precision, Recall and F1-score of various algorithms on Gold Standard data

The algorithm **cr6** achieves the highest F1-score (0.9165), with **cr14** as the nearest competitor (F1 = 0.9134).

3.7.3. Cochran's Q Test

Before conducting pairwise comparisons, an omnibus test is applied to determine whether any statistically significant differences exist across the full set of algorithms. Cochran's Q test (Cochran, 1950) is appropriate here because:

1. The outcome variable is binary (correct / incorrect prediction per gold pair).
2. All algorithms are evaluated on the same set of instances (matched/dependent samples).
3. More than two classifiers are compared simultaneously.

A binary concordance matrix $\mathbf{X}$ of dimension $N \times k$ is constructed, where N = number of gold-standard pairs and k = number of algorithms. Cell $X[i][j] = 1$ if algorithm j correctly predicted gold pair i (i.e., the pair appears in the algorithm's output), and 0 otherwise.

The Q statistic is defined as:

$$Q = \frac{(k-1) \left[k \sum_{j=1}^k L_j^2 - \left(\sum_{j=1}^k L_j \right)^2 \right]}{k \sum_{j=1}^k L_j - \sum_{i=1}^N R_i^2}$$

where $L_j = \sum_{i=1}^N X_{ij}$ is the column sum (number of pairs correctly predicted by

algorithm j), and $R_i = \sum_{j=1}^k X_{ij}$ is the row sum (number of algorithms that correctly predicted pair i).

Under the null hypothesis H_0 (all algorithms equally likely to predict any given pair correctly), Q follows a chi-squared distribution with $k - 1$ degrees of freedom.

Parameter	Value
Gold-standard pairs (N)	2,022
Algorithms (k)	25
Degrees of freedom	24
Q statistic	789.80
p -value	2.95×10^{-151}

Table 10: Cochran's Q Test results on Gold Standard data

Thus, the null hypothesis (all 25 algorithms are equivalent) is rejected with overwhelming evidence ($p < 0.001$). At least one algorithm differs significantly from the others. This result justifies proceeding to pairwise post-hoc testing.

3.7.4. Post-hoc Pairwise Test: McNemar's Test

Having established that the algorithms are not uniformly equivalent, pairwise comparisons involving the best-performing algorithm are conducted using McNemar's test (McNemar, 1947). McNemar's test is the standard non-parametric method for comparing two classifiers evaluated on the same test set (Dietterich, 1998).

For each pair (best algorithm, comparator algorithm), all instances drawn from the union of gold-standard pairs and both algorithms' predictions are examined. Two counts are computed:

- **n_{10}** : instances where the best algorithm was correct and the comparator was wrong
- **n_{01}** : instances where the comparator was correct and the best algorithm was wrong

The test statistic with continuity correction is:

$$\chi^2 = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}} \sim \chi^2(1)$$

To control the family-wise error rate across 24 simultaneous comparisons, the Bonferroni correction (Bonferroni, 1936) is applied, yielding an adjusted significance threshold of $\alpha/24 = 0.05/24 \approx 0.0021$.

Results of McNemar's Test comparing cr6 versus all other algorithm is shown below. Bonferroni-corrected threshold: $p < 0.0021$. Column n_{10} favours cr6; n_{01} favours the comparator.

Algorithm	n₁₀	n₀₁	χ^2	<i>p</i>-value	Significant?
cr1	220	157	10.20	0.001407	Yes ***
cr2	192	140	7.83	0.005126	Yes *
cr3	253	168	16.76	4.24×10^{-5}	Yes ***
cr4	250	167	16.12	5.93×10^{-5}	Yes ***
cr5	120	53	25.18	5.22×10^{-7}	Yes ***
cr7	216	146	13.15	0.000287	Yes ***
cr8	207	143	11.34	0.000759	Yes ***
cr9	231	156	14.15	0.000169	Yes ***
cr10	203	136	12.85	0.000338	Yes ***
cr11	262	163	22.60	2.00×10^{-6}	Yes ***
cr12	265	162	24.37	7.97×10^{-7}	Yes ***
cr13	183	98	25.11	5.41×10^{-7}	Yes ***
cr14	115	103	0.56	0.456263	No
cr15	270	161	27.06	1.97×10^{-7}	Yes ***
cr16	248	160	18.55	1.65×10^{-5}	Yes ***
damerau	323	120	92.11	8.21×10^{-22}	Yes ***
hamming	591	143	272.22	3.73×10^{-61}	Yes ***
indel	238	102	53.60	2.45×10^{-13}	Yes ***

Algorithm	n_{10}	n_{01}	χ^2	p -value	Significant?
jaro	290	135	55.80	8.01×10^{-14}	Yes ***
jarowinkler	287	136	53.19	3.03×10^{-13}	Yes ***
levenshtein	327	121	93.81	3.48×10^{-22}	Yes ***
osa	326	120	94.23	2.81×10^{-22}	Yes ***
postfix	644	152	302.87	7.82×10^{-68}	Yes ***
prefix	532	158	201.64	9.18×10^{-46}	Yes ***

Table 11: McNemar Test Results of CR6 algorithm versus all other algorithms

* $p < 0.05$ (not significant after Bonferroni correction)

*** $p < 0.0021$ (significant after Bonferroni correction)

The evaluation yields the following principal conclusions:

1. **cr6 is the best-performing algorithm** by F1-score (0.9165), achieving the highest precision (0.9003) and recall (0.9332) of all 25 candidates.
2. **cr6 is statistically indistinguishable from cr14** (McNemar's $\chi^2 = 0.56$, $p = 0.456$). The discordance counts ($n_{10} = 115$, $n_{01} = 103$) are nearly equal, indicating that neither algorithm has a consistent advantage over the other. Either may be recommended as the preferred alignment strategy.

3. **cr6 is significantly superior to all remaining 23 algorithms** after Bonferroni correction, with the weakest performing group being the character-overlap-based methods (postfix, prefix, hamming).
4. **cr2 occupies an intermediate position**: it reaches nominal significance ($p = 0.005$) but does not survive the Bonferroni correction, suggesting its inferiority to cr6.

Thus, CR6 is the best performing algorithm, followed by CR14. As CR6 gave better results on Gold Standard data than CR14 in terms of all parameters measured i.e. absolute numbers of pairs identified, precision, recall and F1-score; the algorithm CR6 was selected out of these algorithms. CR6 makes use of shingles and hence captures various features like terminal case removal, textual similarity between tatsama words and tadbhava words, common verb or common noun in compounds, etc. This may be the reason why CR6 gives better results than other algorithms.

3.8. Merits of the usage of two languages as target languages

Algorithm CR6 was applied to the complete dictionaries ASH (D1) and AP90 (D2). As AP90 definitions are in the English language and ASH definitions are in the Hindi language, both were translated with the help of

Google Translate, and an English version of ASH (E) and a Hindi version of AP90 (H) were created. D1 and H were compared against each other (both with Hindi definitions), and D2 and E were compared against each other (both with English definitions). Having comparisons with both languages helped in a big way. There are cases where one language is insufficient to map satisfactorily, but the other language could map without any difficulty. Let us see such a case with an example.

D1.91	जो चुराये जाने के योग्य न हो, या हटाये जाने अथवा दूर ले जाये जाने के योग्य न हो
D1.92	श्रद्धालु, निष्ठावान्
D1.93	दृढ, अविचल, अननुनेय
D1.94	पहाड़

Table 12: Entry of the word 'अहार्य' in the ASH (D1) dictionary

D2.1916	not to be stolen, removed or carried
D2.1917	not to win (by fraud), devoted, loyal
D2. 1918	firm, steadfast, hard
D2.1919	a mountain

Table 13: Entry of the word 'अहार्य' in the AP90(D2) dictionary

As can be seen from the contents, the four senses of D1 correspond sequentially to the four entries of D2.

Now, let us look at the Google translations of D1 into English and D2 into Hindi.

E.91	Unstealable, or not capable of being removed or taken away
E.92	Devotees, loyal
E.93	Strong, motionless, irresistible
E.94	Mountain

Table 14: Entry of ASH translated to English via Google Translate (E)

H.1916	चोरी, हटाया या ले जाने के लिए नहीं
H.1917	जीतने के लिए नहीं (धोखाधड़ी से), समर्पित, वफादार
H.1918	दृढ़, अडिग, कठोर
H.1919	एक पर्वत

Table 15: Entry of AP90 translated to Hindi via Google Translate (H)

3.8.1. Cases where one language fares better than another

Had we used only the translation of AP90 into Hindi through Google Translate, and compared it with the entries in ASH, the words 'पहाड़' (D1.94) and 'एक पर्वत' (H.1919) would not get a good similarity score. However, the same words when translated to English will be highly similar, viz., 'a mountain' (D2.1919) and 'Mountain' (E.94). Therefore, the similarity score

with English as the destination language will be very high. Similarly, 'श्रद्धालु, निष्ठावान्' will not match 'समर्पित, वफादार' much at the character level, but 'devoted, loyal' will match 'Devotees, loyal' at the character level.

There are also cases where Hindi fares better. Mapping 'फँसा हुआ' with 'फँसा हुआ' is easier than mapping 'trapped' and 'entangled' (D1.342 and D2.556).

Thus, using two languages helps us to take care of some cases where one language uses different synonyms and the other language has only one word for the concept or may have used the same word out of available synonyms.

3.8.2. Confidence levels with two target languages

Creating a mapping with two languages also gives us some more benefits. It gives us more confidence about a given mapping if both languages give the same mapping. Based on these insights, an analysis of the mappings of gold standard data and full dictionary data was carried out. We classify the confidence levels of the machine into 7 different categories. These categories are shown in Table 16 with their corresponding confidence levels.

Category	Description	Confidence
A	Both languages give above sim_threshold, and both languages give the same first match.	High
B	(Language1 above sim_threshold, and Language2 gives lower similarity score) or (Language2 above sim_threshold, and Language1 gives lower similarity score)	High
C	Headword present in only one dictionary, and absent in the other	High
D	Both languages give below sim_threshold, and both languages give the same first match.	Low
E	(Language1 below sim_threshold, but better than Language2) or (Language2 below sim_threshold, but better than Language1)	Low
F	Headword present in both dictionaries, but all entries of dictionary1 have already been assigned to other entries of dictionary2 or vice versa. Hence, there is no mapping	High
G	Force mapped, as this is the only remaining match	Low

Table 16: Categorisation of various mappings along with their confidence level

3.9. Analysis of gold standard data

Analysis of gold standard data with these codes yielded the following results (See Table 17).

Category	Pairs in the category	Percentage
A	636	30.34%
B	221	10.54%
C	738	35.21%
D	60	02.86%
E	59	02.81%
F	360	17.18%
G	22	01.05%
Total	2096	100%

Table 17: Categorisation of gold standard mapping created via algorithm CR6

It is worth noting that the machine generated a total of 2096 mappings. The gold standard data has 2022 mappings. It is because the machine does not know what are the number of mappings present in the gold standard data. Word senses may have one-one, one-many and many-one mappings. Therefore, it is not possible to determine in advance how many word sense mappings are to be generated. Therefore, the machine generated a total of 2096 mappings. Among these, the entries falling in the category of D ,E and G have low confidence. Thus, roughly 6% cases are such which are of low quality and

would improve with human intervention. The rest of the 94% cases can be mechanically aligned, saving precious resources. As the gold standard data was corrected as and when some parsing error or typographic error was seen, the error rate in the gold standard data is much less. Whereas, there was no attempt made to clear these kinds of errors in full dictionaries. Therefore, the error rates in the full dictionaries are more than the gold standard data.

3.10. Analysis of ASH and AP90 full dictionary mapping

The following are the results of application of the above methodology to full dictionaries Apte Sanskrit–Hindi (ASH) and Apte Sanskrit–English (AP90) (See Table 18).

Categories	Pairs in the given category	Percentage
A	51964	32.59%
B	16655	10.44%
C	47131	29.56%
D	9730	06.10%
E	6258	03.92%
F	25710	16.12%
G	2023	01.27%
Total	159471	100.00%

Table 18: Categorisation of word sense mappings generated by algorithm CR6 for Apte Sanskrit-Hindi and Apte Sanskrit-English dictionaries

Roughly 11% of cases fall under the low confidence zone, which may require human intervention. Three random numbers were selected, and 100 entries starting therefrom were examined for false positives. The following is the result obtained from the analysis of these 300 meaning pairs. (See Table 19)

Category	False positives
A	01
B	14
C	00
D	01
E	10
F	00
G	04
Total	30

Table 19: False positives from randomly selected mappings.

Thus, manual examination also yields around a 10% error rate (30/300 mappings generated).

3.11. Mapping other dictionaries

Similar exercise was also tried for different dictionary pairs, like (1) Apte Sanskrit-English and Monnier Williams Sanskrit-English dictionary¹³, (2) Wilson Sanskrit-English and Yates Sanskrit-English dictionary, and (3) English WordNet and Webster's English dictionary.

¹³<https://www.sanskrit-lexicon.uni-koeln.de/scans/MWScan/2020/web/webtc/download.html>

3.11.1. Apte Sanskrit-English and Monnier Williams Sanskrit English dictionaries

Apte Sanskrit-English dictionary and Monier Williams dictionary are two of the most commonly used dictionaries in the Sanskrit language. Both have different levels of granularity. Apte has substantially fewer word senses (133375) than Monier Williams (287606). Even then, many headwords from Apte are absent in MW.

A machine mapping of the word sense level between these two dictionaries was taken up. The following are the results of the same.

Category	Pairs in the given category	Percentage
A	60687	17.84%
B	15398	04.53%
C	180726	53.12%
D	28551	08.39%
E	9465	02.78%
F	43194	12.70%
G	2184	00.64%
Total	340205	100.00%

Table 20: Categorisation of mapping between Apte Sanskrit-English and Monnier Williams dictionaries

3.11.2. Comparison between Wilson and Yates dictionaries

A similar comparison was also drawn between Wilson (1832) and Yates (1846) dictionaries. The result is shown below.

Category	Pairs in the given category	Percentage
A	47947	53.22%
B	10272	11.40%
C	12109	13.44%
D	7195	07.99%
E	4321	04.80%
F	7232	08.03%
G	1016	01.13%
Total	90994	100.00%

Table 21: Categorisation of mapping between Wilson and Yates dictionaries

3.11.3. Comparison between Wordnet and Webster dictionaries

After comparing Sanskrit-Hindi and Sanskrit-English dictionaries, the same approach was tried with English-English dictionaries. Wordnet is a

lexical resource in the English language designed to have an ontological structure and with a thrust on identifying various word senses associated with commonly used words. Its coverage of headwords is not as comprehensive as a regular lexicon would have. Webster's is a classic dictionary. Wordnet has 207272 word senses, whereas Webster has 166980 word senses. The result of the comparison between them is shown below.

Category	Pairs in the given category	Percentage
A	29967	09.43%
B	12397	03.90%
C	196512	61.83%
D	21993	06.92%
E	15917	05.01%
F	37118	11.68%
G	3922	01.23%
Total	317826	100.00%

Table 22: Categorisation of mapping between Wordnet and Webster dictionaries

3.11.4. Comparative analysis of different dictionary pairs

The comparative categorisation of the various dictionary pairs is shown in the Table 23 for easy comparison.

Category	Apte-MW	Wilson-Yates	Wordnet-Webster
A	17.84%	53.22%	09.43%
B	04.53%	11.40%	03.90%
C	53.12%	13.44%	61.83%
D	08.39%	07.99%	06.92%
E	02.78%	04.80%	05.01%
F	12.70%	08.03%	11.68%
G	00.64%	01.13%	01.23%

Table 23: Categorisation of mappings for various dictionary pairs

Thus, in almost all dictionary pairs studied, the error rate (D+E+G) is roughly to the tune of 11-14%. These are the places where human annotators can make maximum impact by manual examination and correction.

3.12. Experiments with graph-based methods for finding similarity

Graph-based methods require some sort of existing graphs to work. These methods are usually employed in cases where there are some relationships between the nodes. These methods have been tried with WordNet due to its predefined ontological relationships, and with Wikipedia because of the presence of links between two articles. These relationships form the graph edges.

In graph theory, a graph G is defined as a set of vertices V and edges E belonging to the set $V \times V$. The edges can be weighted or unweighted. The edges can also be directed or undirected. For finding out similarity between two vertices, usually the 'shortest path' method is used. It identifies the shortest path required for travelling from the source vertex to the target vertex.

In our case, there is no such preexisting graph structure. The graph structure has to be generated first before any of the tools related to graph theory can be put to use. As we are using only textual similarity scores, which are based on the character / ngram / word level, it sometimes becomes

difficult to find a similarity between two entries because they may have used separate synonyms or words altogether. Therefore, there is not much textual similarity between them. In such cases, a graph-based approach may fare better.

Let us understand the matter with a real example. The following are two of the meanings of the same word 'कमल' in AP90 and MW dictionaries.

AP90

1. The Sārasa bird.
2. A lotus

MW

1. The Indian crane (*Ardea Sibirica*)
2. a lotus, lotus-flower, *Nelumbium*

As can be seen from the definitions, textual similarity-based algorithms would be able to identify the similarity between the 2nd sense, i.e. related to the lotus, because both the definitions use the same word 'lotus'.

A human being will be able to identify that "The Sārasa bird" and "the Indian crane (*Ardea Sibirica*)" are the same entities, but there is actually no same or similar word which is being used in these definitions. Therefore, textual similarity-based algorithms will not be able to identify correspondence between these two entries. Similarity score will be abysmally low.

Let us see if the graph-based method may help us in such cases. We have used the data of AP90 as training data for the graph and then queried the entries of MW. We searched for the word 'Sārāsa' in the AP90 text. There are many Sanskrit words that have this word in their meaning.

AP90

कमल - 1. The Sārāsa bird. 2. A lotus

अम्भोज - 1. the (Indian) crane or Sārāsa. 2. a lotus

The graph nodes are abbreviated in the following manner.

SW stands for "Sanskrit Word".

SN stands for "Sense".

EW stands for "English Word".

The numbers suffixed are identifiers to identify the nodes.

By following the convention mentioned above, the nodes of the graph are as follows.

Nodes

SW_1 – 'kamala'

SN_1:1 - <sense referring to the bird for the word kamala>.

SN_1:2 - <sense referring to the aquatic flower>.

EW_1 – sārāsa

EW_2 – bird

EW_3 – lotus

SW_2 – ‘amBoja’

SN_2:1 - <sense referring to the bird for the word amBoja.>

SN_2:2 - <sense referring to the aquatic flower.>

EW_4 – indian

EW_5 – crane

Note that the words sārāsa and lotus have already been assigned EW ids and therefore, the same Ids are used. No new IDs are allotted to them. Thus, an ID is assigned to unique Sanskrit Words or English Words. Based on these nodes and the relationships between them, A graph for the first entry ‘kamala’ will be as follows.

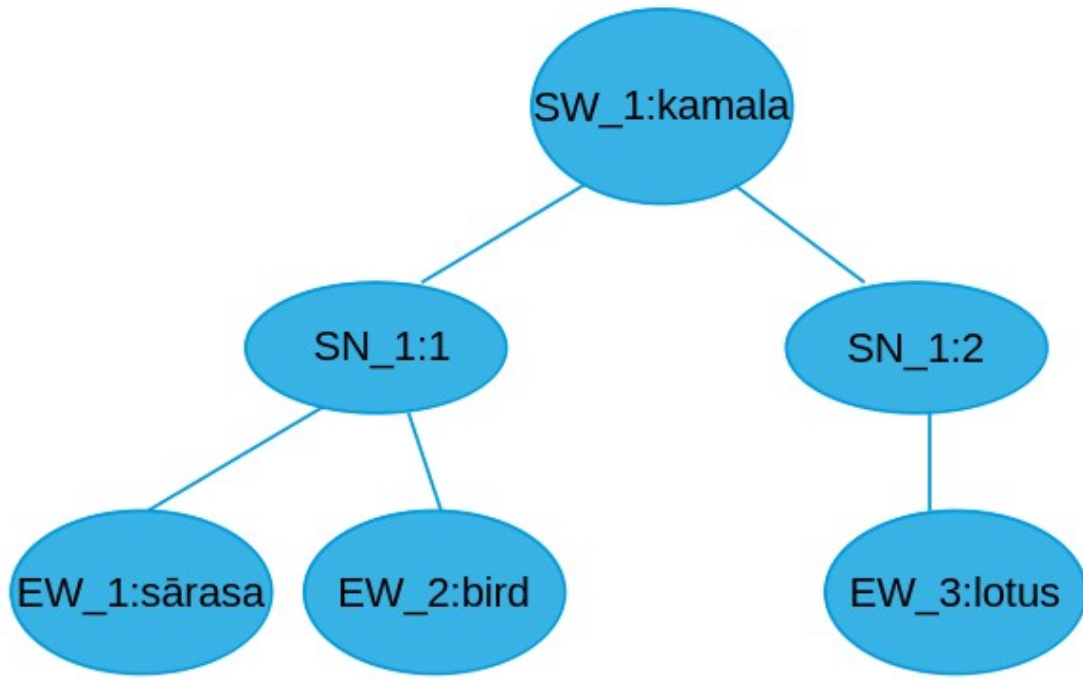


Figure 1: Graph for the word “kamala”

Similarly, the graph for the second word ‘amBoja’ will be like the following.

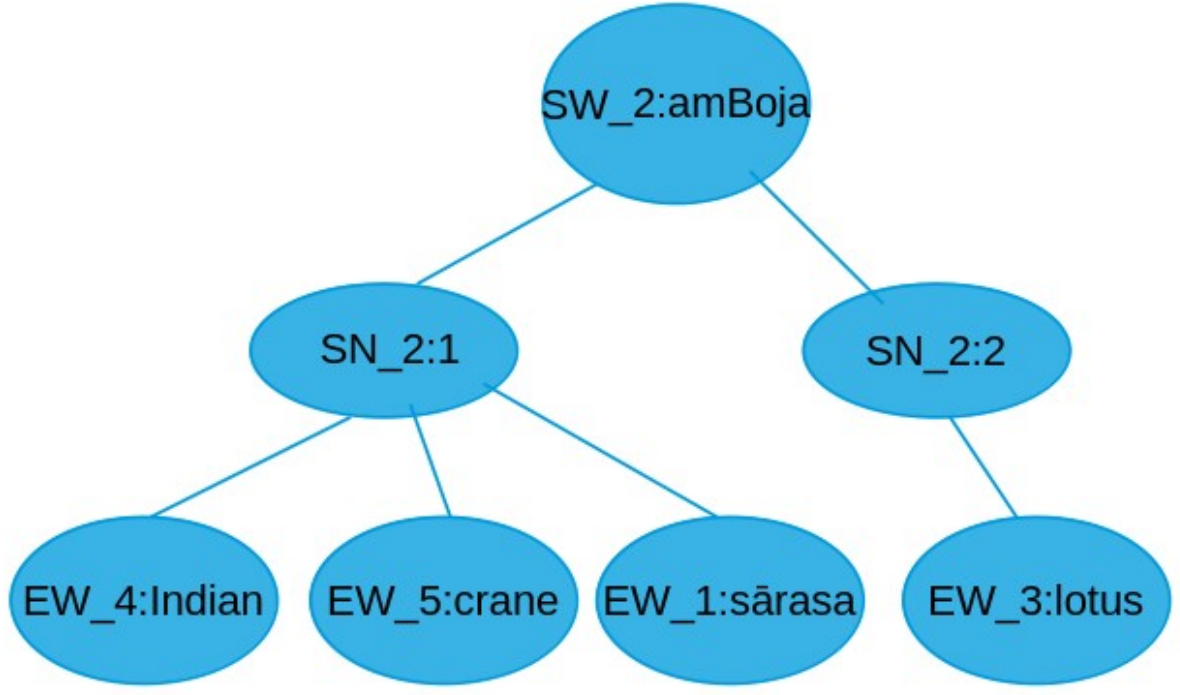


Figure 2: graph for the word “amBoja”

For combining these two graphs, EW nodes were treated as the bridges. Note that the linking between these two graphs happens only at the EW nodes, i.e. with the English words in their definition, and not between SW nodes (Sanskrit words) or SN nodes (Sense nodes). When these two graphs are combined, the resultant graph looks like the following.

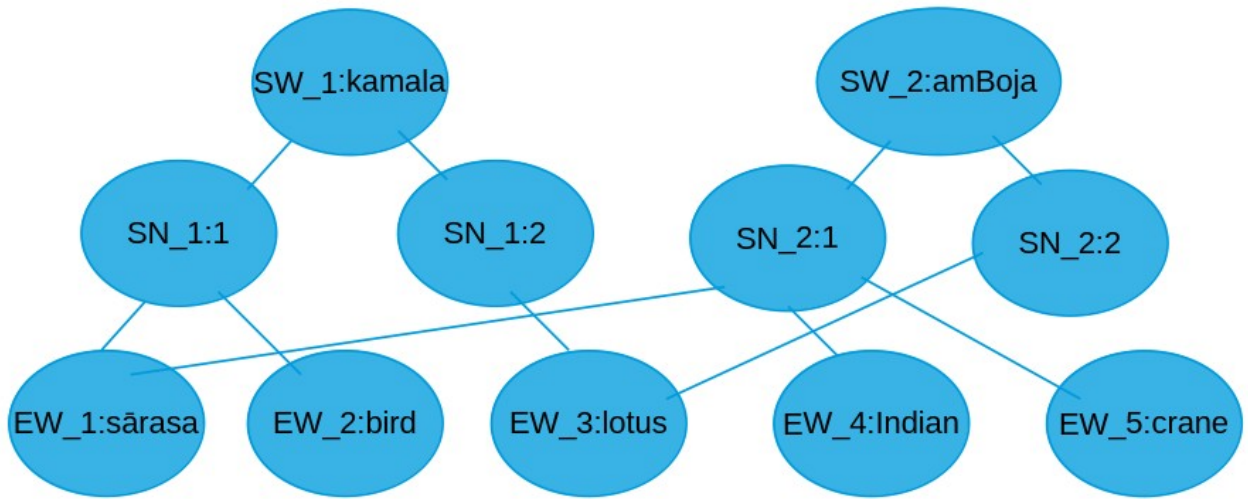


Figure 3: Combined graph generated by amalgamating the graphs for “kamala” and “amBoja” with English words as a bridge between graphs.

See that the nodes “EW_1: sārasa” and “EW_3: lotus” appear in both definitions and help to create a bridge that joins these two graphs.

Now, the following query string from MW is processed, namely ‘the Indian crane (Ardea Sibirica)’. This gets to the following list of words after preprocessing, like removal of stopwords and conversion to lower case, etc. - [‘indian’, ‘crane’, ‘ardea’, ‘sibirica’]

Now, each of these English words is searched in the graph of AP90. e.g. word ‘indian’ is EW_4 and ‘crane’ is EW_5. The rest of the words do not exist in our toy graph. The problem of identifying the most similar sense of the two given senses of ‘kamala’ can be translated into a graph theory problem as

‘finding out through which SN node (Sense node), the shortest path to reach from the given EW (English word) to the given SW (Sanskrit word) passes’.

The shortest paths are as shown below.

1. ‘indian’ – [‘EW_4: indian’, ‘SN_2:1’, ‘EW_1: sārāsa’, ‘SN_1:1’, ‘SW_1: kamala’]
2. ‘crane’ - [‘EW_5: crane’, ‘SN_2:1’, ‘EW_1: sārāsa’, ‘SN_1:1’, ‘SW_1: kamala’]

Thus, the shortest paths from both the words ‘indian’ and ‘crane’ reach the ‘SW_1: kamala’ node from the ‘SN_1:1’ node. Therefore, sense SN_1:1 is more similar to the query string than SN_1:2.

Thus, seemingly dissimilar words that occur in the definitions of the given dictionaries can be linked together to form a graph, and it will give some meaningful relationships to help us identify the similarity.

In our toy example, both the English words in the definition gave the same sense as the similar sense. There may be cases where different English words in the definition may give rise to shortest paths that pass through different senses.

Another measure that may require closer scrutiny is the length of the shortest path. In some cases, the word may be directly linked to the given sense, i.e. path length of only 1. Whereas some words may be linked with the

given sense with a much more circuitous path, let us say path length of 7. The more adjacent nodes are likely to be more similar than distant nodes.

To arrive at a metric that will help us in sorting these two issues, a metric can be used. The metric should have (1) the number of words showing affinity to the same `sense_id` from the given definition in direct proportion and (2) path lengths in inverse proportion.

This graph-based approach was not found to yield reasonable results as a standalone algorithm empirically. This approach may be considered in future work as a supplement to the text-similarity-based approach. The system may be developed in such a way that the word meaning pairs that do not show sufficient similarity score based on text similarity-based algorithms may be subjected to graph-based algorithms to see if it yields some improvement. For our study, we did not go into the graph-based methods in detail.

3.13. Latent Semantic Analysis

For testing whether the Latent Semantic Analysis yields better results or not, we subjected our Gold Standard data to LSA and tried to find out the mapping. We have used the open-source Gensim software to run the LSA for our experiment. Term Frequency – Inverse Document Frequency (TF-IDF) modelling is implemented in Gensim. Thereafter, the data was subjected to LSA with increasing number of features from 64 to 1028. To check whether

the Bag of word (BoW/TF-IDF), Doc2Vec or Dense Embeddings give better results, we ran those models. The results are as below.

Model / Strategy	Dimensions	Precision	Recall	F1-score
LSA	64	75.58%	81.90%	78.61%
LSA	128	79.35%	84.97%	82.06%
LSA	256	81.64%	87.09%	84.28%
LSA	512	83.73%	88.58%	86.09%
LSA	1024	84.58%	89.27%	86.86%
BoW / TF-IDF	Full Vocab	77.97%	85.95%	81.77%
Doc2Vec	128	71.50%	78.54%	74.85%
Dense Embeddings	128 (SVD)	78.36%	84.37%	81.26%

Table 24: Precision, Recall and F1-Score for Latent Semantic Analysis and BoW, Doc2Vec, Dense Embeddings

It is usually not useful to increase the LSA features beyond 1024, unless the dataset is extremely large and with varied vocabulary, which is not the case in our analysis. Therefore, the exploration with more features was not done. Thus, the LSA approach also does not yield better results than the text-based similarity models for the task at hand. LSA-based approach may be used as a fallback mechanism when the text-based similarity models do not yield more

than a threshold similarity score. It is worth noting that BoW/TF-IDF approach gives better precision,

3.14. Frontend for correction submission

The system has been designed to take word-meaning pairs in two dictionaries as input and generate auto-aligned pairs across dictionaries as output, along with some similarity score. A sample is shown in Figure 4.

35	aMSuH	A point or end.	0.742	0.518	61	aMSu	a point, end	0.913	0.713	61	aMSu	a point, end	61	A
36	aMSuH	A small or minute particle.	0.516	0.338	60	aMSu	end of a thread, a minute particle	0.618	0.486	60	aMSu	end of a thread, a minute particle	60	A
37	aMSuH	End of a thread.	0.271	0.034	60	aMSu	end of a thread, a minute particle	0.75	0.386	59	aMSu	thread	59	B
38	aMSuH	A filament, especially of the Soma plant	0.485	0.207	57	aMSu	a filament (especially of the Soma plant)	1.0	0.856	57	aMSu	a filament (especially of the Soma plant)	57	A
39	aMSuH	Garment; decoration.	0.061	0.003	57	aMSu	a filament (especially of the Soma plant)	0.15	0.049	58	aMSu	a kind of Soma libation,	58	E

Figure 4: Visual aid to the annotator alongwith the similarity scores and deltas

Columns are (1) ID_dict1, (2) Headword_dict1, (3) Definition_dict1, (4) Similarity_score with Hindi as the destination language, (5) Delta_score with Hindi as the destination language, (6) ID_dict2 of the most similar entry with Hindi as the destination language, (7) Headword_dict2, (8) Definition_dict2,

(9) Similarity_score with English as the destination language, (10) Delta_score with English as the destination language, (11) ID_dict2 of the most similar entry with English as the destination language, (12) Headword_dict2, (13) Definition_dict2, (14) Final ID_dict2 mapping with given ID_dictd1, (15) Category of the mapping based on confidence of mapping

The last column shows the categories of mapping based on confidence of mapping in accordance with Table 16. The entries that fall in categories D or E will require more attention from a human annotator for manual correction. A frontend has been generated through which the human annotator may submit their corrections. See Figure 5. The entries with low confidence are shown in dark colour. The annotator may provide the correct mapping for the same and submit their correction to the database.

ID : 675, Category: A Meaning : इक्कन, म्यान	ID : 1296 Meaning : A covering, sheath.	675	1296
ID : 676, Category: B Meaning : कपड़ा, वस्त्र	ID : 1297 Meaning : Cloth, clothes;	676	1297
ID : 677, Category: E Meaning : छाजन	ID : 1296 Meaning : A covering, sheath.	677	1296

Figure 5: Frontend for correction submission

Once these corrections are submitted to the database, the dictionaries aligned at the word sense level are generated.

Chapter 4

Synset concordance of Samānāarthaka Sanskrit kośas

The method referred to in the previous part of the thesis was more suited to the dictionaries that are fashioned in the Western style of dictionaries, providing word–definition pairs. The same may be extended to the anekāarthaka kośas too. But the samānāarthaka kośas deserve a different treatment, as the knowledge representation in them is in a different format. The next part of the thesis tries to explore the possibility of creating a concordance of Samānāarthaka Sanskrit kośas.

4.1. Amarakośa

Amarakośa (AMAR) is a well-renowned Sanskrit kośa that 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 is 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.

4.2. Abhidhānacintāmaṇi of Hemacandra

Hemacandra has authored many lexicographic works. The most notable among them are Abhidhānacintāmaṇi - a samānārthaka kośa and Anekārthasaṃgraha - a nānārthaka kośa. Unlike Amarasimha, Hemacandra thought it fit to compose two separate works for synonymic and homonymic lexica. Therefore, in the present comparison, only the synonymic part of Amarakośa is compared with the 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 auto-commentary on this work named Abhidhānacintāmaṇisvopajñavṛtti (ACVH) and a commentary on ABCH, namely Vyutpattiratnākara by Devasāgaragaṇi (AVRD). Both these commentaries provide gender information explicitly.

4.3. Abhidhānacintāmaṇipariśiṣṭa and Abhidhānacintāmaṇiśiloṇcha

Abhidhānacintāmaṇi also has two supplementary treatises, namely Abhidhānacintāmaṇipariśiṣṭa by Hemacandra (ACPH) himself (also known as Śeṣasaṃgrahanāmamālā) and Abhidhānacintāmaṇiśiloṇcha by Jinadeva (ACSJ). These two works supplement words that are missing in Abhidhānacintāmaṇi. ACPH is a compilation of extra words that Hemacandra has provided in his auto-commentary ACVH. It is a drawback of this dictionary 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 Haimanāmamālāśiloṇchadīpikā (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 can have some errors.

4.4. Reason for selection of these kośas

Amarakośa is the most widely used Sanskrit lexicon. ABCH, ACPH and ACSJ put together provide a substantial coverage of Sanskrit synonyms. Amarakośa and commentaries thereon usually follow Pāṇini's grammar, whereas ABCH, ACPH and ACSJ and commentaries thereon (except AVRĐ) follow a different system of grammar known as Haimavyākaraṇa propounded by Hemcandra himself in his work Haimaśabdānuśāsanam. Grammar and lexicons of Hemcandra are widely used in the Jaina literature as a standard text. Therefore, a concordance and comparison between AMAR and ABCH would provide a nice way to synthesise these two systems of grammar and lexicography.

4.5. Limitation of existing digital resources

Currently, the dictionaries which are digitised and available online on CDSD are mostly made in the Western style of lexicography, which list a word and its meanings. Some Sanskrit-Sanskrit dictionaries like Śabdakalpādruma and Vācaspatyam are also made on that pattern. Therefore, there exists no such digital resource that has tried to map the Sanskrit words at a synonym set (synset) level and tried to align various synsets across the Sanskrit samānārthaka kośas. The present work is an attempt to provide a framework for the alignment of 5 Sanskrit samānārthaka kośas. The framework is easily extensible to add new lexical resources as and when required.

4.6. Development of e-resource

The e-resource was developed in two-pronged ways.

1. Generation of a lexicon marked up with synset information and
2. Alignment of synsets across lexica

4.6.1. 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 “<syms>”, and it is followed by ‘word-gender’ pairs separated by a comma. Every synset is also given an extra ID “<eid>” which can help in identifying and aligning this synset across lexica. There are cases where a single ślokārdha may have more than one synset. To take care of such issues, another identifier “<L>” was assigned. It will uniquely identify chunks of text and also the synsets contained in them. Whenever the page-column detail was available, it was provided by the “<pc>” tag. A sample entry from AMAR is shown below for reference.

<L>8<pc>

<eid>28<syms> विष्णु-पुं, नारायण-पुं, कृष्ण-पुं, वैकुण्ठ-पुं, विष्टरश्रवस्-पुं, दामोदर-पुं, हृषीकेश-पुं, केशव-पुं, माधव-
पुं, स्वभू-पुं, दैत्यारि-पुं, पुण्डरीकाक्ष-पुं, गोविन्द-पुं, गरुडध्वज-पुं, पीताम्बर-पुं, अच्युत-पुं, शार्ङ्गिन्-पुं, विष्वक्सेन-पुं, जनार्दन-
पुं, उपेन्द्र-पुं, इन्द्रावरज-पुं, चक्रपाणि-पुं, चतुर्भुज-पुं, पद्मनाभ-पुं, मधुरिपु-पुं, वासुदेव-पुं, त्रिविक्रम-पुं, देवकीनन्दन-पुं, शौरि-
पुं, श्रीपति-पुं, पुरुषोत्तम-पुं, वनमालिन्-पुं, बलिध्वंसिन्-पुं, कंसाराति-पुं, अधोक्षज-पुं, विश्वम्भर-पुं, कैटभजित्-पुं, विधु-
पुं, श्रीवत्सलाञ्छन-पुं, पुराणपुरुष-पुं, यज्ञपुरुष-पुं, नरकान्तक-पुं, जलशायिन्-पुं, विश्वरूप-पुं, मुकुन्द-पुं, मुरमर्दन-
पुं, लक्ष्मीपति-पुं, मुरारि-पुं

<eid>29<syms> वसुदेव-पुं, आनकदुन्दुभि-पुं

विष्णुनारायणः कृष्णो वैकुण्ठो विष्टरश्रवाः ।

दामोदरो हृषीकेशः केशवो माधवः स्वभूः ॥ १८ ॥

दैत्यारिः पुण्डरीकाक्षो गोविन्दो गरुडध्वजः ।

पीताम्बरोऽच्युतः शार्ङ्गी विष्वक्सेनो जनार्दनः ॥ १९ ॥

उपेन्द्र इन्द्रावरजश्चक्रपाणिश्चतुर्भुजः ।

पद्मनाभो मधुरिपुर्वासुदेवस्त्रिविक्रमः ॥ २० ॥

देवकीनन्दनः शौरिः श्रीपतिः पुरुषोत्तमः ।

वनमाली बलिध्वंसी कंसारातिरधोक्षजः ॥ २१ ॥

विश्वम्भरः कैटभजिद्विधुः श्रीवत्सलाञ्छनः ।

पुराणपुरुषो यज्ञपुरुषो नरकान्तकः ।

जलशायी विश्वरूपो मुकुन्दो मुरमर्दनः ।

वसुदेवोऽस्य जनकः स एवानकदुन्दुभिः ॥ २२ ॥

<LEND>

In this case, the `L` id 8 encompasses two `eid`s 28 and 29, referring to Viṣṇu and Vāsudeva, and the text of AMAR, which has these entries. Similarly, the marked-up text was prepared for all four lexical resources.

Corresponding entries in ABCH are as follows:

<L>173<pc>13

<eid>692<syms>विष्णु-पुं, जिष्णु-पुं, जनार्दन-पुं, हरिकेश-पुं, हृषीकेश-पुं, अच्युत-पुं, केशव-पुं, दाशार्ह-
पुं, पुरुषोत्तम-पुं, अब्धिशयन-पुं, उपेन्द्र-पुं, अज-पुं, इन्द्रानुज-पुं, विष्वक्सेन-पुं, नारायण-पुं, जलशय-पुं, नारायण-पुं, श्रीपति-
पुं, दैत्यारि-पुं, पुराणपुरुष-पुं, यज्ञपुरुष-पुं, तार्क्ष्यध्वज-पुं, अधोक्षज-पुं, गोविन्द-पुं, षड्बिन्दु-पुं, मुकुन्द-पुं, कृष्ण-पुं, वैकुण्ठ-
पुं, पद्मेशय-पुं, पद्मनाभ-पुं, वृषाकपि-पुं, माधव-पुं, वासुदेव-पुं, विश्वम्भर-पुं, श्रीधर-पुं, विश्वरूप-पुं, दामोदर-पुं, शौरि-
पुं, सनातन-पुं, विधु-पुं, पीताम्बर-पुं, मार्ज-पुं, जिन-पुं, कुमोदक-पुं, त्रिविक्रम-पुं, जह्नु-पुं, चतुर्भुज-पुं, पुनर्वसु-पुं, शतावर्त-
पुं, गदाग्रज-पुं, स्वभू-पुं, मुञ्जकेशिन्-पुं, वनमालिन्-पुं, पुण्डरीकाक्ष-पुं, बभ्रु-पुं, शशबिन्दु-पुं, पृश्निशृङ्ग-पुं, धरणीधर-
पुं, पाण्डवायन-पुं, सुवर्णबिन्दु-पुं, श्रीवत्स-पुं, देवकीसूनु-पुं, गोपेन्द्र-पुं, विष्टरश्रवस्-पुं, सोमसिन्धु-पुं, जगन्नाथ-पुं, गोवर्धनधर-
पुं, यदुनाथ-पुं, गदाभृत्-पुं, शार्ङ्गभृत्-पुं, चक्रभृत्-पुं, श्रीवत्सभृत्-पुं, शङ्खभृत्-पुं
विष्णुर्जिष्णुजनार्दनौ हरिहृषीकेशाच्युताः केशवो
दाशार्हः पुरुषोत्तमोऽब्धिशयनोपेन्द्रावजेन्द्रानुजौ ।
विष्वक्सेननारायणौ जलशयो नारायणः श्रीपति-
दैत्यारिश्च पुराणयज्ञपुरुषस्ताक्ष्यध्वजोऽधोक्षजः ॥ २१४ ॥
गोविन्दषड्बिन्दुमुकुन्दकृष्णा वैकुण्ठपद्मेशयपद्मनाभाः ।
वृषाकपिर्माधववासुदेवौ विश्वम्भरः श्रीधरविश्वरूपौ ॥ २१५ ॥
दामोदरः शौरिसनातनौ विधुः पीताम्बरो मार्जजिनौ कुमोदकः ।
त्रिविक्रमो जह्नुचतुर्भुजौ पुनर्वसुः शतावर्तगदाग्रजौ स्वभूः ॥ २१६ ॥
मुञ्जकेशिवनमालिपुण्डरीकाक्षबभ्रुशशबिन्दुवेधसः ।
पृश्निशृङ्गधरणीधरात्मभूः पाण्डवायनसुवर्णबिन्दवः ॥ २१७ ॥

श्रीवत्सो देवकीसूनुर्गोपेन्द्रो विष्टरश्रवाः ।

सोमसिन्धुर्जगन्नाथो गोवर्धनधरोऽपि च ॥ २१८ ॥

यदुनाथो गदाशार्ङ्गचक्रश्रीवत्सशङ्खभृत् ।

<LEND>

and

<L>178<pc>13

<eid>728<syns>वसुदेव-पुं, भूकश्यप-पुं, दुन्दु-पुं, आनकदुन्दुभि-पुं

वसुदेवो भूकश्यपो दुन्दुरानकदुन्दुभिः ॥ २२३ ॥

<LEND>

In the present case, the chunks of text could be segregated easily. Therefore, they have separate 'L' ids.

4.6.2. Alignment of synsets across lexica

The methodology adopted for alignment was quite straightforward. It is a semi-automated process.

The automatic part of the process is as follows. As a preprocessing step, the headwords were normalized for words ending with terminal 'क' e.g. 'पारिजातक' was changed to 'पारिजात' in both the dictionaries.

1. The synsets of AMAR and ABCH were checked for words that are common in prātipadika form and also in gender. So (word, gender) tuple was taken as a unit for comparison. The synsets that have the maximum

number of common entities were aligned with each other. For example, in the example cited above, AMAR.eid.28 would have a lot of common (word, gender) pairs with ABCH.eid.692 e.g. (विष्णु, पुं), (जनार्दन, पुं), (हृषीकेश, पुं) etc. Therefore, they will be automatically aligned with each other computationally. It may be cursorily checked by the annotator manually if one wishes.

2. There are also a lot of cases where there are many common words, but the gender may not match exactly. Both the resources may have ascribed different genders to a given word. Therefore, the next step involved mapping the synsets based on the maximum number of headwords matched across the synsets.
3. After these two steps, if there are cases where the number of matching (headword, gender) pairs across the synsets are the same, it was matched to the first occurrence in the second dictionary.
4. After step3, if there are cases where the number of matching headwords across the synsets are the same, it was matched to the first occurrence in the second dictionary.
5. The remaining synsets in both the dictionaries which do not have any mapping based on the above four steps were treated as having no corresponding entries in the other dictionary.

The alignment obtained at the end of every step was evaluated for precision, recall and F1-score. The baseline was treated to be all the synsets being unmapped i.e. every synset in AMAR was shown without a mapping in ABCH and vice versa. The results are as follow.

Step	Precision	Recall	F1-score
Step 0	0.2837	0.4342	0.3431
Step 1	0.7787	0.8318	0.8044
Step 2	0.8293	0.8499	0.8394
Step 3	0.8496	0.8386	0.8441
Step 4	0.8517	0.8381	0.8484

Table 25: Precision, Recall and F1-score for synonymic dictionaries

As can be seen from the above data, it seems that the performance of step1 and step2 account for the majority of the improvement. Thereafter, the addition of step 3 and step 4 bring marginal improvement in F1-score by improving the precision, but at the cost of recall.

Thereafter, a display for the 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.

At a time, only two koshas are considered, and the synonyms are aligned. It becomes the primary source with which the next dictionary of

synonymous words is aligned. This process is repeated as many times as there are dictionaries. Thus, incrementally, new words and new synonyms get added to the resource.

By following the same process, AMAR, ABCH, ACPH and ACSJ and ACVH were added sequentially. ACPH and ACSJ linking was relatively easy, because they provide the reference to the 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 that are unique to ACPH or ACSJ, which are absent from ABCH, their numbers being 142 and 8, respectively. *Abhidhānacintāmaṇisvopajñavṛtti* (ACVH) of Hemacandra provided great help in locating these unique synsets to the appropriate place in the concordance.

The following example will serve as a demonstration regarding the utility of this concordance. The synsets in all four dictionaries regarding the meaning “kāmadeva” (cupid) are shown below.

AMAR -

<eid>31<syms>मदन-पुं, मन्मथ-पुं, मार-पुं, प्रद्युम्न-पुं, मीनकेतन-पुं, कन्दर्प-पुं, दर्पक-पुं, अनङ्ग-पुं, काम-पुं, पञ्चशर-पुं, स्मर-पुं, शम्बरारि-पुं, मनसिज-पुं, कुसुमेषु-पुं, अनन्यज-पुं, पुष्पधन्वन्-पुं, रतिपति-पुं, मकरध्वज-पुं, आत्मभू-पुं, पञ्चबाण-पुं

ABCH -

<eid>733<syms>मदन-पुं, जराभीरु-पुं, अनङ्ग-पुं, मन्मथ-पुं, कमन-पुं, कलाकेलि-पुं, अनन्यज-पुं, अङ्गज-
पुं, मधुदीप-पुं, मार-पुं, मधुसारथि-पुं, स्मर-पुं, विषमायुध-पुं, दर्पक-पुं, काम-पुं, हृच्छय-पुं, प्रद्युम्न-पुं, श्रीनन्दन-पुं, कन्दर्प-
पुं, पुष्पकेतन-पुं, पुष्पेषु-पुं, कुसुमबाण-पुं, पुष्पचाप-पुं, कुसुमधन्वन्-पुं, पुष्पास्त्र-पुं, कुसुमायुध-पुं, मीनकेतन-पुं, मकरकेतन-
पुं, मीनध्वज-पुं, मकरध्वज-पुं, झषकेतन-पुं, झषध्वज-पुं, पञ्चबाण-पुं, विषमेषु-पुं, मनोयोनि-पुं, चेतोभव-पुं, शृङ्गारयोनि-
पुं, शृङ्गारजन्मन्-पुं, सङ्कल्पयोनि-पुं, स्मृतिभू-पुं, आत्मयोनि-पुं, आत्मभू-पुं, मधुसुहृद्-पुं, चैत्रसख-पुं

ACPH -

<eid>733<syms>यौवनोद्भेद-पुं, शिखिमृत्यु-पुं, महोत्सव-पुं, शमान्तक-पुं, सर्वधन्विन्-पुं, रागरज्जु-पुं, प्रकर्षक-
पुं, मनोदाहिन्-पुं, मथन-पुं

ACSJ -

<eid>733<syms>कन्तु-पुं

The concordance generated from the above data is as follows:

HEADWORD	GENDE R	AMAR	ABCH	ACPH	ACSJ
मन्मथ	पुं	✓	✓	x	x
मार	पुं	✓	✓	x	x
प्रद्युम्न	पुं	✓	✓	x	x
मीनकेतन	पुं	✓	✓	x	x
कन्दर्प	पुं	✓	✓	x	x
दर्पक	पुं	✓	✓	x	x
अनङ्ग	पुं	✓	✓	x	x
काम	पुं	✓	✓	x	x
पञ्चशर	पुं	✓	x	x	x
स्मर	पुं	✓	✓	x	x
शम्बरारि	पुं	✓	x	x	x
मनसिज	पुं	✓	x	x	x
कुसुमेषु	पुं	✓	x	x	x
अनन्यज	पुं	✓	✓	x	x
पुष्पधन्वन्	पुं	✓	x	x	x
रतिपति	पुं	✓	x	x	x
मकरध्वज	पुं	✓	✓	x	x
आत्मभू	पुं	✓	✓	x	x

पञ्चबाण	पुं	✓	✓	X	X
जराभीरु	पुं	X	✓	X	X
कमन	पुं	X	✓	X	X
कलाकेलि	पुं	X	✓	X	X
अङ्गज	पुं	X	✓	X	X
मधुदीप	पुं	X	✓	X	X
मधुसारथि	पुं	X	✓	X	X
विषमायुध	पुं	X	✓	X	X
हृच्छय	पुं	X	✓	X	X
श्रीनन्दन	पुं	X	✓	X	X
पुष्पकेतन	पुं	X	✓	X	X
पुष्पेषु	पुं	X	✓	X	X
कुसुमबाण	पुं	X	✓	X	X
पुष्पचाप	पुं	X	✓	X	X
कुसुमधन्वन्	पुं	X	✓	X	X
पुष्पास्त्र	पुं	X	✓	X	X
कुसुमायुध	पुं	X	✓	X	X
मकरकेतन	पुं	X	✓	X	X
मीनध्वज	पुं	X	✓	X	X
झषकेतन	पुं	X	✓	X	X

--	--	--	--	--	--

झषध्वज	पुं	x	✓	x	x
विषमेषु	पुं	x	✓	x	x
मनोयोनि	पुं	x	✓	x	x
चेतोभव	पुं	x	✓	x	x
शृङ्गारयोनि	पुं	x	✓	x	x
शृङ्गारजन्मन्	पुं	x	✓	x	x
सङ्कल्पयोनि	पुं	x	✓	x	x
स्मृतिभू	पुं	x	✓	x	x
आत्मयोनि	पुं	x	✓	x	x
मधुसुहृद्	पुं	x	✓	x	x
चैत्रसख	पुं	x	✓	x	x
यौवनोद्भेद	पुं	x	x	✓	x
शिखिमृत्यु	पुं	x	x	✓	x
महोत्सव	पुं	x	x	✓	x
शमान्तक	पुं	x	x	✓	x
सर्वधन्विन्	पुं	x	x	✓	x
रागरज्जु	पुं	x	x	✓	x
प्रकर्षक	पुं	x	x	✓	x
मनोदाहिन्	पुं	x	x	✓	x
मथन	पुं	x	x	✓	x

कन्तु	पुं	x	x	x	✓

Table 26: Sample entry of the samānāṛthaka concordance

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 new additional words attested in the commentary ACVH were also added to the concordance subsequently.

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

Currently, only the samānāṛthaka kośas are taken up. Anekāṛthaka kośas have a different type of knowledge representation and may not be suitable for 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. Whenever a new dictionary attests a new meaning, the concordance may be extended accordingly.

4.7. Structure of information in Amarakośa

During the investigation for finding alignment between synsets of AMAR and ABCH, the knowledge representation structures of both dictionaries came to the fore. Nair (2011) has already studied the knowledge structure of Amarakośa.

The Amarakośa has arranged its content in the following parts.

1 प्रथमकाण्डम्

- 1.1 स्वर्गवर्गः
- 1.2 व्योमवर्गः
- 1.3 दिग्वर्गः
- 1.4 कालवर्गः
- 1.5 धीवर्गः
- 1.6 शब्दादिवर्गः
- 1.7 नाट्यवर्गः
- 1.8 पातालभोगिवर्गः
- 1.9 नरकवर्गः
- 1.10 वारिवर्गः

2 द्वितीयकाण्डम्

- 2.1 भूमिवर्गः
- 2.2 पुरवर्गः
- 2.3 शैलवर्गः
- 2.4 वनौषधिवर्गः

- 2.5 सिंहादिवर्गः
- 2.6 मनुष्यवर्गः
- 2.7 ब्रह्मवर्गः
- 2.8 क्षत्रियवर्गः
- 2.9 वैश्यवर्गः
- 2.10 शूद्रवर्गः
- 3 तृतीयकाण्डम्
 - 3.1 विशेष्यनिघ्नवर्गः
 - 3.2 सङ्कीर्णवर्गः
 - 3.3 नानार्थवर्गः
 - 3.4 अव्ययवर्गः
 - 3.5 लिङ्गादिसङ्ग्रहवर्गः

4.8. Structure of information in Abhidhānacintāmaṇi

ABCH has arranged its contents in the following manner.

- 1 देवाधिदेवकाण्डः
- 2 देवकाण्डः
- 3 मर्त्यकाण्डः
- 4 तिर्यक्काण्डः

4.1 एकेन्द्रियाः

4.1.1 पृथ्वीकायः

4.1.2 अप्कायः

4.1.3 तेजस्कायः

4.1.4 वायुकायः

4.1.5 वनस्पतिकायः

4.2 द्वीन्द्रियाः

4.3 त्रीन्द्रियाः

4.4 चतुरिन्द्रियाः

4.5 पञ्चेन्द्रियाः

4.5.1 स्थलचरपञ्चेन्द्रियाः

4.5.2 खचरपञ्चेन्द्रियाः

4.5.3 जलचरपञ्चेन्द्रियाः

5 नारककाण्डः

6 सामान्यकाण्डः

7 अव्ययानि

4.9. Comparison between the knowledge representation of Amarakośa and Abhidhānacintāmaṇi

The arrangement of words into various kāṇḍas or Vargas in AMAR and ABCH does not match one-to-one. Based on the empirical observations, the rough correspondence between these two lexicons is as follows.

AMAR	ABCH
No corresponding entries, as देवाधिदेवकाण्ड is dedicated to Jaina deities	१. देवाधिदेवकाण्डः (श्लो. १-८६)
१.१. स्वर्गवर्गः	२. देवकाण्डः (श्लो. ८७-१२६, १७१-२४०)
१.२. व्योमवर्गः	२. देवकाण्डः (श्लो. १६३)
१.३. दिग्वर्गः	२. देवकाण्डः (श्लो. १६४-१७०)
१.४. कालवर्गः	२. देवकाण्डः (श्लो. १२७-१६२)
१.५. धीवर्गः	२. देवकाण्डः (श्लो. ३०८-३२४)
१.६. शब्दादिवर्गः	२. देवकाण्डः (श्लो. २४१-२७८)
१.७. नाट्यवर्गः	२. देवकाण्डः (श्लो. २७८-३०७, ३२५-३३६)
१.८. पातालभोगिवर्गः	४.५.१. स्थलचरपञ्चेन्द्रियाः (श्लो. १३०२-१३१५)
१.९. नरकवर्गः	५. नारककाण्डः (श्लो. १३५८-१३६४)
१.१०. वारिवर्गः	४.१.२. एकेन्द्रियेषु अप्कायः (श्लो. १०६९-१०९६)
२.१. भूमिवर्गः	४.१.१. एकेन्द्रियेषु पृथ्वीकायः (श्लो. ९३५-९७०)
२.२. पुरवर्गः	४.१.१. एकेन्द्रियेषु पृथ्वीकायः (श्लो. ९७१-१०२६)
२.३. शैलवर्गः	४.१.१. एकेन्द्रियेषु पृथ्वीकायः (श्लो. १०२७-१०३६)
२.४. वनौषधिवर्गः	४.१.५. एकेन्द्रियेषु वनस्पतिकायः (श्लो. १११०-१२०१)
२.५. सिंहादिवर्गः (श्लो. १-१२)	४.५.१. स्थलचरपञ्चेन्द्रियाः (श्लो. १२८३-१३०१)

२.५. सिंहादिवर्गः (श्लो. १३)	४.२. द्वीन्द्रियाः (श्लो. १२०२-१२०६)
-	४.३. त्रीन्द्रियाः (श्लो. १२०६-१२०९)
२.५. सिंहादिवर्गः (श्लो. १३-१४)	४.४. चतुरिन्द्रियाः (श्लो. १२१०-१२१६)
२.५. सिंहादिवर्गः (श्लो. १४-४३)	४.५.२. खचरपञ्चेन्द्रियाः (श्लो. १३१६-१३४३)
१.१. स्वर्गवर्गः (श्लो. ५३-५७)	४.१.३. एकेन्द्रियेषु तेजस्कायः (श्लो. १०९७-११०५)
१.१. स्वर्गवर्गः (श्लो. ६१-६३)	४.१.४. एकेन्द्रियेषु वायुकायः (श्लो. ११०६-११०९)
१.१०. वारिवर्गः (श्लो. १७-२५)	४.५.२. जलचरपञ्चेन्द्रियाः (श्लो. १३४३-१३५७)
२.६. मनुष्यवर्गः	३. मर्त्यकाण्डः (श्लो. ३३७-६८८)
२.७. ब्रह्मवर्गः	३. मर्त्यकाण्डः (श्लो. ८०७-८६३)
२.८. क्षत्रियवर्गः	३. मर्त्यकाण्डः (श्लो. ६८९-८०६)
२.९. वैश्यवर्गः	३. मर्त्यकाण्डः (श्लो. ८६४-८९४)
२.१०. शूद्रवर्गः	३. मर्त्यकाण्डः (श्लो. ८९४-९३४)
३.१. विशेष्यनिघ्नवर्गः	Spread throughout the work.
३.२. सङ्कीर्णवर्गः	Spread throughout the work.
३.३. नानार्थवर्गः	Absent, as there is a separate work authored by Hemacandra for Nānārtha.
३.४. अव्ययवर्गः	७. अव्ययानि (श्लो. १५२५-१५४२)
३.५. लिङ्गादिसङ्ग्रहवर्गः	Absent, as there is a separate work authored by Hemacandra for

	Liṅgānuśāsanam.
Spread throughout the work.	६. सामान्यकाण्डः (श्लो. १३६५-१५२४)
२.९. वैश्यवर्गः (श्लो. ९१-१११)	४.१.१. एकेन्द्रियेषु पृथ्वीकायः (श्लो. १०३६-१०६८) metals and minerals
२.८. क्षत्रियवर्गः (श्लो. ३४-५०)	४.१.५. स्थलचरपञ्चेन्द्रियाः (श्लो. १२१७-१२५२) Entries related to elephant and horse
२.९. वैश्यवर्गः (श्लो. ५०-७७)	४.१.५. स्थलचरपञ्चेन्द्रियाः (श्लो. १२५३-१२८३) Domesticated animals

Table 27: Mapping of knowledge representation between Amarakośa and Abhidhānacintāmaṇi

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

4.10. Statistics of the concordance

1. Total number of synsets: 5489
2. Synsets in individual dictionaries AMAR, ABCH, ACPH, ACSJ and ACVH
 1. AMAR – 3692
 2. ABCH – 4791
 3. ACPH – 347

4. ACSJ – 534
5. ACVH – 591
3. Synset co-occurrence across dictionaries

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 28: Synset co-occurrence across samānārthaka kośas

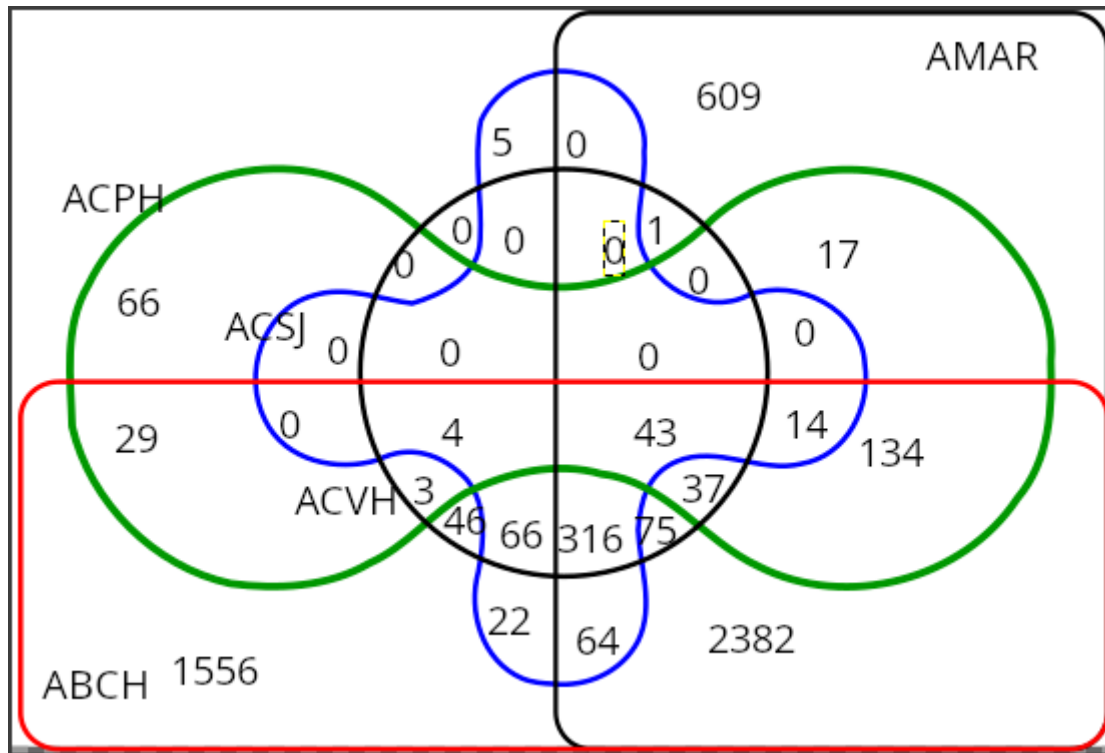


Figure 6: Edwards-Venn diagram showing synset co-occurrence across dictionaries

4. Total number of Word-gender pairs in concordance: 20975
5. Word-gender pairs in individual dictionaries
 1. AMAR – 10157
 2. ABCH – 14704
 3. ACPH – 1506
 4. ACSJ – 693
 5. ACVH - 949
6. Word-gender pair co-occurrence across dictionaries

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 29: Word-gender pair co-occurrence across samānārthaka kośas

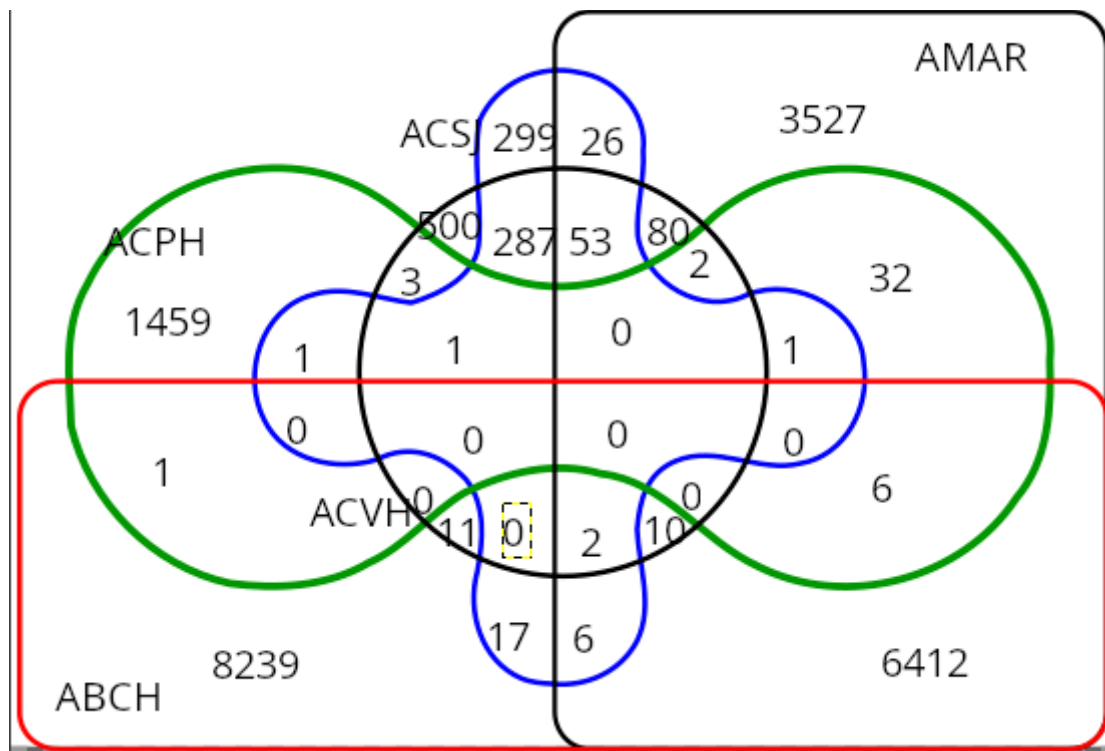


Figure 7: Edwards-Venn diagram showing (word, gender) pair co-occurrence across dictionaries

7. Gender differences across dictionaries

The cases where the gender of a given word in a given synset is different in different dictionaries are noted in Annexure 1. They may be manually examined to decide whether there is some data entry error, or it genuinely is a case where lexicographers differ regarding the gender of the word.

8. Differences arising due to the 'ka' suffix at the end

They refer to the same word, but one with the reflexive 'ka' suffix at the end and one without it, e.g. पारिजातक-पारिजात, डमरु-डमरुक, etc. If these cases are aligned, there may be greater correspondence between the dictionaries. Currently, they are treated as separate words, but in effect, it is not a separate word. पारिजातक refers to पारिजात itself. It is required to decide whether this change is to be made in the data itself or if it can be aligned during the computational process in the next step. The complete list is given in Annexure 2.

4.11. Differences between AMAR, ABCH, ACPH, ACSJ and ACVH dictionaries

The concordance generated was manually examined to identify the differences between constituent dictionaries. To demystify the problem, help was obtained from commentaries, too. The analysis is presented in Annexure 3.

4.12. The difference between lexicographers

As the analysis of concordance brought forth many differences between lexicographers, we decided to study further the differences of opinion between lexicographers. It was not restricted to Amarasiṃha and Hemacandra. The references to various lexicographers were searched from

ACVH, and their content was studied to identify the differences of opinion and to check whether they have any bearing on the concordance. The observations are noted in Annexure 4.

Chapter 5

Conclusion and Way Ahead

Undertaking the task of mapping dictionaries at the level of word sense seems daunting at first, but after experimenting with a few dictionary pairs, it was only 11-14% of word senses that required manual examination by a human expert. The thesis has given a few similarity score metrics and also an evaluation method to ascertain the confidence level of machine-mapped dictionary entries.

Once a quick implementation having an accuracy of 86-89% is created by the machine, human annotators/users can be given an option to change the mapping if they feel that the mapping generated by the machine is incorrect. It holds immense potential to expand sense-mapped text resources from one language to another. It will particularly help the users of languages that have scarce resources, e.g. a Sanskrit work which has been disambiguated and sense-mapped in English with the help of a Sanskrit-English dictionary can be extended to the users of, say, the French language by mapping the Sanskrit-English dictionary to a Sanskrit-French dictionary at the word sense level.

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

The following are a few directions in which further work may be undertaken in future.

We can explore the possibility of using graph-based similarity scores or semantic measures such as LSA or ESA to find out similarity in cases where text-based similarity scores are below a threshold. These approaches are computationally heavy and may require more computational resources.

In the present thesis, the thresholds of similarity scores were chosen empirically or, rather, arbitrarily. It may be possible to learn these thresholds by optimising its F-scores. As the gold standard (training data) is quite small, this exercise has not yet been tried. Once we have large manually validated data, it will be worthwhile to find out the optimum thresholds with statistical methods.

The present methodology can be expanded to other language pairs and check whether findings in different language pairs are similar or otherwise.

The effect of the quality of translation services like Google Translate between different language pairs may add a cascading effect on the performance.

The concordance of samānāṛthaka kośa provides a useful tool to see all the synonyms of a given concept (synset) at a single point. This can be expanded with the inclusion of more and more samānāṛthaka 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. This can be taken up as a systematic approach to highlight differences between different lexical resources.

Once we have enough lexical resources in the concordance, we may be able to understand when new words were added in or dropped from a synset. This may give us some idea about the development of language over a period of time. Over a period of time, some new words may come in vogue and some may fade away in oblivion. Comparative analysis between two lexical resources compiled in particular time period will give interesting insight into this linguistic phenomenon.

Annexure 1

Gender differences between AMAR and ABCH

Headword	amar_eid	amar_gender	abch_eid	abch_gender
नाक	1	पुं	341	पुंक्ली
त्रिदिव	1	पुं	341	पुंक्ली
सुदर्शन	46	पुंक्ली	724	पुं
पिनाक	62	पुं	667	पुंक्ली
अर्चिस्	125	स्त्री	3310	स्त्रीक्ली
अतिवेल	145	क्ली	4469	पुं
निर्भर	145	क्ली	4469	पुं
पद्म	154	पुंक्ली	653	पुं
वर्ष	223	क्ली	564	पुंक्ली
करक	228	पुंस्त्री	566	त्रि
नीहार	243	पुं	3228	पुंक्ली
तुषार	243	पुं	3228	पुंक्ली
हिम	243	क्ली	3228	पुंक्ली
शीत	245	क्ली	4138	पुं
तारा	251	स्त्री	402	पुंस्त्री
तारका	251	स्त्रीक्ली	402	त्रि

मृगशिरस्	259	क्ली	407	पुंक्ली
लग्न	269	क्ली	433	पुंक्ली
मण्डल	278	क्ली	392	त्रि
रुच्	280	स्त्री	389	क्ली
भास्	280	स्त्री	389	पुंस्त्री
समय	285	पुं	469	पुंक्ली
दिन	288	क्ली	489	पुंक्ली
प्रत्यूष	289	पुं	490	पुंक्ली
साय	290	पुं	492	क्ली
गणरात्र	300	क्ली	499	पुंक्ली
तपस्	333	पुं	527	पुंक्ली
शुक्र	337	पुं	531	पुंक्ली
वसन्त	345	पुं	540	पुंक्ली
अब्द	351	पुं	546	पुंक्ली
पैत्र	352	पुं	547	क्ली
दैवत	353	पुं	548	क्ली
ब्राह्म	354	पुं	550	क्ली
प्रकाण्ड	364.02	क्ली	4319	पुंक्ली
दैव	366	क्ली	4121	पुंक्ली
जन्तु	377	पुं	4088	पुंक्ली

द्वापर	389	पुं	4111	पुंक्ली
मेचक	426	पुं	4166	पुंक्ली
नामन्	466	क्ली	854	पुंक्ली
आग्नेडित	480	क्ली	876	त्रि
काकु	482	स्त्री	4212	पुंस्त्री
सुप्रलाप	494	पुं	904	क्ली
सूनृत	506	त्रि	323	क्ली
मणित	515	त्रि	4202	क्ली
वितथ	518	त्रि	867	क्ली
अनृत	518	त्रि	867	क्ली
कोलाहल	524	पुं	4188	पुंक्ली
कल	537	त्रि	4209	पुं
मन्द्र	538	त्रि	4178	पुं
मन्द्र	538	त्रि	4208	पुं
तार	539	त्रि	4180	पुं
तार	539	त्रि	4207	पुं
पटह	552	पुंक्ली	975	पुं
आङ्गिक	594	क्ली	923	पुं
सात्त्विक	594.01	क्ली	924	पुं

शृङ्गार	595	पुं	978	पुंक्ली
बीभत्स	608	त्रि	984	पुं
रौद्र	611	त्रि	981	पुं
दर	612	पुं	1006	पुंक्ली
मान	616	पुं	1044	पुंक्ली
शील	629	क्ली	2568	पुंक्ली
स्नेह	631	पुं	4116	पुंक्ली
काम	632	पुं	1325	पुंक्ली
अहि	683	पुं	3934	पुंस्त्री
कञ्चुक	688	पुं	3967	पुंक्ली
गरल	689	क्ली	3594	पुंक्ली
काकोल	690	पुंक्ली	3602	पुं
कालकूट	691	पुंक्ली	3597	पुं
हलाहल	692	पुंक्ली	3595	पुं
दुर्गति	701	स्त्री	4071	पुंस्त्री
सिन्धु	714	पुं	3230	पुंस्त्री
वार्	717	क्ली	3222	स्त्री
तट	725	त्रि	3246	क्ली
अन्तरीप	728	पुंक्ली	3248	क्ली

जम्बाल	731	पुं	3279	पुंक्ली
प्रोष्ठी	766	पुंस्त्री	4040	स्त्री
कुलीर	782	पुं	4054	पुंक्ली
निपान	801	क्ली	3286	पुंक्ली
कूप	802	पुं	3283	पुंस्त्री
अखात	806	पुंक्ली	3291	क्ली
कासार	808	पुं	3293	पुंक्ली
आलवाल	813	क्ली	3296	पुंक्ली
प्रणाली	830	पुंस्त्री	3275	त्रि
उत्पल	835	क्ली	3485	पुंक्ली
शैवाल	840	क्ली	3496	पुंक्ली
शैवल	840	पुं	3496	पुंक्ली
नलिन	843	क्ली	3481	पुंक्ली
कमल	843	क्ली	3481	पुंक्ली
मृणाल	847	पुंक्ली	3491	त्रि
किञ्जल्क	850	पुं	3492	पुंक्ली
विष्टप	862	क्ली	4084	पुंक्ली
भुवन	862	क्ली	4084	पुंक्ली
नीवृत्	869	पुं	2893	स्त्री

कच्छ	876	पुं	2907	पुंक्ली
विपथ	892	पुं	3004	क्ली
कापथ	892	पुं	3004	क्ली
गव्यूति	897	स्त्री	2719	पुंस्त्री
रोदसी	901	स्त्री	2867	क्लीद्वि
पुरी	903	स्त्रीक्ली	2970	त्रि
नगरी	903	स्त्रीक्ली	2970	स्त्री
विपणि	908	पुंस्त्री	3057	स्त्री
गृह	915	क्ली	3018	पुंक्ली
भवन	915	क्ली	3018	पुंक्ली
मन्दिर	915	क्ली	3018	स्त्रीक्ली
मन्दुरा	920	स्त्री	3043	स्त्रीक्ली
मठ	923	पुं	3027	त्रि
वातायन	929	क्ली	3088	पुंक्ली
अट्ट	940	पुं	2995	पुंक्ली
अलिन्द	941	पुं	3080	पुंक्ली
पटल	949	क्ली	3082	त्रि
शिखर	984	क्ली	3152	पुंक्ली
शृङ्ग	984	क्ली	3152	पुंक्ली

कन्दर	990	पुंस्त्री	3155	त्रि
गह्वर	991	क्ली	3156	पुंक्ली
श्रेणी	1008	स्त्री	4272	पुंस्त्री
दल	1040	क्ली	3360	पुंक्ली
छद	1040	पुं	3360	पुंक्ली
फल	1043	क्ली	3376	पुंक्ली
क्षारक	1047	पुं	3367	पुंक्ली
पुष्प	1051	क्ली	3366	पुंक्ली
जम्भ	1069	पुं	3448	पुंक्ली
वेतस	1081	पुं	3401	पुंस्त्री
शिगु	1087	पुं	3388	पुंक्ली
असन	1117	पुं	3426	पुंक्ली
साल	1118	पुं	3404	पुंक्ली
इङ्गुदी	1121	स्त्री	3422	त्रि
आमलकी	1141	त्रि	3431	स्त्री
तमाल	1165	पुं	3435	पुंक्ली
करीर	1191	पुं	3450	पुंक्ली
चुक्र	1305	पुं	1299	पुंक्ली
मुस्ता	1346	स्त्री	3584	त्रि

नड	1353	पुं	3583	पुंक्ली
रौहिष	1362	क्ली	3579	पुंक्ली
नालिकेर	1372	पुं	3456	पुंस्त्री
शलल	1392	क्ली	3920	त्रि
कन्दली	1398.01	स्त्री	3912	त्रि
आखु	1419	पुं	3928	पुंस्त्री
वृश्चिक	1427	पुं	3664	पुंस्त्री
अलिन्	1477	पुं	3666	पुंस्त्री
द्विरेफ	1477	पुं	3666	पुंस्त्री
भृङ्ग	1477	पुं	3666	पुंस्त्री
भ्रमर	1477	पुं	3666	पुंस्त्री
इन्दिन्दिर	1477	पुं	3666	पुंस्त्री
चञ्चरीक	1477	पुं	3666	पुंस्त्री
रोलम्ब	1477	पुं	3666	पुंस्त्री
गरुत्	1497	पुं	3970	पुंक्ली
तनूरुह	1497	क्ली	3970	पुंक्ली
अण्ड	1501	क्ली	3973	पुंक्ली
नीड	1502	पुंक्ली	3974	त्रि
वार	1506	पुं	4214	पुंक्ली

यूथ	1510	पुंक्ली	4215	क्ली
कलल	1646	पुंक्ली	1678	क्ली
पलित	1654	क्ली	1758	पुंक्ली
अमांस	1663	पुं	1383	त्रि
दुर्बल	1663	पुं	1383	त्रि
मांसल	1664	पुं	1382	त्रि
अंसल	1664	पुं	1382	त्रि
व्याधि	1690	पुं	1427	स्त्री
पिटक	1700	त्रि	1437	पुं
पादवल्मीक	1713	क्ली	1435	पुंक्ली
मांस	1737	क्ली	1906	पुंक्ली
आमिष	1737	क्ली	1906	पुंक्ली
मस्तिष्क	1746	क्ली	1912	पुंक्ली
किट्ट	1747	क्ली	1927	पुंक्ली
मल	1747	पुंक्ली	1927	क्ली
पुरीतत्	1748	पुंक्ली	1856	क्ली
गुल्म	1749	पुं	1855	पुंक्ली
कुल्य	1759	क्ली	1913	पुंक्ली
कङ्काल	1760	पुं	1919	पुंक्ली

शरीर	1765	क्ली	1742	पुंक्ली
गुल्फ	1768	पुं	1883	पुंस्त्री
पार्ष्णि	1769	पुं	1886	पुंस्त्री
ऊरु	1772	पुं	1876	पुंस्त्री
गुद	1774	क्ली	1873	पुंक्ली
कट	1776	पुं	1861	पुंक्ली
श्रोणि	1777	स्त्री	1861	पुंस्त्री
भग	1783	क्ली	1867	पुंक्ली
शिशन	1784	पुं	1868	पुंक्ली
मेढ्र	1784	पुं	1868	पुंक्ली
मुष्क	1785	पुं	1872	पुंक्ली
वृषण	1785	पुं	1872	पुंक्ली
दोस्	1798	पुं	1809	पुंक्ली
कफोणि	1799	पुंस्त्री	1812	स्त्री
कूर्पर	1799	पुंस्त्री	1812	पुं
नख	1811	पुंक्ली	1825	पुं
नखर	1811	पुंक्ली	1825	पुं
मुख	1829	क्ली	1761	पुंक्ली
हनु	1834	स्त्री	1789	पुंस्त्री

गोधि	1839	पुं	1762	स्त्री
श्रवण	1847	पुं	1763	पुंक्ली
तनूरुह	1861	क्ली	1922	पुंक्ली
वेष	1863	पुं	1937	पुंक्ली
हार	1878.01	पुं	1995	पुंस्त्री
पारिहार्य	1887	पुं	2013	पुंक्ली
केयूर	1888	पुंक्ली	2012	क्ली
अङ्गद	1888	पुंक्ली	2012	क्ली
अङ्गुलीयक	1889	पुंक्ली	2014	क्ली
कङ्कण	1891	क्ली	2013	पुंक्ली
करभूषण	1891	क्ली	2013	पुंक्ली
रशना	1892	स्त्री	2016	स्त्रीक्ली
तुलाकोटि	1894	पुं	2019	पुंक्ली
हंसक	1894	पुं	2019	पुंक्ली
पादकटक	1894	पुं	2019	पुंक्ली
कार्पास	1902	त्रि	2031	क्ली
बादर	1902	त्रि	2031	क्ली
कौशेय	1903	त्रि	2032	क्ली
राङ्गव	1904	त्रि	2033	क्ली

कर्पट	1915	पुं	2048	पुंक्ली
वस्त्र	1916	क्ली	2020	पुंक्ली
पट	1917	पुंक्ली	2020	त्रि
निचोल	1919	त्रि	2049	पुं
कम्बल	1920	पुं	2034	पुंस्त्री
कूर्पासक	1923	पुंक्ली	2043	पुं
आप्रपदीन	1926	त्रि	2051	क्ली
अगुरु	1942	क्ली	1952	पुंक्ली
कालागुरु	1943	क्ली	1954	पुं
राल	1945	पुं	1965	पुंक्ली
तुरुष्क	1947	पुं	1966	पुंक्ली
मृगनाभि	1949	पुं	1960	स्त्री
मलयज	1952	पुं	1955	पुंक्ली
तैलपर्णिक	1953	क्ली	1956	पुं
गोशीर्ष	1953	क्ली	1956	पुं
कन्दुक	1973	पुं	2081	पुंक्ली
दीप	1974	पुं	2075	पुंक्ली
दर्पण	1980	पुंक्ली	2067	पुं
बलि	2035	पुं	2495	पुंस्त्री

सदस्	2038	स्त्रीक्ली	1489	स्त्री
पूत	2079	क्ली	2538	पुंक्ली
और्ध्वदेहिक	2083	त्रि	1187	क्ली
कुतप	2087	पुंक्ली	495	पुं
सनि	2089	स्त्री	1217	पुंस्त्री
आनुपूर्वी	2106	स्त्री	4465	स्त्रीक्ली
तपःक्लेशसह	2124	पुं	2462	त्रि
दान्त	2124	पुं	2462	त्रि
पवित्र	2131	पुं	4301	त्रि
पूत	2131	पुं	4301	त्रि
वपन	2146	त्रि	2827	क्ली
उपवीत	2147	क्ली	2571	पुंक्ली
ब्रात्य	2160	पुं	2593	त्रि
अवकीर्णिन्	2163	पुं	2592	त्रि
क्षतव्रत	2163	पुं	2592	त्रि
अमित्र	2201	क्ली	2202	पुं
मित्र	2203	पुं	2204	क्ली
अषडक्षीण	2235	त्रि	2251	क्ली
भागधेय	2248	पुं	2260	पुंस्त्री

बलि	2248	पुं	2260	पुंस्त्री
उपग्राह्य	2251	क्ली	2235	पुं
छत्र	2263	क्ली	2172	त्रि
कुञ्जर	2274	पुं	3680	पुंक्ली
गात्र	2297	क्ली	3725	स्त्रीक्ली
अन्दुक	2301	पुं	3728	पुंक्ली
सृणि	2302	स्त्री	3733	पुंस्त्री
कुथ	2305	पुंस्त्री	2057	त्रि
आश्वीन	2323	क्ली	3796	पुं
शफ	2332	क्ली	3779	पुंक्ली
लाङ्गूल	2333	क्ली	3780	पुंक्ली
रथ	2336	पुं	2280	पुंस्त्री
शकट	2339	पुंक्ली	2286	त्रि
कम्बलिवाह्यक	2340	क्ली	2287	पुंक्ली
पाण्डुकम्बलिन्	2344	त्रि	2290	पुं
चक्र	2349	क्ली	2293	पुंक्ली
प्रधि	2350	पुं	2294	पुंस्त्री
वरुथ	2353	पुं	2303	पुंक्ली
कूबर	2354	पुं	2297	पुंक्ली

युगन्धर	2354	पुं	2297	पुंक्ली
पत्र	2357	क्ली	2308	पुंक्ली
कञ्चुक	2369	पुं	2325	पुंक्ली
अधिकाङ्ग	2370	पुंक्ली	2326	क्ली
काण्डपृष्ठ	2378	त्रि	2332	पुं
आयुधीय	2378	त्रि	2332	पुं
आयुधिक	2378	त्रि	2332	पुं
कृतहस्त	2379	त्रि	2340	पुं
कृतपुङ्ख	2379	त्रि	2340	पुं
धानुष्क	2381	त्रि	2338	पुं
काण्डवत्	2382	त्रि	2339	पुं
काण्डीर	2382	त्रि	2339	पुं
शाक्तीक	2383	त्रि	2336	पुं
याष्टीक	2384	त्रि	2337	पुं
पारश्वधिक	2385	त्रि	2334	पुं
नैस्त्रिंशिक	2386	त्रि	2335	पुं
प्रासिक	2387	त्रि	2333	पुं
कौन्तिक	2387	त्रि	2333	पुं
पताकिन्	2389	त्रि	2320	पुं

वैजयन्तिक	2389	त्रि	2320	पुं
अनुप्लव	2390	त्रि	1535	पुं
सहाय	2390	त्रि	1535	पुं
अनुचर	2390	त्रि	1535	पुं
अभिचर	2390	त्रि	1535	पुं
पुरोग	2391	त्रि	1538	पुं
अग्रेसर	2391	त्रि	1538	पुं
प्रष्ठ	2391	त्रि	1538	पुं
अग्रतःसर	2391	त्रि	1538	पुं
पुरःसर	2391	त्रि	1538	पुं
पुरोगम	2391	त्रि	1538	पुं
पुरोगामिन्	2391	त्रि	1538	पुं
मन्दगामिन्	2392	त्रि	1532	पुं
मन्थर	2392	त्रि	1532	पुं
जङ्घाल	2393	त्रि	1528	पुं
अतिजव	2393	त्रि	1528	पुं
जङ्घाकरिक	2394	त्रि	1529	पुं
जाङ्घिक	2394	त्रि	1529	पुं
त्वरित	2395	त्रि	1530	पुं

जवन	2395	त्रि	1530	पुं
जय्य	2396	त्रि	2412	पुं
जेय	2397	त्रि	2413	पुं
जेतृ	2398	त्रि	2411	पुं
अभ्यमित्य	2399	त्रि	2407	पुं
अभ्यमित्रीय	2399	त्रि	2407	पुं
अभ्यमित्रीण	2399	त्रि	2407	पुं
ऊर्जस्वल	2400	त्रि	2409	पुं
ऊर्जस्विन्	2400	त्रि	2409	पुं
उरस्वत्	2401	त्रि	2408	पुं
उरसिल	2401	त्रि	2408	पुं
रथिर	2402	त्रि	2311	पुं
रथिक	2402	त्रि	2311	पुं
रथिन्	2402	त्रि	2311	पुं
कामङ्गामिन्	2403	त्रि	1533	पुं
अनुकामीन	2403	त्रि	1533	पुं
अत्यन्तीन	2404	त्रि	1534	पुं
जेतृ	2406	त्रि	2411	पुं
जिष्णु	2406	त्रि	2411	पुं

जित्तर	2406	त्रि	2411	पुं
सांयुगीन	2407	त्रि	2410	पुं
चक्र	2408	क्ली	2262	पुंक्ली
गुल्म	2414	पुं	2270	पुंक्ली
आयुध	2423	क्ली	2344	पुंक्ली
शस्त्र	2424	पुंक्ली	2344	क्ली
अस्त्र	2425	पुंक्ली	2344	क्ली
कोदण्ड	2426	क्ली	2350	पुंक्ली
इष्वास	2426	पुं	2350	पुंक्ली
तल	2430	क्ली	2354	स्त्रीक्ली
बाण	2436	पुं	2362	पुंक्ली
शर	2436	पुं	2362	पुंक्ली
इषु	2436	पुंस्त्री	2362	त्रि
तूण	2441	पुं	2375	त्रि
रिष्टि	2442	पुं	2376	पुंस्त्री
मुद्गर	2447	पुं	2387	पुंस्त्री
रेणु	2469	पुंस्त्री	2967	स्त्री
चूर्ण	2470	क्ली	2966	पुंक्ली
समुत्पिञ्ज	2471	पुं	1176	त्रि

पिञ्जल	2471	पुं	1176	त्रि
बल	2477	क्ली	2421	पुंक्ली
जन्य	2480	पुं	2422	पुंक्ली
सङ्ख्य	2480	क्ली	2422	पुंक्ली
समर	2480	पुं	2422	पुंक्ली
अनीक	2480	पुं	2422	पुंक्ली
रण	2480	पुं	2422	पुंक्ली
संयुग	2480	पुं	2422	पुंक्ली
संयत्	2480	स्त्री	2422	क्ली
पटह	2488	पुं	2424	पुंक्ली
आडम्बर	2488	पुं	2424	पुंक्ली
अजन्य	2491	क्ली	467	पुंक्ली
कुणप	2507	क्ली	1743	पुंक्ली
जीव	2511	पुं	4090	त्रि
जीवातु	2513	पुं	4091	पुंक्ली
कुसीद	2526	क्ली	2688	पुं
माष्य	2539	क्ली	2958	त्रि
उम्य	2540	क्ली	2955	त्रि
औमीन	2540	क्ली	2955	त्रि

अणव्य	2541	क्ली	2953	त्रि
भङ्ग्य	2542	क्ली	2954	त्रि
कौद्रवीण	2544	क्ली	2951	त्रि
बीजाकृत	2551	क्ली	2962	त्रि
केदार	2560	पुं	2945	पुंक्ली
मेधि	2573	पुं	2740	पुंस्त्री
मसूर	2579	पुं	3506	पुंस्त्री
बुस	2598	क्ली	3546	पुंक्ली
शूर्प	2611	पुंक्ली	3109	क्ली
पिट	2614	पुं	3107	पुंक्ली
आरालिक	2618	त्रि	2187	पुं
सूद	2618	त्रि	2187	पुं
औदनिक	2618	त्रि	2187	पुं
गुण	2618	त्रि	2187	पुं
अम्बरीष	2624	क्ली	3114	पुंक्ली
भ्राष्ट्र	2624	पुं	3114	पुंक्ली
अलिञ्जर	2626	पुं	3121	पुंक्ली
मणिक	2626	पुं	3121	पुंक्ली
पिठर	2628	पुं	3111	क्ली

कुतुप	2634	पुं	3128	पुंक्ली
पात्र	2635	क्ली	3132	त्रि
चुक्र	2641	क्ली	1299	पुंक्ली
जीरक	2643	पुं	1308	पुंक्ली
आर्द्रक	2645	क्ली	3574	पुंक्ली
कुस्तुम्बुरु	2646	क्ली	1302	पुंक्ली
वशिर	2652	क्ली	2877	पुं
सैन्धव	2653	पुंक्ली	2878	क्ली
सौवर्चल	2656	क्ली	2881	पुंक्ली
धानाः	2675	स्त्री	1253	स्त्रीब
पिष्टक	2676	पुं	1245	पुंक्ली
कृसर	2684	क्ली	1244	पुंस्त्री
क्षीर	2688	क्ली	1265	पुंक्ली
मण्ड	2698	क्ली	1238	पुंक्ली
गो	2715	पुं	3810	पुंस्त्री
गो	2739	पुं	3835	पुंस्त्री
आपीन	2761	क्ली	3861	पुंक्ली
मेष	2773	पुं	3873	पुंक्ली
वस्न	2782	पुं	2639	पुंक्ली

पल	2798	क्ली	2704	पुंक्ली
आचित	2804	पुंक्ली	2709	पुं
आढक	2807	पुंक्ली	2712	त्रि
प्रस्थ	2813	पुं	2711	पुंक्ली
धन	2816	क्ली	650	पुंक्ली
हिरण्य	2816	क्ली	650	पुं
कोश	2817	पुं	3177	क्ली
अश्मगर्भ	2820	पुं	3212	क्ली
पद्मराग	2821	पुं	3213	पुंक्ली
स्वर्ण	2825	क्ली	3176	पुंक्ली
सुवर्ण	2825	क्ली	3176	पुंक्ली
हिरण्य	2825	क्ली	3176	पुंक्ली
पारद	2835	पुं	3188	पुंक्ली
कुसुम्भ	2853	क्ली	3478	पुंक्ली
कापोत	2860	क्ली	2885	पुं
श्वेतमरिच	2862	क्ली	3389	पुं
भृतक	2914	पुं	1162	त्रि
भृतिभुज्	2914	पुं	1162	त्रि
कर्मकर	2914	पुं	1162	त्रि

वैतनिक	2914	पुं	1162	त्रि
वार्तावह	2915	पुं	1168	त्रि
वैवधिक	2915	पुं	1168	त्रि
भारवाह	2916	पुं	1167	त्रि
भारिक	2916	पुं	1167	त्रि
भृत्य	2918	पुं	1161	त्रि
दास	2918	पुं	1161	त्रि
नियोज्य	2918	पुं	1161	त्रि
किङ्कर	2918	पुं	1161	त्रि
भुजिष्य	2918	पुं	1161	त्रि
परिचारक	2918	पुं	1161	त्रि
पराचित	2919	पुं	1161	त्रि
परिस्कन्द	2919	पुं	1161	त्रि
परजात	2919	पुं	1161	त्रि
परैधित	2919	पुं	1161	त्रि
वटी	2942	त्रि	2842	स्त्री
शाण	2960	पुं	2778	पुंस्त्री
मैरेय	2982	क्ली	2763	पुंक्ली
इभ्य	3021	पुं	1157	त्रि

धनिन्	3021	पुं	1157	त्रि
स्वामिन्	3022	पुं	1160	त्रि
ईश्वर	3022	पुं	1160	त्रि
पति	3022	पुं	1160	त्रि
ईशितृ	3022	पुं	1160	त्रि
अधिभू	3022	पुं	1160	त्रि
नायक	3022	पुं	1160	त्रि
नेतृ	3022	पुं	1160	त्रि
प्रभु	3022	पुं	1160	त्रि
परिवृढ	3022	पुं	1160	त्रि
अधिप	3022	पुं	1160	त्रि
कुटुम्बव्यापृत	3024	पुं	1481	त्रि
अभ्यागारिक	3024	पुं	1481	त्रि
उपाधि	3024	पुं	1481	त्रि
लक्ष्मण	3030	पुं	1156	त्रि
स्निग्ध	3031	पुं	1480	त्रि
कर्मकार	3045	त्रि	1163	पुं
देवद्रूयञ्च	3095	त्रि	1365	अ
विष्वद्रूयञ्च	3096	त्रि	1364	अ

सधूयञ्च	3097	त्रि	1366	अ
तिर्यञ्च	3098	त्रि	1367	अ
उद्धिद्	3157	त्रि	4067	पुं
उद्धिज्ज	3157	त्रि	4067	पुं
उद्धिद	3157	त्रि	4067	पुं
रेफ	3161	पुं	4322	त्रि
प्रधान	3168	क्ली	4307	त्रि
अणु	3174	पुं	4281	त्रि
भूयस्	3176	अ	4280	त्रि
सवेश	3183	त्रि	4335	पुं
वृत्त	3189	क्ली	4368	त्रि
दृढसन्धि	3204	त्रि	4378	पुं
संहत	3204	त्रि	4378	क्ली
अनुपद	3210	क्लीअ	4349	त्रि
अनन्यवृत्ति	3213	त्रि	4351	स्त्री
अन्त	3215	पुंक्ली	4353	त्रि
सामान्य	3218	क्ली	4377	त्रि
स्वनित	3267	त्रि	4196	क्ली
पारायण	3325	त्रि	2554	क्ली

डिम्ब	3386	पुं	2437	पुंक्ली
जागरा	3404	पुंस्त्री	1363	स्त्री
सङ्गम	3445	पुं	4473	पुंक्ली
मार्गण	3446	क्ली	1216	त्रि
अवसान	3468	क्ली	1063	पुं
मध्य	5518	त्रि	4576	क्ली
अल्प	5544	त्रि	4589	पुं
महत्	5545	त्रि	4590	पुं

Annexure 2

**The difference between lexica is only due to
'ka' at the end**

amar_headword	amar_eid	abch_headword	abch_eid
पारिजातक	106	पारिजात	619
गुणवृक्षक	743	गुणवृक्ष	2679
केनिपातक	745	केनिपात	2684
कर्कटक	782	कर्कट	4054
खात	805	खातक	3292
देवखातक	806	देवखात	3291
गोस्थानक	885	गोस्थान	2942
ध्रुव	1020	ध्रुवक	3356
गुच्छक	1049	गुच्छ	3370
करञ्जक	1126	करञ्ज	3412
अतिमुक्त	1175	अतिमुक्तक	3438
हयमारक	1190	हयमार	3399
मातुलुङ्गक	1194	मातुलुङ्ग	3449
रसोनक	1320	रसोन	3565
तेजनक	1352	तेजन	3581

कान्तारक	1357.01	कान्तार	3588
देवजग्धक	1362	देवजग्ध	3579
शल्य	1391	शल्यक	3919
शिखण्ड	1482	शिखण्डक	3977
चकोरक	1492	चकोर	4023
कोयष्टिक	1493	कोयष्टि	4021
टिट्टिभक	1494	टिट्टिभ	3999
अर्भक	1503	अर्भ	1104
शावक	1503	शाव	1104
गृह्यक	1519	गृह्य	4033
ऊर्ध्वजानु	1675	ऊर्ध्वजानुक	1402
खञ्ज	1684	खञ्जक	1400
मण्डलक	1703	मण्डल	1442
वर्चस्क	1757	वर्चस्	1936
माणवक	1884	माणव	2006
आवापक	1887	आवाप	2013
स्थूलशाटक	1918	स्थूलशाट	2038
कूर्पासक	1923	कूर्पासक	2043
चित्रक	1937	चित्र	1981
कृत्रिमधूपक	1946	कृत्रिमधूप	1966

मञ्च	1972	मञ्चक	2064
समुद्गक	1976	समुद्ग	3098
सम्पुटक	1976	सम्पुट	3098
तालवृन्तक	1981	तालवृन्त	2076
राजन्यक	2180	राजन्य	4243
द्वारपाल	2186	द्वारपालक	2183
पद्मक	2294	पद्म	3727
बिन्दुजालक	2294	बिन्दुजाल	3727
अक्षाग्रकीलक	2352	अक्षाग्रकील	2295
कङ्कटक	2372	कङ्कट	2324
पदाति	2376	पदातिक	1537
शरव्य	2434	शरव्यक	2361
लिप्तक	2440	लिप्त	2365
वीरपाण	2479	वीरपाणक	2433
वार्षुषिक	2531	वार्षुषि	2687
कोरदूष	2578	कोरदूषक	3528
प्रसेव	2613	प्रसेवक	2791
कण्डोल	2614	कण्डोलक	3107
पिट	2614	पिट	3107

किलिञ्जक	2615	किलिञ्ज	3105
अश्मन्त	2620	अश्मन्तक	3110
वर्धमानक	2630	वर्धमान	3125
शिग्रु	2638	शिग्रुक	3549
चिपिटक	2674	चिपिट	1255
तर्णक	2723	तर्ण	3818
धौरेय	2736	धौरेयक	3825
शिवक	2762	शिव	3867
कीलक	2762	कील	3867
छगलक	2772	छगल	3870
वैदेहक	2778	वैदेह	2636
वण्टक	2815	वण्ट	4298
दुर्वर्ण	2827	दुर्वर्णक	3175
ताम्रक	2829	ताम्र	3172
शस्त्रक	2830	शस्त्र	3168
हरितालक	2845	हरिताल	3202
तूल	2852	तूलक	3409
रथकार	2878	रथकारक	2754
लोहकारक	2891	लोहकार	2815

पाणिवाद	2908	पाणिवादक	2833
गोप्यक	2918	गोप्य	1161
चेटक	2918	चेट	1161
उष्ण	2921	उष्णक	1210
मृगदंशक	2928	मृगदंश	3877
शुनक	2928	शुन	3877
भीरु	3068	भीरुक	1173
रिक्तक	3167	रिक्त	4327
क्षुल्लक	3174	क्षुल्ल	4281
विप्रकृष्टक	3186	विप्रकृष्ट	4338

Annexure 3

Analysis of concordance between AMAR, ABCH, ACPH, ACSJ and AVCH

1. सम्भव

ABCH enumerates only शम्भव, but ACSJ adds सम्भव also. It is a case in which the tālavya ūṣman consonant is replaced by a mūrdhanya one.

2. मुनि

ABCH mentions मुनिसुव्रत as the name of a tīrthaṅkara. ACSJ adds मुनि. HSDS explains it like this - “मुनिसुव्रतैकदेशो वा मुनिः”. Just like भीम is used instead of भीमसेन, ekadeśa justifies this usage. There are other such cases too e.g. बलराम-राम.

3. नेमिन्

ACSJ adds this word over and above नेमि. HSDS justifies it thus “नेमः-संयमरूपा मर्यादा सः अस्यास्तीति नेमी । नीयते श्रेयोऽनेनेति वा नेमी । "बहुलम्" [५।१।२] इति वचनाद् मिन् ।”

4. मण्डितपुत्र

ABCH entry reads “मण्डितमौर्यपुत्रौ” as one of the 11 gaṇādhīpas. In ACVH, Hemacandra himself splits this compound in two ways - मण्डित and मौर्यपुत्र, as well as मण्डितमौर्ययोः पुत्रौ. Thus, another entry मण्डितपुत्र is added by ACVH and ACSJ.

5. मरुदेवी

ABCH entry has “मरुदेवा”, but ACSJ also attests another form मरुदेवी by mentioning “नवा शोणादेः” [२।४।३९] इति विकल्पेन डीः ।’.

6. अजिता

ACSJ justifies it as ekadeśa of अजितबला.

7. सुतारा

सुतारका is also known by the name of सुतारा. ACVH and ACSJ both attest to this form. HSDS gives its etymology as “शोभना तारा अस्याः सुतारा । सुतराम् तारयति भक्तान् विघ्नादिभ्य इति वा ।”.

8. भद्र

Synonym of भद्रकृत् as per ACSJ.

9. भन्द्र

Synonym of भद्र as per ACSJ. “भन्देर्वा [उ.३९९] इति रः नलुक् च विकल्पेन, नलोपाभावे भन्द्रम्”.

10. श्रवण

श्रमण is also known as श्रवण as per ACSJ. Etymology is given as ‘श्रूयते सर्वलोकानामत्यन्तमान्यत्वेन इति श्रवणः’.

11. स्वर

This indeclinable is enumerated by AMAR along with the other words denoting svarga, whereas ABCH has skipped it, saying ‘स्वरव्ययेषु वक्ष्यते’ in ACVP.

12. त्रिदिव

AMAR treats it as masculine, whereas ACVH and AVRDR treat it as masculine as well as neuter. Nāmacandrikā commentary on Amarakośa by Viśṇudattaśarman (ANVS) attests its existence as

neuter by saying ' "त्रिदिवं तु खे । स्वर्गे च त्रिदिवा नद्याम्" इति दर्शनात् क्लीबेऽपि ।'. Amarapadapārijāta of Mallinātha (APPM) also cites Vaijayantī of Yādavaprakāśa (VJYP) to justify both masculine and neuter gender of the word. "वैजयन्त्यां (पृ. ३, श्लो. २) तु त्रिदिवशब्दः पुंनपुंसक इति कथितः ।".

13. सुरालय, वासवावास

AMAR has त्रिदशालय. ABCH adds सुरालय and ACPH adds वासवावास. Here, आलय and आवास are synonyms, and त्रिदश and सुर are synonyms. वासव is इन्द्र, i.e. king of gods. Therefore, this kind of alternate headwords can be generated by replacing त्रिदश with its synonyms and आलय with its synonyms. The only restriction that applies here is that the usage should be attested in literature. ABCH has very categorically dealt with such an extension of vocabulary in his paribhāṣa section.

“वाध्यादिषु पदे पूर्वे वडवाग्निषूत्तरे ॥ १७ ॥

द्वयेऽपि भूभृदाद्येषु पर्यायपरिवर्तनम् ।

एवं परावृत्तिसहा योगात् स्युरिति यौगिकाः ॥ १८ ॥

मिश्राः पुनः परावृत्त्यसहा गीर्वाणसन्निभाः ।

प्रवक्ष्यन्तेऽत्र"

Thus, there are three types of पर्यायपरिवर्तनम् which can happen. Words like वार्धि allow replacement with synonyms in pūrvapada only e.g. वार्धि, जलधि, तोयधि etc. Words like वडावाग्नि allow replacement in uttarapada only e.g. वडवाग्नि, वडवानल, वडवावह्नि etc. Words like भूभृत् allow replacement in both pūrvapada and uttarapada e.g. भूभृत्, उर्वीभृत्, भूधर, उर्वीधर etc. These words are known as यौगिक words.

Unlike यौगिक words, मिश्र words do not allow for any replacement at all e.g. गीर्वाण. We cannot replace either pūrvapada like वाग्बाण, uttarapada like गीःशर, or both like वाक्शर. None of these words would have the meaning of 'god'. Therefore, the users of the lexica have to be very cautious when they extend the word list by replacing it with synonyms of the words that have been provided in the lexica.

In the present case, both pūrvapada and uttarapada are replaceable. There are multiple such cases in which one dictionary has used one word out of the synonym set and the second dictionary has used another. An example in which uttarapada is replaced is चतुरानन – चतुर्मुख. In the following, the pūrvapada is replaced, कमलासन – सरोरुहासन. In the following, both pūrvapada and uttarapada are replaced, अब्जयोनि - कमलोद्भव.

14. महाराजिक, माहाराजिक

AMAR has given the word as महाराजिक, and ABCH has given it as माहाराजिक. Kṣīrasvāmin mentions in his Amarakośodghāṭanam commentary (AKKS) that both are valid forms and gives the rules for deriving both “महाराजशब्दोऽस्त्येषामिति । सास्य देवता (सू०) इति माहाराजिका इत्येके, महाराजप्रोष्ठपदाद्वञ् (सू०) ।”.

15. Difference in form due to compound v/s taddhita forms

There are multiple such cases where one of the dictionaries has given a compound form, and one has given taddhita form, e.g. दितिज – दैत्य, शुद्धोदनसुत – शौद्धोदनि, etc.

16. Additional words in ABCH

सौनन्द and संवर्तक are said to be मुसल and हल of बलराम. This information is absent in AMAR. Names of enemies of कामदेव are enumerated in ABCH. Similarly, there are other such words too.

17. Additional words in AMAR

Names of flowers which make up the arrows of Kāmadeva, like अरविन्द, अशोक, चूत, नवमल्लिका, नीलोत्पल and the effect of Kāma on human beings like उन्मादन, तापन, शोषण, स्तम्भन, सम्मोहन are mentioned in AMAR, but absent from AMAR. Similarly, there are other such words too.

18. सुदर्शन

AMAR treats sudarśana as a neuter word - चक्रं सुदर्शनम् । (१.२९). ABCH treats it as a masculine word - चक्रं सुदर्शनः । (२.२२२). Abhidhānacintāmaṇisvopajñavṛtti (ACVH) holds that it is masculine despite the authority of Vācaspati (VCSP). वाचस्पतिस्तु- "चक्रं सुदर्शनोऽस्त्रियाम्" इत्याह । (२.२२२). Abhidhānacintāmaṇivyutpatttiratnākaraṭīkā (AVRD) holds that this word is both a masculine and neuter word - "सुदर्शनोऽस्त्रियां चक्रे" इति नामनिधानात् पुंनपुंसकलिङ्गोऽपि । वाचस्पतिरपि "चक्रं सुदर्शनोऽस्त्रियाम्" इत्याह । (२.२२२). Therefore, there are multiple authorities who treat this word as a masculine and neuter word.

19. कौमोदकी - कौपोदकी

The club of कुमोदक i.e. विष्णु is called कौमोदकी. ACSJ mentions that it is also known as कौपोदकी - "कूपोदके भवा कौपोदकी ।".

20. अजगव, अजगाव, अजकाव, आजगव

AMAR, ABCH and ACSJ put together have brought out these four usages of the word denoting Śiva's bow. Śabdabhedaparakāśa of Maheśvara

(SBPM) summarises the four valid alternatives thus “अजगवाऽजगावे च भवेदाजगवं तथा । पिनाकेऽजकावम् चाथ सैहिकेये तमस्तमाः ॥ ११८ ॥”.

21. पिनाक

AMAR has it as masculine, whereas ABCH treats it as masculine and neuter. AVRД attests occurrence in both masculine and neuter genders. “यद्वा पातीति पिनाकः” and “यद्वा पिनष्ट्यरीन् पिनाकम्”.

22. पारिषद, पार्षद

AMAR has पारिषद, whereas ABCH has पार्षद. Hemaandra also acknowledges the existence of the word पारिषद in ACVH by saying “पारिषदा इत्यपि”.

23. वशित्व, वशिता

This kind of difference is caused by the application of different suffixes conveying the same meaning.

24. भृङ्गिन्, रिति, भृङ्गीरिति, भृङ्गीरीति

AMAR has the following entry – “शृङ्गी भृङ्गी रितिस्तुण्डी” whereas ABCH has “भृङ्गी भृङ्गीरितिभृङ्गीरीतिर्नाड्यस्थिविग्रहः”. It requires further examination to check whether the editor of AMAR wrongly put space between भृङ्गी and रिति.

25. शुनासीर, सुनासीर, शुनाशीर

AMAR has सुनासीर, ABCH has शुनासीर. AVRД also has शुनाशीर. Amaraviveka of Maheśvara (AVMH) summarises the occurrences as 'शुनासीरः "द्वितालव्यः द्विदन्त्यश्च”.

26. व्योमयान

AMAR treats it as neuter, whereas HSDS treats it as both masculine and neuter.

27. पीयूष, पेयूष

ABCH has पीयूष, and ACSJ has पेयूष. Hemacandra himself also admits पेयूष in ACVH. HSDS treats it as having both masculine and neuter genders.

28. पारिजातक, पारिजात

This is a common occurrence. One dictionary has enumerated an entry with 'ka' suffixed at the end, whereas the other may have enumerated it without the 'ka' suffix. This phenomenon is already analysed in Annexure 2.

29. नासत्यौ, दस्रौ, नासत्यदस्रौ

Both AMAR and ABCH attest नासत्यौ and दस्रौ. ACPH adds नासत्यदस्रौ to the list.

30. बर्हिस्, शुष्मन्, बर्हिःशुष्मन्

Some commentaries on AMAR interpret the word 'बर्हिःशुष्मा' as बर्हिस् and शुष्मन्, whereas some other commentaries allow बर्हिःशुष्मन् also as a separate entry. See Vyākhyāsudhā commentary (AVSB) for example which says, "शुक्रो वैश्वानरो वह्निर्बर्हिःशुष्मा तनूनपात्" इति शब्दार्णवात् "बर्हिःशुष्मा" इति व्यस्तं समस्तं नाम इति कश्चित्. ABCH takes it as बर्हिःशुष्मन्.

31. अर्चिस्

According to AMAR, this word is feminine in nature, whereas ACVH treats it as feminine and neuter. "अर्च्यतेऽर्चिः स्त्रीक्लीबलिङ्गः". AMAR itself also mentions it as feminine and neuter in nānāarthavarga. "ज्वालाभासौ न पुंस्यर्चिस्" (नानार्थवर्गः २३१). Hence, this word can be conclusively treated as taking both feminine and neuter genders.

32. झञ्झावात, झञ्झा

झञ्झा is a special category of वात. AMAR prefers to call it झञ्झावात, whereas ABCH prefers to call it only झञ्झा.

33. झम्पा, झम्प

Both feminine and masculine word mean the same thing. ACVH mentions “स्त्रीलिङ्गोऽयं 'झम्पः सम्पातपाटवम्' इति पुंस्यन्ये”.

34. पद्म

Out of 9 treasures, पद्म is treated as masculine in both AMAR and ABCH, but according to Vācaspati, it is used both in masculine and neuter gender. This view of Vācaspati is cited in ACVH as “वाचस्पतिस्तु- पद्मोऽस्त्रियाम्, इत्याह ॥”. Rasālā commentary on Amarakośa by Śaktidharaśarman (ARSS) quotes the full quotation thus “पद्मोऽस्त्रियां महापद्मः शङ्खो मकरकच्छपौ । मुकुन्दकुन्दनीलाश्च खर्वश्च निधयो नवेति (शब्दार्णवः)”. APPM mentions the same verse and ascribes it to Hārāvalī.

35. खर्व, चर्वस्

ABCH prefers चर्वस् over खर्व. Even different commentaries cite हारावली to have both खर्व and चर्वस्. Tikāsarvasva commentary on Amarakośa (ATSS) mentions the following “पद्मोऽस्त्रियां महापद्मः शङ्खो मकरकच्छपौ । मुकुन्दकुन्दनीलाश्च चर्वश्च निधयो नव ॥ इति हारावली”. Whereas APPM mentions the following - “पद्मोऽस्त्रियां महापद्मः शङ्खो मकरकच्छपौ । मुकुन्दकुन्दनीलाश्च खर्वश्च निधयो नव ॥ इति तल्लिङ्गे तत्संख्यायां च हारावलिः ॥”.

36. अन्तरिक्ष, अन्तरीक्ष

AMAR has अन्तरिक्ष, whereas ACSJ also adds अन्तरीक्ष to the list.

37. अवाची, अपाची

ABCH mentions दक्षिणाऽपाची, and ACVH adds अवाची to the synset by saying “जपादित्वाद् वत्वे अवाची इत्यपि”. जपादि is a gaṇa in Siddhahemaśabdānuśāsanam which allows replacement of ‘p’ with ‘v’ by the rule जपादीनां पो वः (२.३.१०५) like जपा-जवा, पारावत-पारापत. ACVH has 14 such occurrences where alternate words are explained with the help of this rule. They are अपाची-अवाची, लिपि-लिवि, पोटा-वोटा, करपाल-करवाल, कपाट-कवाट, पारापार-पारावार, अकूपार-अकूवार, प्रतति-व्रतति, जपा-जवा, शेपाल-शेवाल, अपरा-अवरा, तम्पा-तंवा, पारापत-पारावत and याप्य-याव्य.

38. अवाचीन, अपाचीन

AMAR prefers अवाचीन, whereas ABCH prefers अपाचीन.

39. उत्तरा-अपाचीतरा, अपाची-उत्तरेतरा

ACPH has added names of directions by mentioning the diagonally opposite direction and appending इतरा at the end. उत्तरा is termed as अपाचीतरा and अपाची is termed as उत्तरेतरा. This is a way to increase the lexicon by taking advantage of the antonymic relationships.

40. Directions and their grahas

AMAR mentions that "रविः शुक्रो महीसूनुः स्वर्भानुर्भानुजो विधुः । बुधो बृहस्पतिश्चेति दिशां चैव तथा ग्रहाः ।" i.e. starting from the East, South-East and so on counterclockwise, these grahas preside over those quarters. This information is missing in ABCH.

41. Consorts of Diggajas

AMAR mentions the 8 consorts of 8 diggajas, whereas ABCH does not provide this information.

42. चमर, चामर

ACPH adds चमर as an alternate to चामर.

43. वर्ष

AMAR treats it as neuter, whereas ACVH treats it as neuter and masculine, based on the authority of Vācaspati “यद् वाचस्पतिः- अथ वृष्टिर्वर्षमस्त्री केचिदिच्छन्ति वर्षणम् ।”.

44. करक, करका

AMAR has करका as feminine, whereas ABCH has it as a word that can take all three genders. ANVS cites the following: “करकस्तु पुमान्पक्षिविशेषे दाडिमेषु च । द्वयोर्मेषोपले न स्त्री करङ्गे च कमण्डलौ ॥”। Thus, in the meaning of मेषोपल, this word seems to take only masculine and feminine gender, and not neuter, as claimed by Hemacandra. In neuter gender, this word can be used in sense of करङ्ग or कमण्डलु.

45. अपिधान, पिधान

Both these words are attested in AMAR. ATSS explains this phenomenon with the quote “वष्टि वागुरिरल्लोपमवाप्योरुपसर्गयोः”.

46. हिम, नीहार, तुषार

AMAR treats हिम as neuter and the rest two as masculine, whereas ABCH treats all three of them as taking both masculine and neuter gender, “एते त्रयः पुंस्त्रीबलिङ्गाः”.

47. तारा

AMAR treats it as feminine, whereas ABCH treats it as masculine as well as feminine. Vyākhyāsudhā commentary on Amarakośa by Bhānujīdikṣita (AVSB) clarifies the situation thus “तारयतेः पचाद्यचि (३.१.१३४) । तारोऽपि 'नक्षत्रे नेत्रमध्ये च तारा स्यात्तार इत्यपि' इति व्याडिः । ('तारो

वानरभिन्मुक्ताविशुद्ध्योः शुद्धमौक्तिके । ना, नक्षत्रेऽसिमध्ये च न ना, रूप्ये नपुंसकम् । स्त्री बुद्धदेवताभेदे
वालिगीर्पतिभार्ययोः । त्रिलिङ्गोऽत्युच्चशब्दे च')

48. तारका

AMAR treats it as feminine, whereas ACVH treats it as having all three genders. AVSB clarifies the situation thus “अपिशब्दात्तारकापि वा स्त्रियाम् । स्त्रीत्वात् क्लीबमुदाहर्यति । 'नक्षत्रे चाक्षिमध्ये च तारकं तारकापि च' इति शाश्वतः । ('तारको दैत्यभित्कर्णधारयोर्न द्वयोर्दृशि । कनीनिकायामृक्षे च न पुमांस्त्रातरि त्रिषु) ॥”.

49. मृगशिरस्

AMAR treats it as neuter, whereas ACVH treats it as having both masculine and neuter gender, “पुंक्लीबलिङ्गोऽयम्, यद्वाचस्पतिः- सौम्यं मृगशिरोऽस्त्रियाम् ।”. AVSB has totally different view, and claim that it has both feminine and neuter gender on the authority of Vopālita “रूपभेदात्क्लीबम् । 'सौम्या मृगशिरः स्युर्मृगशिराः' इति बोपालितः स्त्रीत्वमप्याह ॥”.

50. भास्

AMAR treats it as feminine, whereas ACVH treats it as both masculine and feminine. ATSS cites a quotation from Veṇīsaṃhāra to justify that this word also takes up masculine gender. ' “जृम्भारम्भप्रविततदलोपान्तजालप्रविष्टैर्भाभिर्भानोर्नृपतय इव स्पृश्यमाना विबुद्धाः ।” इति वेणीसंहारे भाशशब्दः पुंलिङ्गोऽपि ।’

51. Jain measurement of time

Entries like अवसर्पिणी, उत्सर्पिणी, अर, एकान्तसुषमा, सुषमा, सुषमदुःषमा, दुःषमसुषमा, दुःषमा, एकान्तदुःषमा are present only in ABCH and absent from AMAR.

52. दिन

AMAR has it as neuter. ANVS emphatically declares that it is neuter only. “दिनं क्लीबमेव”. ACVH treats it as both masculine and neuter genders.

53. साय

AMAR treats साय as masculine, whereas ABCH treats it as neuter. ACVH justifies it thus, ‘स्यति दिनं सायं, क्लीबलिङ्गः ॥ गौडस्तु-उत्सूरश्च विकालश्च सायश्च सबलिश्च स इति पुंस्याह ।’.

54. गणरात्र

AMAR has it as neuter, whereas ABCH has it as both masculine and neuter. It is a quite hotly debated issue in the commentarial literature of both AMAR and ABCH.

55. तपस्, शुक्र

AMAR – m., ACVH – mn. VJYP mentions ‘शुक्रस्सिते कवौ चन्द्रे मासभेदेऽनले द्विजे । शुक्रं मेध्येऽप्सु पुण्येऽर्थे रुक्मेऽक्षिरुजि रेतसि ॥’. So, it seems that the words indeed take two genders, but in different senses. Therefore, in the present synset, it takes only masculine gender.

56. प्रकाण्ड

AMAR - n., AVR D – mn. Amarapadapārijāta by Liṅgāyasūrin (APPL) mentions ‘ “प्रकाण्डोऽस्त्री पुमानुद्धः” इति त्रिकाण्डशेषः ।’.

57. द्वापर

AMAR - m., ACVH - mn. AVSB mentions the following without providing the source, which also attests the neuter gender ‘द्वापरं संशये युगे’.

58. द्वादशाङ्ग

ABCH provides the names of Jain āgamas. AMAR does not.

59. Tenets of Jainism

ABCH provides them, AMAR does not. See उत्पाद, अग्रायणीय etc.

60. हरिण, पाण्डुर, पाण्डु

AMAR treats them as a separate synset, whereas ACVH subsumes them under the श्वेत synset.

61. मेचक

AMAR - m., ACVH - mn. ACVH also quotes दुर्ग to say that the word means a slightly different shade according to some lexicographers ‘ “मेचकः शितिकण्ठाभः” इति दुर्गः’. AVSB also mentions a different author “शब्दार्णवे तु ‘मेचकः कृष्णनीलः स्यादतसीपुष्पसन्निभः’ इत्युक्तम् ।”.

62. उपन्यास-वाङ्मुख, उपोद्धात-उदाहार

AMAR treats them as two separate synsets, whereas ACVH has taken them to be a single synset.

63. समज्ञा, समाज्ञा

AMAR enlists समज्ञा, whereas ABCH enlists समाज्ञा. AVSB also attests “ ‘समाज्ञा’ इति समाङ्पूर्वाज्ज्ञः”.

64. पटह

AMAR – mn., ACVH – m. ACVH also cites Vācaspati’s view that it takes both the genders ‘वाचस्पतिस्तु-“पटहस्त्वानकोऽस्त्रियाम्” इत्याह ।’.

65. Types of नाट्य

व्यायोग, समवकार, etc., are mentioned in ABCH, and not shown in AMAR.

66. शृङ्गार

AMAR – m., ACVH – mn. AVSB also mentions the following 'शृङ्गारो गजमण्डने । सुरते रसभेदे च शृङ्गारं नागसम्भवे । (चूर्णे लवङ्गपुष्पे च)' इति हैमः ।'. Therefore, this word is used in different genders for different intended meanings. In the present case, it takes masculine gender.

67. दर

AMAR – m., ACVH – mn. AVSB mentions “‘दरः स्याद्भयगर्तयोः । दरी तु कन्दरे' इति हैमः”.

68. मान

AMAR – m., ACVH – mn. AVSB mentions “‘मानं प्रमाणे प्रस्थादौ मानश्चित्तोनतौ ग्रहे' इति हैमः”.

69. Gender differences in the words denoting poisons

वत्सनाभः कालकूटो ब्रह्मपुत्रः प्रदीपनः ।

सौराष्ट्रिकः शौलिकेयः काकोलो दारदोऽपि च ॥११९६॥

अहिच्छत्रो मेषशृङ्गः कुष्ठ-चालूक नन्दनाः ।

कैराटको हैमवतो मर्कटः करवीरकः ॥११९७॥

सर्षपो मूलको गौरार्द्रकः सक्तु-कर्दमौ ।

अङ्गोलसारः कालिङ्गः शृङ्गिको मधुसिक्थकः ॥११९८॥

इन्द्रो लाङ्गलिको विस्फुलिङ्ग-पिङ्गल-गौतमाः ।

मुस्तको दालवश्चेति स्थावरा विषजातयः ॥११९९॥

According to ACVH, they also take neuter gender as per the authority of Vācaspati. 'एते सर्वेऽपि पुंक्लीबलिङ्गाः' इति वाचस्पतिः'.

70. Names of hells

AMAR enlists as per Hindu tradition, whereas ABCH does as per Jain tradition.

71. द्वीपोऽस्त्रियामन्तरीपं

There is difference of opinion about the interpretation about gender of these two words. Whether अस्त्रियाम् applies to द्वीप or अन्तरीप or both. Amarapadacandrikā of Rāyamukūṭa (APCR) mentions this clash as follows: “अप्-शब्दस्यादेरीत्त्वेन च साधिते "अस्त्रियामि"ति काकाक्षिवत् द्वाभ्यां सम्बन्धः । "अन्तरीपेणैवेति केचिद्" इति स्वामी ॥".

72. तरणि-तरणी, तरि-तरी

AMAR prefers तरणि, तरि. ACVH prefers तरणी, तरी.

73. शङ्खनख, शङ्खनक

AMAR has शङ्खनख, whereas ABCH has शङ्खनक. ANVS mentions “शङ्खस्य नखा इव (शङ्खनखाः)--- इति तु मुकुटः ॥ शङ्खनकाः-- इत्यन्ये ॥".

74. Coverage about names of insects

ABCH has more coverage of insects like दुर्नामा, पिपीलक, पिपीलीका, ब्राह्मणी, घृतेली etc.

75. अखात

Both AMAR and ABCH mention अखात to be neuter. But ANVS also mentions “अखातो देवखातः स्यात्" इति पुंस्काण्डेऽमरदत्तात् पुंस्यपि ॥".

76. आलवाल

AMAR – n., ACVH – mn. AVRĐ cites Abhidhānaratnamālā of Halāyudha (ARMH) for masculine gender thus, “आवाल आलवालः स्यादावापः स्थानकं तथा" [हलायुधकोशः २ । १८५ ॥] इति हलायुधः ।

77. शैवाल

AMAR – n., ACVH – mn. APCR cites Vācaspati to attest masculine version of the word ‘ “शैवलश्चैव शैवालः शैवलो जलनीलिका" इति वाचस्पतिः ।'.

78. शैवल

AMAR – m., ACVH – mn. AKKS cites “शेते जले शेवलं शेवालं च, शीडो-
वलज्वलनवालनः (उ०), शैवलं शैवालमिति तु द्रविडाः ।” to attest neuter gender.

79. विष्टप

AMAR – n., ACVH – mn. AVRĐ, ARSS and APCR all cite Amaramālā
to justify the masculine gender ‘ “पिष्टपो विष्टपोऽप्यस्त्री भुवनं तु नपुंसकम्”
इत्यमरमालोक्तेः ।’

80. नीवृत्

AMAR – m., ACVH – f. Hemacandra himself is aware of the
difference of opinion, as he himself notes in ACVH ‘ नियतं वर्तन्तेऽस्यां
नीवृत् स्त्रीलिङ्गः, अमरस्तु पुंस्याह ।’.

81. सेतुरालौ स्त्रियां पुमान् ।

The anvaya of this verse is explained by AVMH as “स्त्रीलिङ्गायामालौ सेतुः
पुमानित्यन्वयः ।”.

82. विपथ, कापथ

AMAR – m., ACVH – n. Hemachandra is aware about his difference
of opinion as he says “पथः सङ्ख्याऽव्ययोत्तरः” ॥ (हैमलिङ्गानु० नपुं० श्लो-८) ॥ इति
क्लीबत्वम् । अमरस्तु – विपथकापथशब्दौ पुंस्याह, गौडश्च यदाह- “व्यध्वो विपथकापथौ” इति ॥’

83. गव्यूति

AMAR – f., ACVH – mf. APCR also mentions ‘गव्यूतिः पुंसि स्त्रियां च ।
“धन्वन्तरसहस्रं तु क्रोशः क्रोशद्वयं पुनः । गव्यूतं स्त्री तु गव्यूतिर्गोरुतं गोमतं च तत् ॥” इति
वाचस्पतिः । “धनुर्हस्तचतुष्टयम्” इत्युक्त्वा “द्वाभ्यां धनुःसहस्राभ्यां गव्यूतिः पुंसि भाषित” इति
शब्दार्णवः ॥’.

84. List of countries

ABCH enlists the countries from कामरूप to कुन्तल and गया to निषधा.

AMAR does not.

85. मन्दिर

AMAR – n., ACVH – fn. APCR mentions “मदेः किरच् (उ० १ । ५२) । 'मन्दिस्म्' । "स्त्रीनपुंसकयो" रित्यगोक्तेः (अरुणदत्तोक्तः ?)- स्त्रियां टाप् (पा० ४ । १ । ४) 'मन्दिरा' च । प्रयोगश्च-“मन्दिरातस्तवारिः” इति सुभूतिः ।". ATSS also cites the following detail - “'मन्दिरं शकटं पीठं क्रोडं रशनं तुम्बं नगरं मन्दिरापुर्यो' इत्यरुणदत्तेन स्त्रीक्लीबयोर्मन्दिरं पठितम् । तथा चास्यैव प्रयोगः- 'व्रजति पथि मनाहं(?) मन्दिरातस्तवारिः' इति ।"

86. वातायन

AMAR – n., ACVH – mn. AVR D mentions ‘वातस्याऽयनं गमनमनेनेति वातायनः, पुंक्लीबलिङ्गः । पुंसि वामनः, क्लीबे माला ।'.

87. शिखर

AMAR – n., ACVH – mn. ATSS mentions ‘अस्त्री शिखरमिति सम्बन्धः'.

88. शृङ्ग

AMAR – n., ACVH – mn. ARSS mentions ‘सानुः शृङ्गः पर्वताग्रभागो'.

89. श्रेणि , श्रेणी

AMAR – f., ACVH – mf. ARSS mentions “'श्रेणिः स्त्रीपुंसयोः पङ्क्तौ समानशिल्पिसंहतौ" (इति मेदिनी)'. ACVH mentions ‘श्रीयते श्रेणिः पुंस्त्रीलिङ्गः, इयां श्रेणी ।'.

90. कुट्मल, कुड्मल

AMAR – कुट्मल, ABCH – कुड्मल. ARSS mentions “कुडति, कुड्यते वा कुड्मलः कुटति कुट्यते वा कुट्मलश्च".

91. पुष्प

AMAR – n., ACVH – mn. ACVH mentions “पुष्प्यति पुष्पं पुंक्लीबलिङ्गः,
वाचस्पतिस्तु ‘पुष्पोऽस्त्री’ इति पुंस्यप्याह ।”.

92. प्रियाल, पियाल

AMAR – प्रियाल, ACVH – पियाल. ACVH also admits प्रियाल too.

93. कर्कन्धू, कर्कन्धु

AMAR – कर्कन्धू, ACVH – कर्कन्धु. ACSJ adds कर्कन्धू as a supplement.

HSDS mentions “कर्को लोहितः अन्धुः, कर्कस्य अन्धुरिव वा कर्कन्धूः ।

“उतोऽप्राणिनश्चायुरज्वादिभ्य ऊङ्” [२।४।७३] इति ऊङि दीर्घः । स्त्रीलिङ्गः । पक्षे कर्कन्धुः,
कुवली । पुंस्त्रीलिङ्गोऽयम् ।”

94. इङ्गदी

AMAR – f., ACVH – mfn. ATSS mentions ‘पुंसि इङ्गदः ।’. So, masculine
gender is taken care of, but the neuter gender remains unattested
in commentarial literature other than ACVH.

95. मलयु, मलयू, मलपू

ACVH has मलयु, but AVRД has मलयू. Various commentaries of AMAR
have मलयु, मलयू or मलपू. ATSS has मलपू (“मलं पुनातीति”). AKKS seems to
have both मलपू and मलयू (“मलं श्वित्रं पावयति मलपूः श्वित्रभेषजम् , (मलयूः) यु
जुगुप्सायाम् ।”).

96. तमाल

AMAR – m., ACVH – mn. AKSS mentions ‘तमालं वृक्षभेदः’.

97. निर्गुण्डी, निर्गुण्ठी, निर्गुण्टी

AMAR has निर्गुण्डी. ACVH has निर्गुण्डी and निर्गुण्ठी. HSDS adds निर्गुण्टी too.

98. कुरण्टक, कुरुण्टक, कुरुण्ट, कुरुण्डक, कुरुण्डक

ACVH and HSDS put together attests these many variations of the same word.

99. करीर

AMAR - m., ACVH - mn. AMAR himself mentions that "वंशाङ्कुरे करीरोऽस्त्री तरुभेदे घटे च ना ।". Therefore, the word depicting the tree is masculine in nature.

100. मातुलुङ्ग, मातुलिङ्ग

HSDS mentions these two as alternatives.

101. निदिग्धा, निर्दिग्धा

AMAR mentions निदिग्धा, ACVH mentions निर्दिग्धा. But it seems to be an editorial lapse, as AVRĐ mentions it as निदिग्धा and writes in footnote “ 'निर्दिग्धा' इति३, मुद्रितमूले स्वोपज्ञटीकायां च 'निर्दिग्धिका' इति दृश्यते ॥”

102. Names of trees

AMAR enlists many trees which ABCH does not. It may be because Hemacandra has authored a separate work Nighaṇṭuśeṣa dealing with the subject.

103. बल्वजाः, वल्वजाः

AMAR – बल्वजाः, ACVH – वल्वजाः.

104. रौहिष

AMAR – n., ACVH – mn. AVMH cites Viśvaprakāśa - 'रौहिषः "रोहिषः" । "रोहिषं कत्तृणे ज्ञेयं रोहिषो हरिणान्तरे" इति विश्वप्रकाशः ।'. So, in the present usage neuter seems more apt.

Annexure 4

Additional information provided by lexicographers on specific words

1. ABCH – ऋषिर्मुनिः (७६)

ACVH - ऋषति जानाति तत्त्वं ऋषिः ... दर्शनाद्वा ऋषिः ।

Bhaṭṭatota – यद्भट्टतोतः- ऋषिश्च किल दर्शनादिति ।

2. ABCH – भुविस्तविषताविषौ (८७)

ACVH – भुविः स्त्रीलिङ्गः

Vācaspati (VCSP) – यद्वाचस्पतिः - “त्रिपिष्टपं देवलोको भुविः स्त्री द्योदिवौ स्त्रियौ” ।

3. ABCH – ग्रहपुषः (९५)

ACVH - आदित्यो हि सुषुम्णाभिर्नाडिभिर्ग्रहान् पुष्णाति ।

Vyādi (VYDI) – यद् व्याडिः- “सुषुम्णाद्याश्च नाड्योऽस्य पुष्णन्ति सततं ग्रहान् ॥”

4. ABCH – पूषा (९५)

ACVH - पूषति वर्द्धते ऋतुभेदाद्रश्मिभिरिति “उक्षितक्षि-” ॥ (उणा-९००) ॥ इत्यादिना अनि पूषा ।

VYDI - यद् व्याडिः-

"ऋतुभेदात् पुनस्तस्यातिरिच्यन्तेऽपि रश्मयः ।

शतानि द्वादश मधौ त्रयोदशैव माधवे ॥१॥

चतुर्दश पुनर्ज्येष्ठे नभोनभस्ययोस्तथा ।

पञ्चदशैव त्वाषाढे षोडशैव तथाश्विने ॥२॥

कार्तिकके त्वेकादश शतान्येवं तपस्यपि ।

मार्गे तु दश साद्धानि शतान्येवं च फाल्गुने ॥३॥

पौष एव परं मासि सहस्रं किरणा रवेः ॥२१॥"

5. ABCH - रोचिरुस्ररुचिशोचिरंशुगो ज्योतिरर्चिरुपधृत्यभीशवः । (९९)

Gauḍa (GAUD) - अभीषुरिति गौडः ।

6. ABCH - अष्टादश माठराद्याः सवितुः पारिपार्श्विकाः ॥१०३॥

VYDI - यद् व्याडिः--

“सुरावृतस्त्वसावष्टादशभिः पारिपार्श्विकैः ।

इन्द्रादयश्च ते देवाः सर्वे नामान्तरैः स्थिताः ॥१॥

तत्र शक्रो वामपार्श्वे दण्डाख्यो दण्डनायकः ।

वह्निस्तु दक्षिणे पार्श्वे पिङ्गलो वामनश्च सः ॥२॥

पूर्वद्वारे गुहहरौ राजश्रोथौ क्रमेण तौ ।

द्वितीये तु धर्माधर्मौ तौ खरद्वारिकौ क्रमात् ॥३॥

तृतीये तु यमताक्षर्यौ क्रमात् कल्माषपक्षिणौ ।

चित्रगुप्तश्च कालश्च रवेर्दक्षिणपार्श्वगौ ॥४॥

प्रथमो जातृकाराख्यो माठराख्यो द्वितीयकः ।

पश्चिमेनाब्धिवरुणौ तौ प्राप्नुयात् कुतापकौ ॥५॥

उदीच्यां यक्षहेरम्बौ तौ च पिङ्गगजौ क्रमात् ।

प्राच्यां तु रुद्ररेवन्तौ तौ दण्डिपुरुषो क्रमात् ॥६॥

आश्विनौ तु पार्श्वद्वये तौ द्वावेव किशोरकौ ।

इत्यष्टादशको वर्गश्चण्डांशोः पारिपार्श्विकः” ॥७॥

7. ABCH - चन्द्रमाः कुमुदबान्धवो दशश्वेतवाज्यमृतसूस्तिथिप्रणीः । (104)

VYDI - यदाह व्याडिः-

“अश्वास्तु दश चन्द्रस्य यजुश्चन्द्रमना वृषः ।

सप्तधातुर्हयो वाजी हंसो व्योममृगो नरः ॥१॥

- अर्वा चाथ चन्द्रमनः स्थानेऽस्ति त्रिधनाः क्वचित् ।
सप्तधातोः पुनः स्थाने सहरुण्योऽस्ति कुत्रचित्” ॥२॥
8. ABCH - चन्द्रिका चन्द्रगोलिका ॥१०६॥
VCSP - चन्द्रिमेत्यपि वाचस्पतिः ।
9. ABCH - दाक्षायण्यः सर्वाः शशिप्रियाः ॥११५॥
Amaratīkā (AMTK) - दक्षस्यापत्यानि इति अजन्तादायनणि दाक्षायण्य इति अमरटीका, तत्र च स्त्रीणां युवसंज्ञाया अभावादायनण् चिन्त्यः ॥. This quotation in ACVH mentions an Amaratīkā, but does not specify the commentator.
10. ABCH - शुक्रो मघाभवः काव्य उशना भार्गवः कविः ॥११९॥
Vāmanapurāṇa (VMNP) - रुद्रशुक्रद्वारेण निर्यातत्वाद्वा यद्-वामनपुराणे-
इत्येवमुक्त्वा भगवान्मुमोच
शिशनेन शुक्रं स च निर्जगाम ॥१॥
11. ABCH - पुष्पदन्तौ पुष्पवन्तावेकोक्त्या शशिभास्करौ ॥१२४॥
Viśvadatta (VSDT) - यद् विश्वदत्तः-रविशशिनौ पुष्पदन्ताख्याविति ॥
12. ABCH - क्षणैः षड्भिस्तु नाडिका । (१३६)
ACVH - नाल्येव नालिका; डलयोरैक्ये नाडिका ।
Nandin (NAND) – नडण् अवस्कन्दने, इत्यस्य नाडीति नन्दी ।
13. ABCH - दिनावसानमुत्सूरो विकालसबली अपि । सायं (१४०)
ACVH - स्यति दिनं सायं, क्लीबलिङ्गः ।
GAUD - गौडस्तु - उत्सूरश्च विकालश्च सायश्च सबलिश्च स इति पुंस्याह ।
14. ABCH - श्राद्धकालस्तु कुतपोऽष्टमो भागो दिनस्य यः । (१४१)
Smṛti (SMRT) - यत् स्मृतिः-
दिवसस्याष्टमे भागे मन्दीभवति भास्करे ।
स कालः कुतपो यत्र पितृणां दत्तमक्षयम् ॥

This seems to be referring to Mahābhārata verse 1.87.31.

15. ABCH - त्रियामा यामिनी भौती तमी तमा विभावरी । (१४२)

ACVH - ताम्यन्ति चक्रवाका अस्यां, गौरादित्वात् इयां तमी ॥१०॥ तमेति तु अमरटीकाभिप्रायेण ॥

16. ABCH - 'तमिस्रा दर्शयामिनी । ज्यौत्स्नी तु पूर्णिमारात्रिः'

ACVH mentions that according to AMAR, the meanings are not strictly the nights of अमावास्या or पूर्णिमा। 'तमिस्रा तामसी रात्रीरिति अमरः ।' and 'ज्यौत्स्नी चन्द्रिकयाऽन्वितेति अमरः ।'.

17. ABCH - निशीथस्त्वर्धरात्रो महानिशा । (१४५)

Kātya (KATY) - नियतं शेरतेऽस्मिन्निशीथः । “न्युद्भ्यां शीडः” ॥ (उणा-२२८) ॥ इति थः कित् ।
अत एव कात्यो निशीथं सुप्तजनमाह ॥

18. ABCH - 'तमिस्रं तिमिरं तमः ॥१४५॥ ध्वान्तं भूच्छायान्धकारं तमसं समवान्धतः । '

AMAR specifies that these three entities refer to specific type of darkness. ACVH duly notes this phenomenon saying 'अमरस्तु-ध्वान्ते गाढेऽन्धतमसं, क्षीणेऽवतमसं तमः । विष्वक् संतमसमिति विशेषमाह ।'.

19. ABCH - तुल्यनक्तंदिने काले विषुवद् विषुवं च तत् ॥ १४६ ॥

ACVH - विषुवत्, पुंक्लीबलिङ्गः ।

VJYP - यद् वैजयन्ती- पुरी तन्महिमा हेम विषुवत् कर्म लोम दोः नृषण्डलिङ्गाः ।

20. ABCH - दैवे युगसहस्रे द्वे ब्राह्मं (१६०)

Manusmṛti (MANU)- यत् मनुः----

चत्वार्यहुः सहस्राणि वर्षाणां तु कृतं युगम् ।

तस्य तावच्छती संध्या संध्यांशश्च तथाविधः ॥१॥

इतरेषु ससंध्येषु ससंध्यांशेषु च त्रिषु ।

एकापायेन वर्तन्ते सहस्राणि शतानि च ॥२॥

यदेतत् परिसंख्यातमादावेव चतुर्युगम् ।

एतद् द्वादशसाहस्रं देवानां युगमुच्यते ॥३॥
दैविकानां युगानां तु सहस्रं परिसंख्यया ।
ब्राह्ममेकमहर्जयं तावती रात्रिरुच्यते ॥४॥

मानुषमानेन ब्राह्ममहोरात्रं यथा-
मानुषाणि सहस्राणां शतान्याहुश्चतुर्दश ।
चत्वारिंशत् सहस्राश्च वर्षाणां तत् कृतं युगम् ॥१॥
लक्षाणि चतुर्हीनानि सहस्रा द्विगुणास्ततः ।
इति त्रेतायुगे कालगणना मानुषी कृता ॥२॥
नियुतानि पुनः सप्त सहस्राणि च विंशतिः ।
द्वापरं युगमन्ये तु लक्षाण्यष्टौ मृषैव तत् ॥३॥
त्रीणि लक्षाणि षष्टिश्च सहस्राणि कलौ युगे ।
चतुर्युगसहस्राणि द्वादश ब्रह्मणो दिनम् ॥४॥
इयं च मानुषेणैव मानेन गणना कृता ।

21. ABCH - कल्पो युगान्तः कल्पान्तः संहारः प्रलयः क्षयः । (१६१)

ACVH - कल्पते जगदस्मिन् कल्पः ॥१॥ कल्पस्यान्तोऽवधिः कल्पान्तः, अत्रैकः कल्पः क्षयार्थोऽन्यो
ब्राह्मदिनार्थः ।

Anekārthasamuccaya of Śāśvata (ANSK) - यत् शाश्वतः---कल्पः शास्त्रे विधौ न्याये
संवर्त्ते ब्रह्मणो दिने ॥

22. ABCH – दुर्दिनं मेघजं तमः । (१६५)

Bhāguri (BHAG) – यद् भागुरिः-दुर्दिनं ह्यन्धकारोऽब्दैः ।

23. ABCH – वृष्ट्यां वर्षाणिवर्षे (१६६)

ACVH - वर्षं पुंक्लीबलिङ्गः

VCSP - यद् वाचस्पतिः-अथ वृष्टिर्वर्षमस्त्री केचिदिच्छन्ति वर्षणम् ॥

24. ABCH - ऐरावतः पुण्डरीको वामनः कुमुदोऽञ्जनः । पुष्पदन्तः सार्वभौमः सुप्रतीकश्च दिग्गजाः ॥

१७०॥

ACVH - दिक्क्रमेणेति प्रस्तावाद् लभ्यते ।

BHAG - 'ऐरावतः पुण्डरीकः कुमुदाञ्जनवामनाः' इति भागुरिः क्रममाह ।

Mālā (MALA) - माला तु-ऐरावतः सुप्रतीक इति ।

25. ABCH - प्राचीनबर्हिः (१७१)

ACVH - प्राचीनमुखा बर्हिषो दर्भा अस्य प्राचीनबर्हिः ।

Viṣṇupurāṇa (VSPR) - यद् विष्णुपुराणे -

प्राचीनाग्राः कुशास्तस्य पृथिव्यां विश्रुता मुनेः ।

प्राचीनबर्हिरभवत् ख्यातो भुवि महाबलः ॥ (VSPR verse 1.14.4)

26. ABCH - पिनाकं स्यादाजगवमजकावं च तद्धनुः । (२०१)

AMAR - अमरस्तु "पिनाकोऽजगवं धनुः" इत्याह ।

Prācya - प्राच्यास्तु- "अजगावं धनुः प्रोक्तम्" इत्याहुः ॥

27. ABCH - देवानांप्रियजाल्मौ च (३५३)

AMAR - "जाल्मोऽसमीक्ष्यकारी स्यात्" इत्यमरः ।

28. ABCH - अथ कातरः । दरितश्चकितो भीतो भीरुभीरुकभीलुकाः ॥३६५॥

AMAR - "अधीरे कातरः" इत्यमरः ।

29. ABCH -

माया तु शठता शाठ्यं कुसृतिर्निकृतिश्च सा ॥३७७॥

कपटं कैतवं दम्भः कूटं छद्मोपधिश्छलम् ।

व्यपदेशो मिषं लक्षं निभं व्याजः (३७८)

Unknown - मायादयः कपटादयश्च एकार्था इत्येके ।

AMAR - कपटादयो व्यपदेशादयश्च भिन्नार्था इत्यमरः

30. ABCH -

अथार्थनैषणा ।

अर्दना प्रणयो याञ्चा याचनाध्येषणा सनिः ॥३८८॥

AMAR - 'सनिस्त्वध्येषणा' इति गुर्वादिः क्वचिदर्थे प्रार्थनया नियोजनमित्यमरः पृथगाह ।

31. ABCH - पृथुकश्चिपटस्तुल्यौ (४०१)

AMAR - चिपिटक इत्यमरः ।

32. ABCH - उत्कस्तूत्सुक उन्मनाः । उत्कण्ठितः (४३६)

AMAR - 'इष्टार्थोद्युक्त उत्सुकः' इत्यमरः पृथगाह ।

33. ABCH - पिचिण्डिलो बृहत्कुक्षिस्तुन्दितुन्दिकतुन्दिलाः । उदर्युदरिले (४५०)

AMAR - तुन्दिरस्यास्ति तुन्दिभोऽपीत्यमरः ।

ACVH - अत्र तुन्दिशब्दाद् भप्रत्ययश्चिन्त्यः ॥

34. ABCH - कुकरे कुणिः (४५३)

Nairukta (NRKT) - कुत्सितः पाणिरस्येति नैरुक्ताः ।

35. ABCH - पामनः कच्छुरस्तुल्यौ (४६०)

Bhoja (BHOJ) - पामाऽस्त्यस्य पामनः अङ्गादित्वान्नः, पामर इत्येके, द्वावपीति भोजः ।

36. ABCH - मान्द्यं व्याधिरपाटवम् (४६२)

VCSP - यद्वाचस्पतिः - 'विविधान् यः करोत्याधीन् व्याधिः स हि निरुच्यते ॥'

37. ABCH - अरुरीर्म क्षणनुश्च (४६५)

ACVH - ईरयति अङ्गमीर्म "अर्तीरि-" (उणा-३३८) इति मः, क्लीबेऽयम् ।

VJYP - वैजयन्तीकारस्तु- 'व्रणोऽपीर्मोऽपि न स्त्रियाम्' इति पुंस्यप्याह ।

38. ABCH - शोथस्तु श्वयथुः शोफे (४६८)

ACVH - शवति विसर्पति शोथः, पुँल्लिङ्गः ।

VJYP - क्लीबेऽपि वैजयन्ती, यदाह- 'शोथोऽस्त्री श्वयथुः शोफः' ।

39. ABCH - व्याधिप्रभेदा विद्रधिभगन्दरज्वरादयः ॥ ४७१ ॥

ACVH - विद्रो धीयतेऽस्यामिति वा स्त्रीलिङ्गोऽयम् ।

VJYP - पुंस्यपि वैजयन्ती, यदाह - 'विद्रधिर्न षण्' ॥

40. ABCH – सत्कृत्यालङ्कृतां कन्यां यो ददाति स कूकुदः ॥ ४७५ ॥

Manu (MANU) – ब्राह्मे विवाहे दाताऽयम्, यन्मनुः -

“आच्छाद्य चार्हयित्वा च श्रुतशीलवते स्वयम् ।

आहूय दानं कन्याया ब्राह्मे धर्मः प्रकीर्तितः ॥”

41. ABCH – चपलश्चिकुरः (४७६)

Kumārasambhavam - यल्लक्ष्यम् - “तां वीक्ष्य लीलाचिकुरामनङ्गः स्वचापसौन्दर्यमदं मुमोच”. This quotation is from Kumārasambhavam of Kālidāsa (verse 1.47). Only difference is in the word चतुरा and चिकुरा.

42. ABCH - कुलिकस्तु कुलश्रेष्ठी (४८५)

AMAR – कुलं कायति कुलकः, इत्यमरः ।

43. ABCH – शारः शारिश्च खेलनी (४८७)

ACVH – शृणत्येनां शारिः स्त्रीलिङ्गः ।

Śābdika (SHBD) - शाब्दिकास्तु परिणयेन शारीन् हन्ति पुंस्यप्युदाहरन्ति ।

44. ABCH – स्यान्मनोजवसस्तातुल्यः (४८८)

ACVH – ‘पितृसदृशः मनोजव इत्यन्ये’ ।

VYDI – ‘यद्व्याडिः - “जनः पितृसधर्मा यः स तातार्हो मनोजवः” ॥

45. ABCH - अथ सेवा भक्तिः परिचर्या प्रसादना ॥४९६॥

शुश्रूषाराधनोपास्तिवरिवस्यापरीष्टयः । उपचारः (४९६)

AMAR – अमरस्तु- पर्येषणा परीष्टिश्च श्राद्धे द्विजशुश्रूषेति विशेषमाह ॥

46. ABCH - अथावेशिकागन्तू प्राघुणोऽभ्यागतोऽतिथिः । प्राघूर्णके (४९९)

ACVH - अवेशे अप्रतिवेशे भव आवेशिकः ।

SMRT- “नैकग्रामीणमतिथिं विप्रं सांगतिकं तथा” इति स्मृतेः ।

47. ABCH - प्रागल्भ्यौदार्यमाधुर्यशोभाधीरत्वकान्तयः । दीप्तिश्चाऽयत्नजाः (५०९)
 ACVH – यत्नाज्जाता यत्नजाः क्रियात्मानः ।
 Padārthavid - “इच्छातो यत्नस्ततो देहे क्रिया” इति हि पदार्थविदः ।
48. ABCH - मध्यमा तु दृष्टरजास्तरुणी युवतिश्चरी । तलुनी दिक्करी (५११)
 Śiśupālavadham - ‘यल्लक्ष्यम् “परिणतदिक्करिकास्तटीर्बिभर्ति” इति । (4.21)
49. ABCH - सुवासिनी वधूटी स्याच्चिरिण्टी (५१२)
 Dramila (DRML) - स्ववासिनीति द्रमिलाः ।
 VYDI - वध्वटीत्यपि व्याडिः ।
 AMAR – चिरण्टीत्यमरः ।
 Others – चरण्टीत्यन्ये ।
50. ABCH – जनी जाया परिग्रहः (५१३)
 MANU – यन्मनुः - “जायायास्तद्धि जायत्वं यदस्यां जायते सुतः”
51. ABCH – कुलस्त्री कुलबालिका (५१५)
 AMAR – कुलपालिकेत्यमरः ।
52. ABCH - दिधिषूस्तु पुनर्भूर्द्विरूढा (५२५)
 MANU - मनुस्त्वन्यथाऽऽह- “ज्येष्ठायां यद्यनूढायां कन्यायामुह्यतेऽनुजा । सा चाग्रेदिधिषूर्ज्या पूर्वा तु दिधिषूर्मता ॥”
53. ABCH - ज्येष्ठेऽनूढे परिवेत्ताऽनुजो दारपरिग्रही (५२६)
 SMRT – यत्स्मृतिः- “येऽग्रजेष्वकलत्रेषु कुर्वते दारसंग्रहम् । ज्ञेयास्ते परिवेत्तारः परिवित्तिस्तु पूर्वजः” ॥
54. ABCH - अथ पतिव्रता ॥५२७॥ एकपत्नी सुचरित्रा साध्वी सती (५२८)
 MANU – यत्स्मृतिः - नास्ति स्त्रीणां पृथग् यज्ञो न व्रतमिति । (5.153)
55. ABCH - याति या प्रियं साऽभिसारिका । (५२९)
 Bharata (BHRT) - यद् भरतः- “हित्वा लज्जाभये श्लिष्टा मदेन मदनेन वा । अभिसारयते कान्तं सा भवेदभिसारिका” ॥

56. ABCH - दोहदं दोहदं श्रद्धा लालसा (५४९)
 AMAR - अमरस्तु "दोहदमिच्छाम्, इच्छातिरेकं तु लालसाम्' आह ॥
57. ABCH - सूतिमासि तु वैजननः (५४९)
 Śruti (SHRT) - “दशमे मासि सूते वै” इति श्रुतेः ।
 Yājñavalkyasmṛti (YAGY) - “नवमे दशमे चाऽपि प्रबलैः सूतिमारुतैः । निःसार्यते बाण इव यन्त्रच्छिद्रेण स ज्वरः” ॥ इति स्मृतेश्च नवमो दशमो वा मासः । (3.83)
58. ABCH - नन्दनः पुनः ॥५४९॥ उद्धोऽङ्गात्मजः सूनुस्तनयो दारकः सुतः । पुत्रः (५४२)
 MANU - यन्मनुः- “पुत्रान्नो नरकाद् यस्मात् पितरं त्रायते सुतः । तस्मात् पुत्र इति प्रोक्तः स्वयमेव स्वयम्भुवा” ॥
59. ABCH - क्षेत्रजो देवरादिजः ॥५४९॥
 MANU - यन् मनुः- “यस्तल्पजः प्रमीतस्य क्लीबस्य व्याधितस्य वा । स्वधर्मेण नियुक्तायां स पुत्रः क्षेत्रजः स्मृतः” ॥ (9.167)
60. ABCH - स्वजाते त्वौरसोरस्यौ (५५०)
 MANU - यन् मनुः- “स्वक्षेत्रे संस्कृतायां तु स्वयमुत्पादयेद् धियम् । तमौरसं विजानीयात् पुत्रे प्राथमकल्पिकम्” ॥ (6.166)
61. ABCH - मृते भर्तरि जारजः । गोलकोऽथामृते कुण्डः (५५०)
 SMRT - यत् स्मृतिः- “परनार्या प्रजायेते द्वौ सुतौ कण्डगोलकौ । पत्यौ जीवति कुण्डश्च मृते भर्तरि गोलकः” ॥
62. ABCH - भ्राता तु स्यात् सहोदरः ॥५५०॥ समानोदर्यसोदर्यसगर्भसहजा अपि । सोदरश्च (५५१)
 AMAR - “सगर्भ्यः” इत्यमरः ॥
63. ABCH - केलिर्द्रवः परीहासः क्रीडा लीला च नर्म च ॥५५५॥
 Unknown - क्रीडनं क्रीडा “केलिपरिष्वङ्कितोयम्” इति केचित् ॥
64. ABCH - तृतीयाप्रकृतिः पण्डः षण्डः क्लीबो नपुंसकम् ॥५६२॥
 AMAR - ‘शण्डः’ इत्यमरः ।

Others - 'शण्ठः' इत्यन्ये ॥

65. ABCH - श्रुतौ श्रवः । शब्दाधिष्ठानपैञ्चमहानादध्वनिग्रहाः ॥५७३॥ कर्णः श्रोत्रः श्रवणं च (५७४)

VCSP - पुंक्लीबलिङ्गः, यद् वाचस्पतिः- 'श्रवः पैञ्चमस्त्रियाम्' ॥

66. ABCH - सकोपभ्रूविकारे स्याद् भ्रुभ्रुभ्रुपरा कुटिः ॥५७९॥

GAUD - भ्रुकुटिभ्रुकुटी तु भतान्तराश्रयेण, स्त्रीलिङ्गाः, यद् गौडः- "अथ भ्रुकुटिर्भ्रुकुटिभ्रुकुटिभ्रुकुटिः स्त्रियः" इति ॥

67. ABCH - नक्रं नर्कुटकं शिङ्घिनी (५८१)

ACVH - नक्रं क्लीबलिङ्गोऽयम् ।

VCSP - वाचस्पतिस्तु- "नक्रो नासा विकूणिका" इति पुंस्याह ॥

68. ABCH - तत्प्रान्तौ सूक्कणी (५८१)

AMAR - 'सूक्किणी' इत्यमरः ।

69. ABCH - राजदन्तौ तु मध्यस्थानुपरिश्रेणिकौ क्वचित् ॥५८४॥

VJYP - वैजयन्तीकारस्तु- मध्यदन्ता राजदन्ता दंष्ट्रा तत्पार्श्वयोर्द्वयोः' इति चतुरो राजदन्तानाह ॥

70. ABCH - गलो निगरणः कण्ठः (५८८)

ACVH - कणयति शब्दायते कण्ठः कण्ठते वा पुंलिङ्गोऽयम् ।

VCSP - वाचस्पतिस्तु- 'ग्रीवामूले तु कण्ठोऽस्त्री' इति क्लीबेऽप्याह ॥

71. ABCH - कफोणिस्तु भुजामध्यं कफणिः कूर्परश्च सः ।

ACVH - कं फणति कफणिः स्त्रीलिङ्गः ।

VJYP - "कूर्परः कफणिर्नषण्" इति पुंस्यपि वैजयन्ती ॥

72. ABCH - संपिण्डिताङ्गुलिः पाणिर्मुष्टिर्मुस्तुर्मुचुट्यपि । संग्राहश्च (५९७)

ACVH - मुस्यति खण्डयति मुस्तुः "बहुलम्" ॥५११॥ १२॥ इति तुक् पुंस्त्रीलिङ्गौ ।

VJYP - यद्वैजयन्ती- "अक्लीबे मुष्टिमुस्तु द्वौ संग्राहो मुचुटिः स्त्रियाम्" इति ॥

73. ABCH - तौ युतौ पुनरञ्जलिः । (५९८)

NRKT - अम्यते जलमनेनेति नैरुक्ताः ।

Nāṭyaśāstra (NATY) - नाट्ये तु- “पताकाभ्यां तु हस्ताभ्यां संश्लेषादञ्जलिः स्मृतः” ॥ (9.128)

74. ABCH -तुन्दं तुन्दिर्भकुक्षी पिचण्डो जठरोदरे । (६०४)

ACVH - कुक्षिः पुंलिङ्गः ।

GAUD - गौडस्तु- “उदरं कुक्षिरक्लीबे” इति स्त्रियामप्याह ।

VCSP - वाचस्पतिस्तु- “कुक्षिस्तु जठराधारे” इति जठरात् पृथगाह ॥

ACVH - उदरं, क्लीबलिङ्गः ।

VJYP - वैजयन्तीकारस्तु- “जठरं चोदरं न ना” इत्याह;

Buddhisāgara (BDSG) - त्रिलिङ्गोऽयमिति बुद्धिसागरः ।

75. ABCH - अन्त्रं तु पुरीतति ॥६०५॥

ACVH - क्लीबलिङ्गोऽयम् ।

VCSP - वाचस्पतिस्तु-अन्त्रं पुरीतदस्त्रियाम्, इति पुंस्यप्याह ।

76. ABCH - तत्पार्श्वकूपकौ तु कुकुन्दरे ॥६०८॥

ACVH - क्लीबलिङ्गोऽयम् ।

BHAG - भागुरिस्तु- “कुकुन्दरौ समाचष्टे जनो जघनकूपकौ” इति पुंस्याह ।

77. ABCH - द्वयमप्यदः ॥६१०॥ गुह्यप्रजननोपस्थाः (६११)

ACVH - उपतिष्ठते संगच्छत उपस्थः, पुंस्ययम् ।

VJYP - वैजयन्तीकारस्तु- “उक्तं द्वयमुपस्थोऽस्त्री” इति क्लीबेऽप्याह ॥

78. ABCH - स्यादण्डं पेलमण्डकः ॥६११॥ मुष्कोऽण्डकोशो वृषणः (६१२)

ACVH - अण्डं, क्लीबलिङ्गः ।

VJYP - वैजयन्तीकारस्तु-“अस्त्रियो मुष्ककोशाण्डाः” इति पुंस्यप्याह ।

79. ABCH - चरणः क्रमणः पादः पदोऽहिश्चलनः क्रमः । (६१६)

ACVH - पद्यते पादः “पदरुज—” ॥५।३।१६॥ इति घञ् ।

Durga (DURG) - पादपि, यद् दुर्गः- “पादसमानार्थः पादप्यस्ति” ॥३॥

ACVH - पद्यतेऽनेन पदः पुंक्लीबलिङ्गः, वर्षादित्वादल् पदपि” ।

VYDI - यद् व्याडिः-“पत्पादोहिश्वरणोऽस्त्री” इति ॥

80. ABCH - मांसं पललजङ्गले । (६२२)

NRKT - मां स भक्षयिताऽमुत्र, इति नैरुक्ताः ।

MANU - यद् मनुः- “मां स भक्षयिताऽमुत्र यस्य मांसमिहाऽद्म्यहम् । एतन्मांसस्य मांसत्वं प्रवदन्ति मनीषिणः” ॥१॥

NRKT - कुङ् गतौ सौत्रः, क्रूयते क्रव्यम्, कृत्तविकृत्तस्य क्रव्यम्, इति नैरुक्ताः ॥

81. ABCH - बुक्का हृद् हृदयं वृक्का सुरसं च तदग्रिमम् ॥६२३॥

GAUD - बुक्का स्त्रियामाबन्तः, पुंलिङ्ग इति गौडः ।

BHAG - भागुरिस्तु—“अग्रमांसं भवेद् बुक्कम्” इति क्लीबमाह ॥

ACVH - वृज्यते वृक्का “निष्कतुरुष्क-” ॥ (उणा-२६) ॥ इति के निपात्यते, स्त्रीलिङ्गोऽयम् ।

GAUD - यद् गौडः- “स्त्रियां वृक्का बुक्कः सुरसमद्वयोः” इति ।

VJYP - वैजयन्तीकारस्तु-“वृक्कौ पार्श्वगतौ गुलौ” इत्याह ।

AMAR - “बुक्काऽग्रमांसं, हृदयं च जीवाधारपद्मम्, इत्यमरः पृथक् पृथगाह ॥

82. ABCH - मेदोऽस्थिकृद् वपा मांसात् तेजो-जे गौतमं वसा ॥६२४॥

Suśrutasaṃhitā - “शुद्धस्य मांसस्य यः स्नेहः सा वसा” इति वैद्याः ॥ (śarīrasthāna IV.13)

83. ABCH - गोदं तु मस्तकस्नेहो मस्तिष्को मस्तुलुङ्गकः । (६२५)

ACVH - गवते गोदं कुमुदादौ निपात्यते, क्लीबलिङ्गोऽयम् ।

VCSP - वाचस्पतिस्तु-“गोदोऽस्त्री मस्तुलुङ्गकम्” इति पुंस्यप्याह ॥

84. ABCH - मेदोजं कीकसं सारः (६२६)

ACVH - सरति कालान्तरमिति सारः, हड्डं देश्यां, संस्कृतेऽपि ।

VJYP - यद् वैजयन्ती-“अथास्थि कीकसं हड्डम्” इति ॥

85. ABCH - पृष्ठस्याऽस्थि कशेरुका । (६२७)

ACVH - कस्यते कशेरु, क्लीबलिङ्गः “शिगुगे-” ॥ (उणा-८११) ॥ इति साधुः, स्त्रियामपि ।

VJYP - वैजयन्ती यदाह—“पृष्ठस्यास्थि कशेरुका” इति ।

ACVH - के कशेरुका स्त्रीकलीबलिङ्गः ।

Muni (MUNI) - ‘कशारुका’ इति मुनिः ॥

86. ABCH - मज्जा तु कौशिकः शुक्रकरोऽस्थनः स्नेहसम्भवौ ॥६२८॥

ACVH - मज्जत्यस्थनि मज्जा “उक्षितक्षि—” ॥ (उणा-१००) ॥ इत्यन्, नन्तः पुंस्ययम् ।

VCSP - वाचस्पतिस्तु- “अथ मज्जा द्वयोः” इति स्त्रियामप्याह ।

ACVH - वयं तु ब्रूमः- मज्जयतेऽनयाऽस्थिनीति भिदाद्यङि मज्जा स्त्रीलिङ्गः ॥

87. ABCH - त्वक् छविश्छादनी कृत्तिश्चर्माऽजिनमसृग्धरा ॥६३०॥

ACVH - असृग् धरत्यसृग्धरा ।

AMAR - अमरस्तु- “अजिनं चर्म कृत्तिः स्त्री” इति कृत्यादीन् त्रीन् मृगयोनित्वात् पृथगाह; अत एव ‘मृगा अजिनयोनयः’ इत्युक्ताः ।

VCSP - वाचस्पतिरपि- “तत्राजिनं मृगयोनिर्मृगाश्च प्रियकादयः । मृगप्रकरणे तेऽथ प्रोक्ता अजिनयोनयः” ॥

88. ABCH - वस्नसा तु स्नसा स्नायुः (६३१)

ACVH - स्नायुः “कृवापा-” ॥ (उणा-१) ॥ इत्युण्, स्त्रीलिङ्गोऽयं; क्लीबेऽपि ।

VCSP - वाचस्पतिर्यदाह- “अथ स्नायुर्नसा तन्त्रीः” इति ।

89. ABCH - कण्डरा तु महास्नायुः (६३१)

VAID - “स्नायुसंघातः” इति वैद्याः ॥

90. ABCH - सृणीका स्यन्दिनी लालाऽस्यासवः कफकूचिका । (६३३)

ACVH - ‘सुणिका’ इत्यमरः ।

91. ABCH - विङ् विष्ठाऽवस्करः शकृत् । (६३४)

ACVH - ‘विशति पक्वाशये विट्, स्त्रीलिङ्गः ।

VJYP - वैजयन्तीकारस्तु- ‘उच्चारो विङ् न ना” इति क्लीबेऽप्याह ।

AMAR - अमरस्तु-“विष्ठाविषौ स्त्रियाम्” इति मूर्धन्यं षमाह ॥

92. ABCH - कर्पूरागुरुकक्कोलकस्तूरीचन्दनद्रवैः ॥६३८॥ स्याद् यक्षकर्दमो मिश्रैः (६३९)

Dhanvantarinighaṇṭu (DHNV) - धन्वन्तरिस्तु- "कुङ्कुमागुरुकस्तूरी कर्पूरं चन्दनं तथा ।

महासुगन्धमित्युक्तं नामतो यक्षकर्दमः" (7.48)

93. ABCH - अगुर्वगराजार्हं लोहं कृमिजवंशिके । अनार्यजं जोङ्गकं च (६४०)

ACVH - वंशोऽस्या अस्ति वंशिका, स्त्रीक्लीबलिङ्गः ।

AMTK - "वंशाभं वंशकम्" इत्यमरटीका ॥

94. ABCH - हरिचन्दने ॥६४१॥ तैलपर्णिकगोशीर्षो (६४२)

ACVH - गोशीर्षगिरिभवत्वाद् गोशीर्षः, पुंस्येतौ ।

GAUD - गौडस्तु-"गोशीर्ष तैलपर्णिकम्" इति क्लीबमाह ॥

95. ABCH - पत्राङ्गं रक्तचन्दनम् । कुचन्दनं ताम्रसारं रञ्जनं तिलपर्णिका ॥६४२॥

ACVH - तिलस्येव पर्णान्यस्यास्तिलपर्णी सैव तिलपर्णिका ।

Unknown - तिलपर्णी नदी आकरोऽस्या इत्येके ॥

96. ABCH - बाह्वीकं कुङ्कुमं वह्निशिखं कालेयजागुडे । (६४५)

ACVH - कुक्कयते कुङ्कुमं, क्लीबलिङ्गोऽयं ।

VCSP - वाचस्पतिस्तु- "निदाघेऽपि कुङ्कुमः सुखः स्यात्" इति पुंस्याह ।

97. ABCH - मुकुटं पुनः ॥६५०॥ मौलिः किरीटं कोटीरमुष्णीषं (६५१)

ACVH - मङ्कयते मण्डयते शिरोऽनेन मुकुटं "मङ्केर्मकमुकौ च" ॥ (उणा-१५४) ॥ इत्युटः, मुखे कुट्यते वा, क्लीबोऽयम् ।

VJYP - पुंस्यपि वैजयन्ती, यदाह-"मौलिः कोटीरमुष्णीषं किरीटं मुकुटोऽस्त्रियाः" ॥

98. ABCH - उत्क्षिप्तिका तु कर्णान्दुः (६५६)

VJYP - वैजयन्तीकारस्तु-"कर्णान्दूः" इत्यूकारान्तमाह ॥

99. ABCH - गुच्छमाणवमन्दराः ॥६६०॥

अपि गोस्तनगोपुच्छावर्धमर्धं यथोत्तरम् । (६६१)

Unknown - अन्ये तु- "चतुःषष्टिलतो हारोऽथाष्टहीनो यथोत्तरम् । रश्मिकलापो

माणवकोऽर्धहारोऽर्धगुच्छकः ॥१॥ कलापच्छन्दो मन्दरस्य गुच्छसप्ततियष्टिकः ।" इत्याहुः ॥

100. ABCH - पटः प्रोतः (६६७)

ACVH - पटति विस्तीर्यते पटः, त्रिलिङ्गः ।

AMAR - “सुचेलकः पटोऽस्त्री” इत्यमरो विशेषमाह ॥११॥

ACVH - प्रकर्षेणोयते प्रोतः ।

GAUD - पुंसि गौडः ।

VCSP - क्लीबे वाचस्पतिः ॥

101. ABCH - तत्स्यादुद्गमनीयं यद्धौतयोर्वस्त्रयोर्युगम् । (६६८)

ACVH - उद्गम्यतेऽभिलष्यते उद्गमनीयम्, युगमविवक्षितं तेनैकमपि ।

BHAG - यद् भागुरिः- “धीरैरुगमनीयं तु धौतवस्त्रमुदाहृतम्”

Kumārasambhavam - लक्ष्यं च- “गृहीतवत्युद्गमनीयवस्त्रे” इति ॥ (7.11)

102. ABCH - कच्छा कच्छाटिका कक्षा परिधानाऽपराञ्चले ॥६७५॥

ACVH - कच्छैव कच्छाटी ग्रामटीवधूटीवत् के कच्छाटिका ।

VYDI - “कच्छाटी पुंस्त्रियोः” इति व्याडिः ॥

103. ABCH - निचोलः प्रच्छदपटो निचुलश्चोत्तरच्छदः ॥६७६॥

ACVH - निचोलति निचुलः,

GAUD - गौडस्तु-“निचोलस्तु निचुलकम्” इति क्लीबमाह ॥

104. ABCH - कुथे वर्णः परिस्तोमः प्रवेणीनवतास्तराः । (६८०)

ACVH - आस्तीर्यते आस्तरः, आस्तरणमपि; अयं च हरित रथप्रौवरणादावुपयुज्यते ।

VCSP - यद् वाचस्पतिः-“विनियोगस्त्वस्य हस्तिरथप्रावरणादिषु” इति ॥

105. ABCH - पाले पतद्ग्रहः ॥६८३॥ प्रतिग्राहः (६८४)

ACVH - पाल्यते पालः, पुंलिङ्गः तत्र ।

VJYP - क्लीबेऽपि वैजयन्ती यदाह- “पतद्ग्रहः प्रतिग्राहः पालोऽस्त्री” इति ॥

106. ABCH - कसिपुर्भोजनाच्छादौ (685)

AMAR - अमरस्तु- ‘कशति क्लेशं कशिपुः’ इति तालव्यमध्यमाह ।

107. ABCH - यावालक्तौ तु तद्रसः । (६८६)

Dhanapāla (DHAN) - यद् धनपालः-“तद्रागो यावकोऽलक्तकः स्मृतः”

AMAR - अमरादयस्तु-यावालक्तौ लाक्षादिभिः सहैकार्थावाहुः ॥

108. ABCH - सुभूमस्तु कार्तवीर्यः (६९३)

Hemacandra (HEMA) - यदवोचाम- “गृह्णन् भूमिं सुखेनाभूत् सुभूमो नामतस्ततः” ॥

109. ABCH - प्राजापत्यस्त्रिपृष्ठः (६९५)

Triṣaṣṭiśalākāpuruṣacaritam of Hemacandra - यदवोचाम- “दृष्ट्वा वंशत्रयं पृष्ठे त्रिपृष्ठ इति भूपतिः । नाम तस्याकरोत्सूनोरुत्सवेन महीयसा” ॥ (1.234)

110. ABCH - मान्धाता युवनाश्वजः ॥ ७०० ॥

धुन्धुमारः कुवलाश्वः हरिश्चन्द्रस्त्रिशङ्कुजः ।

पुरूरवा बौध ऐल उर्वशीरमणश्च सः ॥ ७०१ ॥

दौष्यन्तिर्भरतः सर्वदमः शकुन्तलात्मजः ।

हैहयस्तु कार्तवीर्यो दोःसहस्रभृदर्जुनः ॥ ७०२ ॥

Unknown - यदाहुः- “मान्धाता धुन्धुमारश्च हरिश्चन्द्रः पुरूरवाः । भरतः कार्तवीर्यश्च षडेते चक्रवर्तिनः” ॥

This quote is also found in Siddhaśabdārṇava of Sahajakīrti (verse 119).

111. ABCH - अजातशत्रुः शल्यारिर्धर्मपुत्रो युधिष्ठिरः । कङ्कोऽजमीढः (७०७)

ACVH - कङ्को ब्राह्मणलिङ्गत्वात् ।

GAUD - यद् गौडः- “कृतान्ते लोहपृष्ठे च कङ्को ब्राह्मणलिङ्गिनि” ।

112. ABCH - विषयानन्तरो राजा शत्रुः मित्रमतः परम् । उदासीनः परतरः पार्ष्णिग्राहस्तु पृष्ठतः ॥

७३२ ॥

Unknown - यदाहुः- “अरिर्मित्रमरेर्मित्रं मित्रमित्रमतः परम् । तथारिमित्रं मित्रं च विजिगीषोः पुरः स्मृताः ॥

१॥ पार्ष्णिग्राहस्तथाक्रन्द आसारौ च तयोः पृथक् । मध्यमोऽथाप्युदासीन इति द्वादशराजकम्” ॥२॥

113. ABCH - सन्धिविग्रहयानान्यासनद्वैधाश्रया अपि । षड्गुणाः (७३५)

ACVH - द्वौ प्रकारौ द्विधा द्विधैव द्वैधमेकेन सन्धायान्यत्र यात्रेत्यर्थः ।

Unknown - यद्वा-- बलिनोर्द्विषितोर्मध्ये वाचात्मानं समर्पयन् । द्वैधीभावेन वर्तेत काकाक्षिवदलक्षितः ॥१॥

114. ABCH - दण्डः स्यात् साहसं दमः ॥७३६॥

ACVH - सहसि बले भवं साहसं क्लीबलिङ्गः ।

VJYP - वैजयन्ती तु-“दण्डो दमः साहसोऽस्त्री” इति पुंस्यप्याह ॥

115. ABCH - यत्कोशदण्डजं तेजः स प्रभावः प्रतापवत् । (७४०)

BHRT - यद् भरतः- अधिकोपापमानादेः प्रयुक्तस्य परेण यत् । प्राणात्ययेऽप्यसहनं तत्तेजः समुदाहृतम् ॥

१॥

116. ABCH - भिया धर्मार्थकामैश्च परीक्षा या तु सोपधा ॥७४०॥

Kautilya (KAUT) - यत्कौटिल्यः- “उपधाभिः शौचाशौचपरिज्ञानममात्यानाम्” इति ॥

117. ABCH - रचना तु स्याद् व्यूहो दण्डादको युधि । (७४७)

Unknown - यदाहुः- “दण्डो मण्डलभागौ चाप्युच्छन्नश्चाबलो दृढः । व्यूहास्तेषां विशेषाः स्युश्चक्रव्यूहादयोऽपि च” ॥

118. ABCH - एकेभैकरथास्त्र्यश्वाः पत्तिः पञ्चपदातिका । (७४८)

Unknown - यदाहुः- “[^1] एगो हत्थी एगो य रहवरो तिन्नि चेव य तुरंगा । पंचेव य पाइक्का एसा पत्ती मुणेयव्वा” ॥

119. ABCH - सेना सेनामुखं गुल्मो वाहिनी पृतना चमूः ॥७४८॥ अनीकिनी च पत्तेः स्यादिभाद्यैस्त्रिगुणैः क्रमात् । (७४९)

Unknown - यदाहुः- “[^2]पत्ती सेणा सेणामुहं च गुम्मं च वाहिणी चेव । पियणा चमू अणीगणी दसगुणीआ अक्खोहिणी होइ” ॥

AMAR - अमरादयस्तु- “सेनामुखं गुन्मगणो वाहिनी पृतना चमूः । अनीकिनी” इत्येवं क्रमेण नामान्याहुः ॥

120. ABCH - दशानीकिन्योऽक्षौहिणी (७४८)

Unkknown - यदाहुः- “स्यात्सेनाऽक्षौहिणी नाम खाऽगाष्टैकद्विकैर्गजैः । रथैश्चैभ्यो हयैस्त्रिघ्नैः पञ्चघ्नैश्च पदातिभिः ॥”

121. ABCH - प्रासङ्गस्तु युगान्तरम् । (७५७)
MUNI - यन्मुनिः-युगं द्वितीयं प्रासङ्गः ॥
122. ABCH - शिबिका याप्ययाने (७५८)
GAUD - “कृपानम्” इति गौडः ।
123. ABCH - सारसनं त्वधिकाङ्गं हृदि धार्य सकञ्चुकैः ॥७६७॥
ACVH - अधिकमङ्गादधिकाङ्गं पुंक्लीबलिङ्गः,
Unknown - अधियाङ्गमित्येके ।
MUNI - यन्मुनिः- “अधियाङ्गं सारसनम्” ।
DURG- दुर्गस्तु-“तस्य सारसनं ज्ञेयं धियाङ्गं च निबन्धनम्” इत्याह ।
124. ABCH - गोधा तु तलं ज्याघातवारणम् ॥७७६॥
ACVH - तलति तलं क्लीबलिङ्गः स्त्रियामपि ।
VJYP - वैजयन्तीकारस्तु यदाह- “गोधा तला च ननरौ” ।
125. ABCH - स्थानान्यालीढवैशाखप्रत्यालीढानि मण्डलम् । समपादं च (७७७)
Dhanurveda (DNVD) - यद् धनुर्वेदः-
अग्रतो वामपादं तु तीक्ष्णं चैवाऽनुकुञ्चितम् ।
आलीढं तु प्रकर्तव्यं हस्तद्वयसविस्तरम् ॥१॥
पादौ सविस्तरौ कार्यौ समहस्तप्रमाणतः ।
वैशाखस्थानके वत्स ! कूटलक्ष्यस्य वेधने ॥२॥
प्रत्यालीढे तु कर्तव्यः सव्यस्तीक्ष्णोऽनुकुञ्चितः ।
तिर्यग्वामः पुरस्तत्र दूरापाते विशिष्यते ॥३॥
समपादे समौ पादौ निष्कम्पौ च सुसङ्गतौ ।
मण्डले मण्डलाकारौ बाह्यतीक्ष्णौ विशेषतः ॥४॥
ACVH - एते क्लीबलिङ्गाः ।
VJYP - वैजयन्तीकारस्तु-“तत्र वैशाखमस्त्रियाम्” इत्याह ॥

126. बाणे पृषत्कविशिखौ (७७८)

NRKT - पुङ्गादीनां पृथक् षट्कमस्येति नैरुक्ताः ।

DNVD- यद् धनुर्वेदः-“पुङ्गः शरस्तथा शल्यं पक्षस्नायुजतूनि षट्” इति ॥

127. ABCH - बाणमुक्तिर्व्यवच्छेदः (७८०)

DNVD - यद् धनुर्वेदः- “पञ्चाङ्गुलीभिर्युगपत् क्षिप्रश्रेष्ठत्वसिद्धये । मोक्षणं यच्च दुर्जेयं तं व्यवच्छेदमादिशत् ॥”

128. ABCH - त्सरुरस्य मुष्टिः ॥७८२॥

ACVH - खड्गस्येत्युपलक्षणम्, तेनान्यस्याऽपि मुष्टिस्त्सरुरुच्यते ।

AMAR - यदमरः-“त्सरुः खड्गादिमुष्टौ स्याद्” इति ॥

129. ABCH - लोहाभिसारो दशम्यां विधिर्नीराजनात् परः । (७८९)

DURG - यद् दुर्गः- “लोहाभिसारस्तु विधिः परो नीराजनाद् नृपैः दशम्यां दंशितैः कार्यः” इति ।

AMAR - अमरस्तु-“लोहाभिसारोऽस्त्रभृतां राज्ञां नीराजनाविधिः । “इति ब्रुवन्नीराजनामेव लोहाभिसारमाह ॥

130. ABCH - प्रसारस्तु प्रसरणं तृणकाष्ठादिहेतवे ।

AMAR - अमरस्तु-“स्यादासारः प्रसरणम्” इत्याह ।

ACVH - तच्चाऽर्थशास्त्रेण न संवदतीति नादृतम् ॥

131. ABCH - वैतालिका बोधकरा अर्थिका सौखसुप्तिकाः ।

AMAR - अमरस्तु-“वैतालिका बोधकराः” इति राज्ञोऽवसरपाठकानाह ॥

132. ABCH - मागधो मगधः (७९५)

ACVH - मगध्यति याचते मगधः, वंशोदीरणेन यो याचते ।

Maṅkha (MANK) - यदाहुः-“मागधाः स्तुतिवंशजाः” इति मङ्गोऽपि ॥

133. ABCH - द्वन्द्वं समाघातसमाह्वयाभिसम्पातसंमर्दसमिप्रघाताः । (७९७)

ACVH - द्वौ द्वौ युध्यतेऽत्र द्वन्द्वं पृषोदरादित्वात्, वन्द्यते वीरैरिति वा “प्रह्वाह्वा-” ॥ (उणा-५१४) ॥ इति वे निपात्यते ।

- BHOJ - 'समासकलहयुग्मेषु' इति भोजः ॥
134. ABCH - पटहाडम्बरौ तुल्यौ (७९९)
ACVH - यद् वाचस्पतिः-“महासमरतूर्ये तु पटहाडम्बरवुभौ” ॥
135. ABCH - नासीरं त्वग्रयानं स्यात् (८००)
ACVH - स्त्रीक्लीबलिङ्गोऽयं ।
VJYP - वैजयन्ती तु- “नासीरोऽस्त्री” इत्याह ॥
136. ABCH - वीराशंसनं त्वाजिभीष्मभूः । (८०१)
AMAR - “वीराशंसनी” इत्यमरः ॥
137. ABCH - बलात्कारस्तु प्रसभं हठः (८०४)
VJYP - वैजयन्ती तु-“प्रसभोऽस्त्री बलात्कारः” इति पुंस्यप्याह ॥
138. ABCH - बन्धां ग्रहकः प्रोपतो ग्रहः ॥८०६॥
Unknown - यदाहुः-‘बन्दी स्यात् पणबन्धस्थः’ ।
139. ABCH - ज्येष्ठाश्रमी गृहमेधी गृहस्थः स्नातको गृही ॥८०८॥
MANU - यन्मनुः- "यस्मात्त्रयोऽप्याश्रमिणो ज्ञानेनान्नेन चान्वहम् ।
गृहस्थैरेव धार्यन्ते तस्माद् ज्येष्ठाश्रमो गृही ॥"
140. ABCH - ब्राह्मणस्तु त्रयीमुखः ॥८११॥
भूदेवो वाडवो विप्रो द्व्यग्राभ्यां जातिजन्मजाः ।
वर्णज्येष्ठः सूत्रकण्ठः षट्कर्मा मुखसम्भवः ॥८१२॥
वेदगर्भः शमीगर्भः सावित्रो मैत्र एतसः ॥ (८१३)
SMRT - “मातुरग्रे विजननं द्वितीयं मौञ्जिवन्धनम्” इति स्मृतेः
Śruti (SHRT) - यत्-श्रुतिः-“ब्राह्मणोऽस्य मुखमासीत्” ।
MANU – यन्मनुः- “तत्र यद् ब्रह्मजन्माऽस्य मौञ्जीबन्धनचिह्नितम् ।
तत्राऽस्य माता सावित्री पिता त्वाचार्य उच्यते ॥”
141. ABCH - औदुम्बर उलूखलः । (८१६)

MANU - मनुस्तु-"ब्राह्मणो बैल्वपालाशौ क्षत्रियो वाटखादिरौ ।

पैलवौदुम्बरौ वैश्यो दण्डानर्हन्ति धर्मतः ॥" इत्याह ।

142. ABCH - अध्ययनं ब्रह्मयज्ञः स्याद् देवयज्ञ आहुतिः ।

होमो होत्रं वषट्कारः पितृयज्ञस्तु तर्पणम् ॥ ८२१ ॥

तच्छ्राद्धं पिण्डदानं च नृयज्ञोऽतिथिभोजनम् ।

भूतयज्ञो बलिः पञ्च महायज्ञा भवन्त्यमी ॥ ८२२ ॥

MANU - यन्मनुः- "अध्यापनं ब्रह्मयज्ञः पितृयज्ञस्तु तर्पणम् ।

होमो देवो बलिर्भूतो नृयज्ञोऽतिथि पूजनम् ॥"

143. ABCH - यूपः स्याद् यज्ञकीलकः ॥८२४॥

144. ACVH - पुंलिङ्गोऽयम् ।

VJYP - वैजयन्ती तु-"यूपोऽस्त्री संस्कृतस्तम्भः" इति क्लीबेऽप्याह ।

145. यूपाग्रभागे स्यात् तर्म् (८२५)

वाचस्पतिस्तु-"यूपाग्रं तर्म् न स्त्रियाम्" इति पुंस्यप्याह ॥

146. ABCH - समिदिन्धनमेधेध्मतर्पणैर्धांसि (८२७)

ACVH - इध्मं क्लीबलिङ्गोऽयम् ।

VJYP - वैजयन्ती तु- "होमेन्धनं स्यादिध्मोऽस्त्री" इति पुंस्यप्याह ॥

147. ABCH - हव्यं सुरेभ्यो दातव्यं पितृभ्यः कव्यमोदनम् । (८३२)

Śrutijña - उभयमुभयत्रेति श्रुतिज्ञाः ।

148. ABCH - विघसो भुक्तशेषके । (८३४)

MANU - यन्मनुः-विघसाशी भवेद् नित्यं नित्यं चामृतभोजनः ।

विघसो भुक्तशेषः स्याद् यज्ञशिष्टमथाऽमृतम् ॥

149. ABCH - पूर्तं वाप्यादि (८३४)

Unknown - आह च-"वापीकूपतडागानि देवतायतनानि च ।

अन्नप्रदानमारामाः पूर्तमार्याः प्रचक्षते ॥"

150. ABCH - इष्टं मखक्रिया ॥८३४॥

Unknown – यदाह- "एकाग्रिकर्म हवनं त्रेतायां यच्च हूयते ।

अन्तर्वेद्यां यद्दानमिष्टं तदभिधीयते ॥"

151. ABCH - उपस्पर्शस्त्वाचमनं (८३७)

MANU - यन्मनु:- "उपस्पृश्य द्विजो नित्यमन्नमद्यात् समाहितः ।

भुक्त्वा चोपस्पृशेत् सम्यगग्निः खानि च संस्पृशेत् ॥

152. ABCH - पाठे स्याद् ब्रह्माञ्जलिरञ्जलिः ॥८३८॥

Unknown – यदाह- "संहत्य हस्तावध्येयं स हि ब्रह्माञ्जलिः स्मृतः ॥"

153. ABCH - दैवतं त्वङ्गुलीमुखे ॥८४०॥

Yājñavalkya (YJVL) - यद् याज्ञवल्क्यः- "कनिष्ठातर्जन्यङ्गुष्ठमूलान्यग्रं करस्य च ।

प्रजापतिपितृब्रह्मदेवतीर्थान्यनुक्रमात् ॥"

154. ABCH - औपवस्त्रं तूपवासः (८४२)

SMRT - यत् स्मृतिः- "माषान् मधुमसूरांश्च वर्जयेदौपवस्त्रके' ॥

155. ABCH - कृच्छ्रं सान्तपनादिकम् ॥८४२॥

SMRT - यत् स्मृतिः- "त्र्यहं सायं त्र्यहं प्रातः त्र्यहमद्यादयाचितम् ।

त्र्यहं परं च नाशनीयात्, कृच्छ्रं सान्तपनं स्मृतम् ॥"

156. ABCH - समास्तु पादग्रहणाभिवादनोपसंग्रहाः ॥८४४॥

MANU - यन्मनु:- "व्यत्यस्तपाणिना कार्यमुपसंग्रहणं गुरोः ।

वामेन वामः स्प्रष्टव्यो दक्षिणेन तु दक्षिणः ॥३॥ "

157. ABCH - निवीतं कण्ठलम्बितम् ॥८४५॥

MANU – यन्मनु:- "प्रोद्धृते दक्षिणे पाणावुपवीत्युच्यते द्विजः ॥

सव्ये प्राचीनमावीती निवीती कण्ठसर्जने ॥"

158. ABCH - वेदव्यासस्तु माठरः ॥८४६॥

ACVH - वेदार्थान् व्यासयति प्रपञ्चयति वेदव्यासः ।

Bhārata - यद् भारते- “वेदार्था भारते न्यस्ताः” इति ॥

159. ABCH - क्षतव्रतोऽवकीर्णी स्यात् (८५४)

SMRT - यत् स्मृतिः- ब्रह्मचार्यवकीर्णी स्यात् कामतस्तु स्त्रियं व्रजन् ॥

160. ABCH - व्रात्यः संस्कारवर्जितः ॥८५४॥

MANU - यन्मनुः-अत ऊर्ध्वं त्रयोऽप्येते यथाकालमसंस्कृताः ।

सावित्रीपतिता व्रात्या भवन्त्यार्यविगर्हिताः ॥

161. ABCH - शिश्विदानः कृष्णकर्मा (८५५)

AMAR - “शिश्विदानः कृष्णकर्मा” इत्यमरः ॥

162. ABCH - वेदहीनो निराकृतिः । (८५६)

SMRT - यत् स्मृतिः- “अनधीत्य तु यो वेदान् अन्यत्र कुरुते श्रमम् ।

स जीवन्नेव शूद्रत्वमाशु गच्छति सान्वयः ॥”

163. ABCH - क्षत्रं तु क्षत्रियो राजा राजन्यो बाहुसंभवः ॥८६३॥

SHRT - यद् श्रुतिः-“बाहू राजन्यः कृतः” ॥

164. ABCH - अर्या भूमिस्पृशो वैश्या ऊरुव्या ऊरुजा विशः । (८६४)

SHRT - यत् श्रुतिः-“ऊरु तदस्य यद्वैश्यः” इति ॥

165. ABCH - सत्यानृतं तु वाणिज्यं वणिज्या (८६७)

MANU - यन्मनुः- “ऋतमुञ्छशिलं ज्ञेयममृतं स्यादयाचितम् ।

मृतं तु याचितं भैक्षं प्रमृतं कर्षणं स्मृतम् ॥१॥

सत्यानृतं तु वाणिज्यं तेनाऽपि खलु जीव्यते ।

सेवा श्ववृत्तिराख्याता तस्मात् तां परिवर्जयेत्” इति ॥

166. ABCH - निक्षेपोपनिधी न्यासे (८७०)

Smārta - स्मार्ते त्वेषां भेदोऽस्ति-

“वासनस्थमनाख्याय हस्तेऽन्यस्य यदर्पितम् ।

द्रव्यं तदुपनिधिर्न्यासः प्रकाश्य स्थापितं तु यत् ॥१॥

विक्षेपः शिल्पिहस्ते तु भाण्डं संस्कर्तुमर्पितम्” इति ।

167. ABCH - गण्यं गण्यं संख्येयं

VCSP - यद् वाचस्पतिः- “गणनीयं तु गण्यम्” इति ॥

168. ABCH - संख्या त्वेकादिका भवत् ॥८७२॥

ACVH - एक आदिरस्या एकादिका, आदिग्रहणात् द्वौ त्रयः चत्वारः एते वाच्यलिङ्गाः, पञ्च सप्तेत्यादयोऽष्टादशान्ता अलिङ्गाः ।

HEMA - यदवोचाम- “नन्ता सङ्ख्या डितिर्युष्मदस्मच्च स्युरलिङ्गकाः” । इति ।

ACVH - एकोनविंशतित्रिंशदित्यादयः शतादवाक् स्त्रीलिङ्गाः ।

HEMA - यदवोचाम स्त्रीलिङ्गनिर्णये- “विंशत्याद्या शताद् द्वन्द्वे” इति ।

ABCH - षष्टिस्तु त्रिलिङ्गनिर्णये पृथक्लिङ्ग उक्तः । अत्र च एकाद्या अष्टा दशान्ता सङ्ख्येये वर्तन्ते, विंशत्याद्यास्तु सङ्ख्याशब्दा एकत्वे वर्तमानाः सङ्ख्येये सङ्ख्याने च वर्तन्ते यथा विंशतिर्घटाः, विंशतिर्घटानाम्; शतं गावः, शतं गवामिति ।

VCSP - यदाह वाचस्पतिः- “अष्टादशभ्य एकाद्याः सङ्ख्याः सङ्ख्येयगोचराः । ऊनविंशत्यादिकास्तु सर्वाः सङ्ख्येयसङ्ख्ययोः” ॥१॥

Bhāṣyakāra (BSHY) - भाष्यकारोऽप्याह- “अष्टादशभ्यः सङ्ख्या सङ्ख्येये वर्तते, ततः परं सङ्ख्याने सङ्ख्येये च” इति ।

169. ABCH - यथोत्तरं दशगुणं भवेदेको दशायुतः ।

शतं सहस्रमयुतं लक्षप्रयुतकोटयः ॥८७३॥

अर्बुदमब्जं खर्वं च निखर्वं च महाम्बुजम् ।

शङ्कुर्वार्द्धिरन्त्यं मध्यं परार्द्धं चेति नामतः ॥८७४॥

Unknown - यदाह- “एकं दश शतमस्मात् सहस्रमयुतं ततः परं लक्षम् ।

प्रयुतं कोटिमथाऽर्बुदमब्जं खर्वं निखर्वं च ॥१॥

तस्माद् महासरोजं शङ्कुं सरितां पतिं ततस्त्वन्त्यम् ।

मध्यं परार्ध्यमाहुर्यथोत्तरं दशगुणं तज्ज्ञाः” ॥१॥ इति ।

170. ABCH - असङ्ख्यं द्वीपवार्ध्यादि (८७५)

Unknown - यदाहुः-“असङ्ख्याता द्वीपसमुद्राः” इति ।

171. ABCH - कर्णधारस्तु नाविकः (८७६)

DURG - यद् दुर्गः-“कर्णः श्रोत्रमरित्रं च”

172. ABCH - अथ द्रोणी काष्ठाम्बुबाहिनी । (८७७)

Kṣīrasvāmin (AKKS) - क्षीरस्वामी तु-काष्ठमुपलक्षणं काष्ठाऽश्मादिमयी जलधारिणी द्रोणी इति
ट ऋ यौ

173. ABCH - गुणवृक्षस्तु कूपकः ॥८७७॥

Unknown - पोते ध्वजपटाद्याधार इत्येके ॥

GAUD - गौडस्तु-जलान्तस्तरुन् कूपकानाह ॥

174. ABCH - स्याद् मङ्गो मङ्गिनीशिरः । (८७८)

ACVH - मङ्गः पुंलिङ्गोऽयम् ।

VJYP - वैजयन्ती तु- “नौशिरो मङ्गमस्त्रियाम्” इति क्लीबेऽप्याह ।

175. ABCH - तत्र स्याद् गुञ्जाः पञ्च माषकः ॥ ८८३ ॥

ते तु षोडश कर्षोऽक्षः पलं कर्षचतुष्टयम् ।

विस्तः सुवर्णो हेम्नोऽक्षे (८८४)

Unknown – यदाह-

“माषो दशार्धगुञ्जः षोडशमाषो निगद्यते कर्षः ।

ससुवर्णस्य सुवर्णस्तैरेव पलं चतुर्भिश्च” इति ॥

176. ABCH - कुरुविस्तस्तु तत्पले ॥८८४॥

MALA - यन्माला- “पलेन हि सुवर्णस्य कुरुविस्तः” ।

177. ABCH - ते दशाचितः ॥८८५॥

Unknown - क्लीबेऽयमित्यन्ये ।

178. ABCH - चतुर्विंशत्यङ्गुलानां हस्तः दण्डश्चतुष्करः ।

तत्सहस्रौ तो गव्यूतं क्रोशः तौ द्वौ तु गोरुतम् ॥ ८८७ ॥

गव्या गव्यूतगव्यूती चतुष्क्रोशं तु योजनम् । (८८८)

Unknown – यदाह-

"हस्तोऽङ्गुलविंशत्या चतुरन्वितया चतुष्करो दण्डः ।

तद्धि सहस्रं क्रोशो योजनमेकं चतुष्क्रोशैः" इति ॥

179. ABCH - गोदारणं तु कुद्वालः (८९२)

ACVH - पुंस्ययम् ।

VJYP - वैजयन्ती तु- "गोदारणं कुद्वालम्" इति क्लीबमाह ॥

180. ABCH - शूद्रोऽन्त्यवर्णो वृषलः पद्यः पज्जो जघन्यजः ॥८९४॥

Nārada (NARD) - नारदस्तु-“वृषो हि भगवान् धर्मस्तस्य यः कुरुते त्वलं वृषलं तं विजानीयात्”
इति ॥

SHRT - यत् श्रुतिः-“पद्भ्यां शूद्रोऽजायत” इति ॥

181. ABCH - माध्वीकं मदना देवसृष्टा कापिशमब्धिजा ॥९०३॥

ACVH - माद्यत्यनया माध्वीकम्, “सृणीकास्तीक-” ॥(उणा-५०)॥ इतीके निपात्यते ।

Unknown - मृद्वीकाया विकारो माध्वीकमित्यन्ये ॥

182. ABCH - गल्बर्कस्तु चषकः स्यात् सरकश्चानुतर्षणम् । (९०६)

AMAR - अमरस्तु सरकानुतर्षशब्दौ सुरापखिषणपर्यायावाह ॥

183. ABCH - व्योकारः कर्मारो लोहकारः (९२०)

BHOJ - व्यो इत्ययःपर्यायः इति भोजः ॥

184. ABCH - चण्डालेऽन्तावसाय्यन्तेवासिश्चपचबुक्कसाः ।

निषादप्लवमातङ्गदिवाकीर्तिजन माः ॥९३३॥

VYDI - यद् व्याडिः- “चण्डमालं मृषा यस्येत्यर्थः शब्दवतां मतः” इति ।

Paurāṇika - मा निषिद्धं तङ्गतीति वा, मातङ्गापत्यमिति पौराणिकाः ॥

185. ABCH - सर्वसहा रत्नगर्भा जगती मेदिनी रसा । (९३७)

BHAG - रत्नवतीति भागुरिः ।

186. ABCH - दिवस्पृथिव्यो रोदस्यौ रोदसी रोदसी च ते (९३९)

Utpala (UTPL) - यद् उत्पलः “द्यावापृथिव्योर्द्विवचने रोदसिशब्द इवर्णान्तादेशः पृषोदरादित्वादिति, रोदसीत्यव्ययमपि” ॥

187. ABCH - रौमकं तु रुमाभवम् । वसुकं वस्तूकं तच्च ।

ACVH - वसति वसु वस्वेव वसुकम् ।

MALA - यन्माला- “रौमकं वसुकं वसु” ॥

188. ABCH - कृष्णे तु तत्र तिलकं (९४३)

VAID - यद् वैद्याः- “कृष्णे सौवर्चलगुणा लवणे गन्धवर्जिते” ॥

189. ABCH - वर्ष वर्षधराद्यङ्गं (९४७)

190. Unknown - लौकिका नववर्षाण्याहुर्यथा-

भारतं प्रथमं वर्षं ततः किंपुरुषं स्मृतम् ।

हरिवर्षं तथैवान्यद् मेरोर्दक्षिणतो द्विजः ॥

रम्यकं चोत्तरं वर्षं तस्यैवानुहिरण्मयम् ।

उत्तराः कुर्वश्चैव यथा वै भारतं तथा ।

भद्राश्वं पूर्वतो मेरोः केतुमालं तु पश्चिमे ।

नवसाहस्रमेकैकमेतेषां द्विजसत्तम ! ॥

इलावृतं च तन्मध्ये तन्मध्ये मेरुरुत्थितः ॥२॥

191. ABCH - विषयस्तूपवर्त्तनम् ।

देशो जनपदो नीवृद् राष्ट्रं निर्गश्च मण्डलम् ॥९४७॥

Kātya (KATY) - “उपावर्त्तनम्” इति कात्यः ॥

ACVH - नियतं वर्त्तन्तेऽस्यां नीवृत् स्त्रीलिङ्गः ।

AMAR - अमरस्तु पुंस्याह ॥

AMAR - अमरस्तु- “नीवृज्जनपदो देशविषयौ तूपवर्त्तनम्” इति पृथगाह ॥

192. ABCH - आर्यावर्त्तो जन्मभूमिर्जिनचक्र्यर्द्धचक्रिणाम् ।

पुण्यभूराचारवेदी मध्यं विन्ध्यहिमाऽगयोः ॥९४८॥

VYDI - यद् व्याडिः-

"आसमुद्राच्च वै पूर्वादासमुद्राच्च पश्चिमात् ।

हिमवद्विन्ध्ययोर्मध्यमार्यावर्तं विदुर्बुधाः ॥" इति ॥

193. ABCH - पश्चिमोत्तरस्तूदीच्यः (९५२)

Unknown - यदाह-

"प्रागुदञ्चौ विभजते हंसः क्षीरोदके यथा ।

विदुषां शब्दसिद्ध्यै या सा वः पायाच्छरावती" ॥

194. ABCH - शाद्वलः शादहरिते (९५५)

ACVH - शादः शष्पं तेन नीले ।

SHAS - यद् शाश्वतः-“शष्पकर्दमयोः शादः” इति ॥

195. ABCH - गोस्थानं गोष्ठम् (९६४)

ACVH - गावस्तिष्ठन्त्यत्र गोष्ठं क्लीबलिङ्गः ।

SHBD - शब्दिकास्तु गोष्ठो भूतपूर्व इति पुंस्यपि प्रयुञ्जते ।

196. ABCH - चूर्णे क्षोदः (९७०)

SHAS - यत् शाश्वतः- “चूर्णानि वासयोगाः स्युश्चूर्णो धूलिः सशर्करः” इति ॥

197. ABCH - नगरी पूः पुरी द्रङ्गः पत्तनं पुटभेदनम् ॥९७१॥

निवेशनमधिष्ठानं स्थानीयं निगमोऽपि च । (९७२)

VCSP - वाचस्पतिस्तु

"स्यात् स्थानीयं त्वतिलम्बो ग्रामो ग्रामशताष्टके ।

तदर्हं तु द्रोणमुखं तच्च कर्वटमस्त्रियाम् ॥१॥

कर्वटार्द्धं कवुटिकं स्यात्तदर्थं तु कार्वटम् ।

तदर्थं पत्तनं तच्च पत्तनं पुटभेदनम् ॥२॥

निगमस्तु पत्तनाद्धे तदद्धे तु निवेशनम् ।

कर्वटादधमो द्रङ्गः पत्तनादुत्तमश्च सः ॥३॥

उद्रङ्गश्च निवेशश्च स एव द्रङ्ग इत्यपि ।" इति

विशेषमाह ॥

198. ABCH - शाखापुरं तूपपुरम् (९७२)

Unknown - यदाह- "शाखा वेदिविभागे स्यात् पादपाङ्गान्तिकेऽपि च" ।

199. ABCH - उज्जयनी स्याद् विशालाऽवन्ती पुष्पकरण्डिनी । (९७६)

VYDI - यद् व्याडिः- "उज्जयन्यां यदुद्यानं तत् स्यात् पुष्पकरण्डकम्" इति ॥

200. ABCH - मथुरा तु मधूपघ्नं मधुरा (९७८)

Unknown - मथुरायाः पिशाच्या अदूरभवं नगरमित्यन्ये ॥

201. ABCH - अथ गजाह्वयं स्याद् हास्तिनपुरं हस्तिनीपुरं हस्तिनापुरम् ॥९७८॥

Bhārata - यद् भारते- "नगरे नागसाह्वये" इति, "तेन गजपुरं नाम नगरम्, इत्यादि ॥

202. ABCH - परिकूटं हस्तिनखो नगरद्वारकूटके । (९८२)

ACVH - हस्तिनख इव हस्तिनखः दुर्गद्वारावतरणार्थः क्रमनिम्नो मृत्कूटोऽपसोपानाख्यः ।

Unknown - बहिरतटमन्तः सोपानयुक्तं युद्धार्थमित्येके ॥

203. ABCH - मुखं निःसरणे (९८२)

KTLY - यत् कौटिल्यः- "मुखसमः संक्रमः" इति ॥

204. ABCH - व्यध्वो दुरध्वः कदध्वा विपथं कापथं च सः ॥९८४॥

ACVH - "पथः सङ्ख्याऽव्ययोत्तरः" ॥ (हैमलिङ्गानु० नपुं० श्लो-८) ॥ इति क्लीबत्वम् ।

AMAR - अमरस्तु – विपथकापथशब्दौ पुंस्याह

GAUD - गौडश्च यदाह- "व्यध्वो विपथकापथौ" इति ॥

205. ABCH - धामाऽगारं निशान्तं च (९९२)

ACVH - निशाम्यन्त्यस्मिन्निशान्तम्, "अद्यर्थाच्चाधारे" ॥५।१।१२॥ इति क ।

Unknown - निशाया अन्तोऽत्रेति केचित् ॥

206. ABCH - चैत्यविहारौ जिनसद्मनि ॥९९४॥
 AMAR - अमरस्तु- “चैत्यमायतनं तुल्ये” इति, “बौद्धानां विहारोऽस्त्री” इति च चैत्यविहारौ भिन्नार्थावाह ॥
207. ABCH - आथर्वणं शान्तिगृहम् (९९७)
 ACVH - शान्त्यै गृहं शान्तिगृहं शान्तीगृहकमपि ।
 VCSP - यद् वाचस्पतिः- “आथर्वणं शान्तिगृहं शान्तीगृहकमप्यदः” इति ॥
208. ABCH - घोषस्त्वाभीरपल्लिका ।
 SHAS - “कुटीग्रामकयोः पल्लिः” इति तु शाश्वतः ॥
209. ABCH - कुड्यं भित्तिः (१००३)
 ACVH - क्लीबलिङ्गोऽयम् ।
 VCSP - वाचस्पतिस्तु- “कुड्यमस्त्रियाम्” इति पुंस्यऽप्याह ॥
210. ABCH - स्तम्भादेः स्यादधोदारौ शिला नासोर्ध्वदारुणि (१००८)
 GAUD - “स्तम्भोर्ध्वे दार्वन्तरस्थापनार्थं यदासज्यते सा शिला” इति गौडः ॥
 MALA - द्वारशाखाया अधोर्ध्व दारुणी शिलानासेति तु माला, यदाह- नासा दारुपरि द्वारस्याऽधो दारुशिला स्त्रियाम्” इति ॥
211. ABCH - पेटा स्याद् मञ्जूषा (१०१५)
 AMAR - ‘पीडण् गहने’ इत्यस्य पेडेत्यमरः ।
212. मुसलोऽयोऽग्रं (१०१७)
 Unknown - अयोनिरित्येके ।
 VJYP - यद्वैजयन्ती- “अयोनिर्मुसलोऽस्त्री स्यात्” इति ॥
213. ABCH - अथाऽन्तिका । चुल्ल्यऽश्मन्तकमुद्धानं स्याद् अधिश्रयणी च सा ॥१०१८॥
 ACVH - अन्तोऽस्त्यऽस्यामन्तिका ।
 MALA - “अन्त्यधिश्रयणी भवेत्” इति तु माला ॥
214. ABCH - मणिकोऽलिञ्जरः । (१०२२)
 SHAS - यत् शाश्वतः- “अलिञ्जरं मणिं विद्याद् मेहनावययो मणिः” ।

215. ABCH - दृतिः खल्लः (१०२५)

ACVH - पुंलिङ्गोऽयम् ।

VJYP - वैजयन्ती तु- “खल्लं क्लीबम्” इत्याह ॥

216. ABCH - सर्वमावपनं भाण्डं पात्राऽमत्रे तु भाजनम् । (१०२६)

ABCH treats them as two synsets सर्वमावपनं भाण्डं and पात्रामत्रे तु भाजनम्.

AMAR - अमरस्तु-आवपनादीन् पञ्चाऽप्येकार्थानाह ॥

217. ABCH - लोकालोकश्चक्रवालः (१०३१)

Kālidāsa (KLDS) - यत्कालिदासः- “प्रकाशश्चान्धकारश्च लोकालोक इवाऽचलः” इति ॥

218. ABCH - गिरिसारं शिलासारं तीक्ष्ण-कृष्णामिषे अयः । (१०३८)

SMRT - यत्स्मृतिः- “अश्यभ्यो लोहमुत्थितम्” ।

ACVH - एतौ क्लीबे ।

AMAR - पुंस्यपि इत्यमरः ॥

219. ABCH - सर्वं च तैजसं लोहं (१०३९)

Unknown - यदाहुः-

“सुवर्णं रजतं ताम्रं रीतिः कांस्यं तथा त्रपु ।

सीसं च धीवरं चैव अष्टौ लोहानि चक्षते” ॥

220. ABCH - विकारस्त्वयसः कुशी । (१०३९)

SHAS - यत्-शाश्वतः- “कुशो दर्भः कुशी फालः” ।

221. ABCH - ताम्रं म्लेच्छमुखं शुल्वं रक्तं द्व्यष्टमुदुम्बरम् ॥१०३९॥

NRKT - ‘उद्गताम्बरम्’ इति नैरुक्ताः ।

Unknown - ओदुम्बरमित्यन्ये ॥

222. ABCH - सीसं तु सीसपत्रकम् ॥१०४०॥

ACVH - क्लीबलिङ्गोऽयम् ।

VJYP - वैजयन्ती तु- “सीसोऽस्त्री” इति पुंस्यप्याह ॥

223. सुवर्ण पुनः ।

स्वर्णं हेम हिरण्यहाटकवसून्पाषाणपदं काञ्चनं

कल्याणं कनकं महारजतरैर्गाङ्गेयरुक्माण्यपि

कलधौतलौहोत्तमवह्निबीजान्यपि

गारुडं

गैरिकजातरूपे

।

तपनीयचामीकरचन्द्रभर्माऽर्जुननिष्ककार्तस्वरकर्बुराणि ॥१०४४॥

जाम्बूनदं शातकुम्भं रजतं भूरि भूतमम् । (१०४५)

ACVH - हटति दीप्यते हाटकम्, हाटकाकरभवत्वाद् वा, क्लीबलिङ्गोऽयम्, पुंस्यपि ।

VCSP - वाचस्पतिः यदाह- “हाटकमस्त्रियाम्, हाटकाकरसंभूतं पलाशकुसुमच्छविः” ॥

Unknown - यदाह:-“अग्रेरपत्यं प्रथमं सुवर्णम्” इति ॥

Purāṇa - यत्पुराणम्- तीरमृत् तद्रसं प्राप्य मुखवायुविशोषिता । जाम्बूनदाख्यं भवति सुवर्णं सिद्धभूषणम् ॥

ACVH - भवत्याकरे भूरि, “भूसूकुशि-” ॥ (उणा-६९३) ॥ इति किद् रिः, क्लीबलिङ्गोऽयम् ।

VCSP - वाचस्पतिस्तु- “भूरिरष्टापदोऽस्त्रियाम्” इति पुंस्यप्याह ॥

224. ABCH - रजतं च सुवर्णं च संश्लिष्टे घनगोलकः । (१०४७)

VCSP - यद् वाचस्पतिः- “तारहेम्नी तु संश्लिष्टे घनगोलकमस्त्रियाम्” इति ॥

225. ABCH - हंसपादकुरुविन्दे तु हिङ्गुलः ॥१०६१॥

ACVH - पुंस्ययम् ।

VCSP - वाचस्पतिस्तु-“हिङ्गुलस्त्वस्त्रियाम्” इति क्लीबेऽप्याह ।

226. ABCH - रत्नं वसु मणिः (१०६३)

VCSP - यद् वाचस्पतिः- “हीरकं मौक्तिकं स्वर्णं रजतं चन्दनानि च । शङ्खश्चर्म च वस्त्रं चेत्यष्टौ रत्नस्य जातयः” ॥

227. ABCH - नीलमणिस्त्विन्द्रनीलः (१०६५)

ACVH - पुंक्लीबलिङ्गः ।

VJYP - “इन्द्रनीलं महानीलम्” इति क्लीबे वैजयन्तिः ।

228. ABCH - सूचीमुखं तु हीरकः । (१०६५)
ACVH - हीरकः पुंलिङ्गः ।
VCSP - वाचस्पतिस्तु- “अथाऽस्त्री हीरकः” इति क्लीबेऽप्याह ।
229. ABCH - क्षीरतैलस्फटिकाभ्यामन्यौ खस्फटिकाविमौ । (१०६८)
ACVH - सूर्यकान्त-चन्द्रकान्तौ द्वौ अपि आकाशस्फटिकाख्यौ इत्यर्थः ।
VCSP - यद् वाचस्पतिः-
“स्फटिकास्तु त्रयस्तेषामाकाशस्फटिको वरः ।
द्वौ क्षीरतैलस्फटिकावाकाशस्फटिकस्य तु ।
द्वौ भेदौ सूर्यकान्तश्च चन्द्रकान्तश्च तत्र च” ॥
230. ABCH - शुक्तिजं मौक्तिकं मुक्ता मुक्ताफलं रसोद्भवम् ॥१०६८॥
ACVH - शुक्तिरूपलक्षणं हस्तिमस्तकाद्युद्भवमपि ।
Unknown - यदाह-
“हस्तिमस्तकदन्तौ तु दंष्ट्रा शुनवरायोः ।
मेघो भुजङ्गमो वेणुर्मत्स्यो मौक्तिकयोनयः” ॥
231. ABCH - नीरं वारि जलं दकं कमुदकं पानीयमभ्रः कुशं
तोयंजीवनजीवनीयसलिलार्णस्यम्बुवाःसंवरम् ।
क्षीरं पुष्करमेघपुष्पकमलान्यापः पयः पाथसी
कीलालं भुवनं वनं घनरसो यादो निवासोऽमृतम् ॥१०६९॥
कुलीनसं कबन्धं च प्राणदं सर्वतोमुखम् । (१०७०)
ACVH - घनस्य रसो घनरसः ।
MALA - ‘घनरसम्’ इति क्लीबे माला ॥
ACVH - कं सुखं वायुं वह्निं वा बध्नाति कबन्धम् ।
Unknown - ‘कम्, अन्धम्’ इति द्वे नाम्नी इत्येके ॥
232. ABCH - अस्थाघास्थागमस्ताघमगाधं चातलस्पृशि ॥ १०७० ॥

निम्नं गभीरं गम्भीरम् (१०७१)

ACVH - अस्थाघादिन् एकार्थान् इत्येके ॥

233. ABCH - द्वीपान्तरा असङ्ख्यास्ते (१०७४)

ACVH - यदाहुः - "असङ्ख्याता द्वीपसमुद्राः" इति ॥

234. ABCH - लवणक्षीरदध्याज्यसुरेक्षुस्वादुवारयः । (१०७५)

ACVH - यदाहुः - "लावणो रसमयः सुरोदकः सार्षिषो दधिजलः पयःपयाः ।

स्वादुवारिरुदधिश्च सप्तमस्तान् परीत्य त इमे व्यवस्थिताः ॥" इति ।

235. ABCH - कूपकास्तु विदारकाः ॥ १०८८ ॥

KATY - यत् कात्यः - विवरान् कूपकानाहुः ॥

236. ABCH - सिकता वालुकाः (१०८९)

ACVH - 'सिकिः सौत्रः' सिक्यन्ते सिकताः "पृषिरञ्जि-" ॥ (उणा-२०८) ॥ । इति किदतः, स्त्रीलिङ्गो वा बहुवचनान्तोऽयम् ।

HEMA - यदवोचाम-"वां तु जलौकाप्पसरसः सिकतासुमनः समाः" (लिङ्गा०२६) इति ॥

237. ABCH - नान्दीमुखो नान्दीपटो वीनाहो मुखबन्धने । (१०९२)

ACVH - वीनाहः पुंक्लीबलिङ्गः ।

VJYP - यद्वैजयन्ती - 'वीनाहमस्त्रियाम्' ।

238. ABCH - पद्माकरस्तडागः स्यात् कासारः सरसी सरः ॥१०९४॥

VCSP - वाचस्पतिस्तु- 'कासारस्त्वल्पसरसि सरसी तु महासरः' इति विशेषमाह ॥

239. ABCH - स्यादालवालमावालमावापः स्थानकं च सः ॥ १०९५ ॥

Unknown - आलवालं विदुर्धिरा धारणं द्रवतोऽम्भसः ।

240. ABCH - निर्झरस्तु झरः सरिः । उत्सः स्रवः प्रस्रवणं (१०९६)

ACVH - उत्सः "ऋजिरिषि-" ॥ (उणा-५६७) ॥ इति कित् सः, पुंस्ययम् ।

VCSP - वाचस्पतिस्तु 'सरिः स्त्री निझरो झरः उत्सरस्वस्त्री' इति क्लीबेऽप्याह ॥

241. ABCH - वह्निर्बृहद्भानु-हिरण्यरेतसौ (१०९७)

ACVH - हिरण्यं रेतोऽस्य हिरण्यरेताः ।

SMRT - यत् स्मृतिः- “अग्रेरपत्यं प्रथमं सुर्वणम्” इति ॥

242. ABCH – कृपीटयोनिर्दमुना विरोचनाऽऽशुशुक्षणी छागरथस्तनूनपात् ॥ १०९७ ॥

ACVH - कृपीटस्याम्भसो योनिर्धूमजत्वाद् मेघानां, कृपीटयोनिः ।

Unknown - ‘कृपीटं योनिरस्य’ इत्येके

Unknown - यदाहुः- “अद्भ्योऽग्निर्ब्रह्मणः क्षत्रमश्मभ्यो लोहमुत्थितम्” ॥

243. ABCH - अर्चिष्मत्-शमीगर्भतमोघ्नशुक्राः ॥ १०९८ ॥

ACVH - शम्या गर्भः शमीगर्भः ।

KLDS - यत् कालिदासः-“शमीमिवाभ्यन्तरलीनपावकाम्” इति ॥

244. ABCH - शोचिष्केशः शुचि-हुतवहो-षर्बुधाः सप्तमन्त्र-

ज्वालाजिह्वो ज्वलन-शिखिनौ जागृविर्जातवेदाः ।

बर्हिःशुष्मा-ऽनिलसख-वसू रोहिताश्वा-ऽऽश्रयाशौ

बर्हिर्ज्योतिर्दहन-बहुलौ हव्यवाहोऽनलोऽग्निः ॥ १०९९ ॥

ACVH - सप्त-मन्त्र-ज्वालशब्देभ्यः परो जिह्वः सप्तजिह्वा अस्य सप्तजिह्वः ।

Unknown - यदाहुः-

“भवति हिरण्या कनका रक्ताकृष्णा वसुप्रभा कन्या ।

रक्ता बहुरुपेति सप्त सप्तर्चिषां जिह्वाः” ॥ १ ॥

ACVH - बर्हिर्दर्भः शुष्म बलमस्य बर्हिःशुष्मा । बर्हते वर्धते बर्हिरिति व्यस्तम्, यथा ‘बर्हिर्मुखा देवाः’

शुष्यत्यनेन शुष्मा ।

Amaramālā - यथा- ‘शुष्मणि प्रणयनाभिसंस्कृते’ इति बर्हिरुक् इत्यमरमाला । यथा ‘दशां बर्हिणो

बर्हिरुक्कः’ इति ॥

ACVH - आश्रयमात्मस्थानमश्नाति आश्रयाशः ।

Unknown - आश्रयमाधारमश्नात्याश्रयाश इत्येके ॥

245. ABCH - तडिदैरावती विद्युच्चला शम्पाऽचिरप्रभा ॥ ११०४ ॥

ACVH - शाम्यति शम्पा “भापाचणि-” ॥ (उणा-२९६) ॥ इति पः, शं पिबतीति वा ।

Prācyā - दन्त्यादिरयमिति प्राच्याः सं पिबति नायनं तेज इति व्याख्यन् ॥

246. ABCH - स्यात् प्ररोहोऽङ्कुरोऽङ्कुरो रोहश्च (१११८)

ACVH - बाहुलकाद् दीर्घत्वेऽङ्कुरः, पुंलिङ्गावेतौ ।

VJYP - क्लीबेऽपि वैजयन्ती, यदाह- ‘स्यादङ्कुरोऽङ्कुरोऽस्त्रियौ’ इति ॥

247. ABCH - प्रसूनं कुसुमं सुमम् ॥११२४॥

ACVH - पुष्प्यति पुष्पं पुंक्लीबलिङ्गः ।

VCSP - वाचस्पतिस्तु ‘पुष्पोऽस्त्री’ इति पुंस्यप्याह ॥

248. ABCH - जालकक्षारकौ तुल्यौ कलिकायां तु कोरकः । ११२५ ॥ कड्मले मुकुलं (११२६)

Hṛdya - हृद्यास्तु-अवान्तरभेदं न मन्यन्ते यदाहुः, “मुकुलाख्या तु कलिका कुड्मलं जालकं तथा । क्षारकं कोरकं च” इति ।

249. ABCH - गुच्छे गुच्छ-स्तबक गुत्सकाः । गुलुञ्छः (११२६)

ACVH - गुड्यते गुलुञ्छः “गुलुञ्छपिलिपिञ्छैधिच्छ-” ॥ (उणा-१२६) ॥ इति छे निपात्यते, पुंस्ययम् ।

VCSP - क्लीबेऽपि वाचस्पतिः, यदाह-‘गुलुञ्छोऽस्त्री’ इति ।

250. ABCH - वृन्तं प्रसवबन्धनम् ।

ACVH - प्रसवः पुष्पादिस्तस्य बन्धनम् ।

KATY - यत् कात्यः ‘बन्धनं पुष्पफलयोर्वृन्तमाहुः’ ॥

251. ABCH - प्लक्षस्तु पर्कटी जटी ॥११३१॥

ACVH - पर्कटाख्यं फलमस्त्यस्य पर्कटी ।

VCSP - यद् वाचस्पतिः- ‘फलं त्वेतस्य पर्कटम्’ इति ।

ACVH - पृच्छयते पर्कटीत्यकारान्तात् ड्यामीदन्तो वा ।

SHAS - यत् शाश्वतः-‘विज्ञेया पर्कटी प्लक्षः प्लक्षः पिप्पलपादपः’ इति ॥

252. ABCH - न्यग्रोधस्तु बहुपात् स्याद् वटो वैश्रवणालयः । (११३२)

ACVH - वैश्रवणस्य यक्षस्यालयो वैश्रवणालयः ।

Unknown - यदाहुः-‘वटे वटे वैश्रवणः’ इति ॥

253. ABCH - अशोकः कङ्कलिः (११३५)

ACVH - कङ्कलिः स्त्रीलिङ्गः ।

Amaraśeṣa - यदमरशेषः - ‘स्त्रियां त्वशोके कङ्कलिः’ ।

254. ABCH - वृषो वासाऽऽटरुषके (११४०)

ACVH - वाति गच्छति श्लेष्माऽनया वासा “मावावदि-” ॥ (उणा- ५६४) ॥ इति सः, दन्त्यान्तोऽयम् ।

AMAR - ‘उष्यते काम्यते वाशेति तालव्यान्तः’ इत्यमरः ॥

255. ABCH - स्नुहिर्वज्रो महातरुः ॥११४०॥

VAID - स्नुहेति वैद्याः ।

256. ABCH - निचुले हिज्जलेज्जलौ । (११४५)

KSRS - ‘हितजलापभ्रंशो हिज्जलः’ इति क्षीरस्वामी, जलवेतसविशेषत्वादस्य ॥

257. ABCH - कलिरक्षो बिभीतकः ॥११४५॥

Unknown - बिभेदक इत्यन्ये ॥

258. ABCH - केतकः क्रकचच्छदः ।

ACVH - केतति वने केतकः पुंस्त्रीलिङ्गः ।

VCSP - यद् वाचस्पतिः ‘केतकस्तु द्वयोः’ इति ॥

259. ABCH - धान्यं तु सस्यं सीत्यं च व्रीहिः स्तम्बकरिश्च तत् । (११६८)

Unknown - यदाहुः-

“व्रीहिर्यवो मसूरो गोधूमा मुद्गमाषतिलवणकाः ।

अणवः प्रियङ्गुः कोद्रवमयुच्छकाः शालिराढक्यः ॥१॥

किञ्च कलाय-कुलत्थौ शणसप्तदशानि धान्यानि ॥”

260. ABCH - कुल्मासस्तु यावकः ॥११७५॥

Unknown - यौत्यम्भसा यावकः धान्यविशेषः इत्यन्ये ॥

261. ABCH - धान्यमावसितं रिद्धं (११८३)

Unknown - 'रक्षार्थमाच्छादितमावसितम्' इत्यन्ये ॥२॥

ACVH - राध्यति स्म रिद्धं सिद्धमित्यर्थः ।

Unknown - 'सुसम्पन्नम्' इत्येके ।

262. ABCH - मूलपत्रकरीराग्रफलकाण्डाविरुढकाः ॥११८३॥

GAUD - 'अविरुढकं तालास्थिमज्जा' इति गौडाः ॥

263. ABCH - भेदाः कान्तार-पुण्ड्राद्यास्तस्य (११९४)

VCSP - आदिशब्दात् कोषकाराद्याः,

यद् वाचस्पतिः-

"पुण्ड्रे पुण्ड्रकः सेव्यः पौण्ड्रकोऽतिरसो मधुः ।

श्वेतकाण्डो भीरुकस्तु हरितो मधुरो महान् ॥१॥

शून्येश्वरस्तु कान्तारः कोषकारस्तु वंशकः ।

शतघोरस्त्वषट्क्षारः पीतच्छायोऽथ तापसः ॥२॥

सितनीलोऽथ नेपालो वंशप्रायो महाबलः ।

अन्वर्थस्तु दीर्घपत्रो दीर्घपर्वा कषायवान् ॥३॥

काष्ठेशुस्तु ह्रस्वकाण्डो घनग्रन्थिर्वनोद्भवः ।

नीलघोरस्तु सुरसो नीलपीतलराजिवान् ॥४॥

अनूपसंभवः प्रायः खर्नटी त्विक्षुबालिका ।

करङ्कशालिः शाकेशुः सूचिपत्रो गुडेक्षवः ॥५॥"

इति ॥

264. ABCH - घासस्तु यवसं (११९५)

ACVH - यूयते यवसं "वहियुभ्यां वा" ॥ (उणा-५७१) ॥ इत्यसः, क्लीबलिङ्गोऽयम् ।

VCSP - वाचस्पतिस्तु- 'अथाऽस्त्री यवसो घासः' इति पुंस्यप्याह ॥

265. ABCH - तृणमर्जुनम् । (११९५)

BHAG - 'सर्वं च तृणमर्जुनम्' इति भागुरिः ॥

266. ABCH - विषः क्ष्वेडो रसस्तीक्ष्णं गरलः (११९५)
 ACVH - वेवेष्टि विषः पुंक्लीबलिङ्गः, स च द्वेधा, स्थावरजंगमभेदात् ।
 Unknown - यदाह-
 “कन्दजं कालकूटादि पुष्पजं दालवादिकम् ।
 फालमङ्गोल्लसारादि दंष्ट्रादिजं तु जंगमम् ॥१॥” इति ॥
267. ABCH - वत्सनाभः कालकूटो ब्रह्मपुत्रः प्रदीपनः । (११९६)
 ACVH - ब्रह्मणः पुत्रो ब्रह्मपुत्रः ।
 YJVL - यद् याज्ञवल्क्यः-‘त्वं विष ! ब्रह्मणः पुत्रः सत्ये धर्मे व्यवस्थितः’ ॥
268. ABCH - अहिच्छत्रो मेषशृङ्गः कुष्ठ-चालूक नन्दनाः । (११९७)
 ACVH - मेषशृङ्गाकारत्वाद् मेषशृङ्गः ।
 VCSP - यद् वाचस्पतिः- मेषशृङ्गं त्वविशृङ्गाकृतिः ॥
269. ABCH - कैराटको हैमवतो मर्कटः करवीरकः ॥११९७॥
 ACVH - करवीरमूल-जत्वात् करवीरः ।
 Unknown - यदाह-‘करवीरं त्वश्वसारमूलजं सर्वतोभवम्’ ॥
270. ABCH - अङ्गोलसारः कालिङ्गः शृङ्गिको मधुसिक्थकः ॥ ११९८ ॥
 ACVH - शृङ्गमस्त्यस्य शृङ्गिकः, अजस्तनाकृतित्वात् ।
 Unknown - यदाह-‘शृङ्गिकं तु कृष्णवर्णमजस्तनसमाकृतिः’ ॥
271. ABCH (1196-1199)
 ACVH - एते सर्वेऽपि स्थावरवनस्पतिभवत्वात् स्थावरविषस्य जातयो भेदा इति ।
 VCSP - ‘एते सर्वेऽपि पुंक्लीबलिङ्गाः’ इति वाचस्पतिः ।
272. ABCH - अस्रपा जलौकसः ॥ १२०३ ॥
 ACVH - अस्रं पिबत्यस्रपा ।
 Unknown - ‘विचका’ इत्येके ।
273. ABCH - माक्षिकादि तु मधु स्यात् ।

ACVH - मक्षिकाभिः कृतं माक्षिकं “नाम्नि मक्षिकादिभ्यः” ॥६॥ ३ ॥१९३॥ इत्यण्, आदिग्रहणात् पौत्तिकादि ।

VCSP - यद् वाचस्पतिः-

“पौत्तिक-भ्रामर-क्षौद्र-दालौ-द्वालक-माक्षिकम् ।

अर्घ्यं छात्रकमित्यष्टौ जातयोऽस्य पृथग्गुणाः ॥१॥

तत्र पौत्तिकमुत्तप्तघृताभं विषकीटजम् ।

भ्रामरं तु भ्रमरजं पाण्डुरं गुरुशीतलम् ॥२॥

क्षौद्रं तु कपिलं दाहि क्षुद्रानीतं मलावहम् ।

लालं तु दलजं सेव्यं दुर्लभं रुक्षवालकम् ॥३॥

उद्वलकं तु शालाकं विषजिद् मधुराम्लकम् ।

माक्षिकं तु मधु ज्येष्ठं विरुक्षं तैलवर्णकम् ॥४॥

अर्घ्यं तु पूज्यमापाण्डु मनाक् तित्तं सवालकम् ।

छात्रं त्वेकान्तमधुरं सर्वार्घ्यं राजसेवितम् ॥५॥”

इति ॥

ACVH - मन्यतेऽभिलष्यते मधु क्लीबलिङ्गः ।

VJYP - पुंस्यपि वैजयन्ती, यदाह-‘मधुलं तु मधुनाऽस्त्री मधुकं तत्पुरातनम्’ इति ॥

274. ABCH - गम्भीरवेद्यवमताङ्कुशः ।

ACVH - गाम्भीर्येण वेदयति गम्भीरवेदी ।

Unknown - यदाहुः-

“त्वग्भेदात् शोणितश्रावादामांसं व्यथनादपि ।

संज्ञां न लभते यस्तु गजो गम्भीरवेद्यसौ ॥”

275. ABCH - कुम्भयोरन्तरं विदुः ॥ १२२६ ॥

Pālakāpya - यत् पालकाप्यः-

“तत्रारक्षविताने द्वे विदू द्वौ श्रवणे गतौ ।

प्राक् पश्चाच्च तिर्यक् च षड्भेदाङ्कुशवारणा ॥”

276. ABCH - वडवाऽश्वा प्रसूर्वामी (१२३३)

ACVH - ‘वड आग्रहणे’ सौत्रः, वडति गर्भं वडवा “वडिवटि-” ॥ (उणा--५१५) ॥ इत्यवः ।

NRKT - ‘वडं वलं हयं वा वातीति वा, अश्वोऽस्यां वेति नैरुक्ताः ॥

277. ABCH - प्रक्षरं प्रखरः समौ ॥१२५१॥

ACVH - प्रावृतः खरत्यनेन प्रखरः “जठर-” ॥ (उणा-४०३) ॥ इत्यरे निपात्यते, प्रखरेति कठोरो वा, आयुधैरभेद्यत्वात् पुंस्ययम् ।

VJYP - क्लीबेऽपि वैजयन्ती, यदाह-प्रक्षरं प्रखरो-ऽस्त्रियाम् ॥

278. ABCH - षण्डतोचित आर्षभ्यः (१२५९)

Unknown - षण्डतायोग्य इत्येके ॥

279. ABCH - षष्ठवाट् तु स्याद् युगपार्श्वगः ॥१२६०॥

Unknown - ‘दमनाय योजितः’ इत्येके ॥

280. ABCH - वृषाक्रान्ता तु सन्धिनी । (१२६७)

KATY - ‘अदुग्धा दोहकाले तु सन्धिनी’ इति कात्यः ।

SHAS - ‘सन्धिन्यकालदुग्धा गौर्वृषा- क्रान्ता च सन्धिनी’ इति तु शाश्वतः ॥

281. ABCH - अजःस्याच्छगलः छागश्छगोवस्तःस्तभः पशुः । (१२७५)

AMAR - ‘तुभः’ इत्यमरः ॥

282. ABCH - विश्वकद्रुस्तु कुशलो मृगव्ये (१२८१)

Unknown - विगतं श्वकदनमस्येत्येके ॥

283. ABCH - सृगालो जम्बुकः फेरु फेरण्डः फेरवः शिवा ॥१२८९॥

NRKT - शेते दिवा शिवा ‘शकुनावेदनी’ इति नैरुक्ताः ।

ACVH - शृगाऽप्ययं स्त्रीलिङ्गः ।

SHAS - यत् शाश्वतः-‘शिवः कीलः शिवा क्रोष्टा भवेदामलकी शिवा’ ॥

284. ABCH - मृगः कुरङ्गः सारङ्गो वातायु-हरिणावपि । (१२९३)

ACVH - वातमेति वातायुः “कृवापाजि-” ॥ (उणा-१) ॥ इत्युण् ।

Unknown - ‘वातं शुष्कमयते वातायुः’ इत्यन्ये ॥

285. ABCH - कदली कन्दली कृष्णशारः पृषत-रोहितौ ॥१२९४॥

ACVH - कद्यते कदली “मृदिकन्दि-” ॥ (उणा-४६५) ॥ । इत्यलः, स्त्रियामयम् ।

Unknown - यदाह- “कदली तु बिले शेते मृदुभक्षैव कर्बुरैः ।

नीलाग्रै रोमभिर्युक्ता सा विंशत्यङ्गुलायता ॥१॥”

ABCH - कन्दान् लाति कन्दली श्यामवर्णा हस्तायामा महोदरी, त्रिलिङ्गोऽयम् ।

AMAR - अमरस्त्वेतौ मत्वर्थीयेनन्तावाह ॥

286. ABCH - बिडाल ओतुर्माजरो ह्रीकुश्च वृषदंशकः ॥१३०१॥

Unknown - ‘ह्रीकुर्वनमाजरोः’ इत्यन्ये ॥

287. ABCH - शेषो नागाधिपोऽनन्तो द्विसहस्राक्ष आलुकः ॥१३०७॥

Unknown - ‘शेतेऽस्मिन् हरिः’ इत्यन्ये ॥

288. ABCH - इत्यादयोऽपरे नागास्तत्तत्कुलसमुद्भवाः ॥१३११॥

Unknown - आदिग्रहणाद् महानीलादयः; यदाह-

“महानीलः कुरहश्च पुष्पदन्तश्च दुर्मुखः ।

कपिलो वामनः शङ्कुरोमा च करवीरकः ॥१॥

एलापत्रः शुक्तिकर्ण-हस्तिभद्र-धनञ्जयाः ।

दधिमुखः समानासोत्तंसको दधिपूरणः ॥२॥

हारिद्रको दधिकर्णो मणिः शृङ्गारपिण्डकः ।

कालियः शङ्खकूटश्च चित्रकः शङ्खचूडकः ॥३॥

इत्यादयोऽपरे नागास्तत्तत् कुलप्रसूतयः ॥” इति ॥

289. ABCH - दंष्ट्राशीः (१३१५)

Unknown - यदाह-“आशीस्तालुगता दंष्ट्रा यया विद्धो न जीवति” ।

Unknown - ‘ईकारान्तोऽयम्’ इत्यन्ये ॥

290. ABCH - दर्वी भोगः फटः स्फटः । फणः (१३१५)

Cāndra - फणं नपुंसकं सस्मरुश्चान्द्राः ॥

291. ABCH - पेशीकोशोऽण्डे (१३१९)

ACVH - पेशीनां मांसखण्डानां कोशः ॥१॥ अमति निःसरत्यस्मादण्डं पुंक्लीबलिङ्गः “पञ्चमाद्-” ॥

(उणा-१६८) ॥ इति डः तत्र ।

GAUD - तिस्रः संज्ञा इति गौडः ॥

292. ABCH - सारसी लक्ष्मणा (१३२८)

ACVH - 'लक्ष्मणी' इत्यन्ये ।

293. ABCH - चाषे किकीदिविः ।

Unknown - 'द्वे नाम्नी' इत्यन्ये । किकिदीविरपि ॥

294. ABCH - दात्यूहे कालकण्टकः ।

Unknown - 'कालः कण्ठोऽस्येति कालकण्ठकः' इत्येके ॥

295. ABCH - बलाहकः स्याद् बलाकः (१३३३)

ACVH - बलाकः, पुंस्ययम् ।

VCSP - यद् वाचस्पतिः- 'बलाकस्तु बलाहकः' ।

Liṅgaśāsanakṛt - 'पुंघ्वजेऽपि स्त्रीलिङ्ग एवायम्' इति लिङ्गानुशासनकृतः ॥

296. ABCH - बलाका विसकण्टिका । (१३३३)

Unknown - 'विसकण्टिका' इत्यन्ये ।

297. ABCH - ज्योत्स्नाप्रिये चलचञ्चु-चकोर-विषसूचकाः ॥१६३९॥

Unknown - यदाह- “चकोरस्य विरज्येते नयने विषदर्शनात्” इति ॥

298. ABCH - नलमीनश्चिलिचिमः (१३४६)

VAID - 'चिलिचीमस्त्रिदोषकृत्' इति तु वैद्याः ।

299. ABCH - कुलीरः कर्कटः पिङ्गचक्षुः पार्श्वोदरप्रियः ।

ACVH - कोलति कुलिरः पुंकलीबलिङ्गः "घसिवशि-" ॥ (उणा-४१९) ॥ इति किदीरः, कुलिनमीरयति वा जनक-भक्ष्यत्वात् ।

KAUT - यत् कौटिल्यः- 'कटकसधर्माणो हि राजपुत्रा जनकभक्ष्याः' इति ॥

300. ABCH - रन्ध्रं बिलं निर्व्यथनं कुहरं शुषिरं शुषिः ॥१३६३॥

ACVH - शुष्यत्यस्यां शुषिः "नाम्युपान्त्य-" ॥ (उणा-६०९) इति किदिः स्त्रीलिङ्गोऽयम् ।

VJYP - वैजयन्ती तु- 'श्वभ्रं नक्ली शुषिर्वपा' इति पुंस्यप्याह ।

Unknown - 'सुष्ठु स्यति सुषि दन्त्यादिः' इत्येके ॥

301. ABCH - अन्तःकरणं मानसं मनः । (१३६९)

Kāmandaki - "आत्मा मनश्च तद्विदैरन्तःकरणमुच्यते" इति कामन्दकिः ।

Unknown - मनो-बुद्धिरहंकार इत्यन्तःकरणं त्रिधेत्यन्ये ।

HEMA - यदवोचाम तर्के- "सर्वार्थग्रहणं मनः" ।

302. ABCH - हृच्चेतो हृदयं चित्तं स्वान्तं गूढपथोच्चले ॥१३६९॥

Unknown - विषयेष्वनाकुलं मनःस्वान्तमित्यन्ये ॥

303. ABCH - दोषे त्वादीनवास्त्रवौ ॥१३७५॥

AMAR - आदीनवास्त्रवौ क्लेश इत्यमरः ॥

304. ABCH - धर्मः पुण्यं वृषः श्रेयः सुकृते (१३७९)

HEMA - यदवोचाम- "दुर्गतिप्रपतत्प्राणिधारणाद्धर्म उच्यते" ॥

305. ABCH - बुद्धीन्द्रियं स्पर्शनादि (१३८५)

HEMA - यदवोचाम प्रमाणमीमांसायाम्- "स्पर्शरसगन्धरूपशब्दग्रहणलक्षणानि स्पर्शनरसनघ्राणचक्षुःश्रोत्राणीन्द्रियाणि द्रव्यभावभेदानि" इति ॥

306. ABCH - कृष्णस्तु मेचकः । (१३९७)

DURG - "मेचकः शिखिकण्ठाभः" इति दुर्गः ॥

307. ABCH - षड्ज-ऋषभ-गान्धारा मध्यमः पञ्चमस्तथा ।

धैवतो निषधः सप्त तन्त्रीकण्ठोद्भवाः स्वराः ॥१४०१॥

VYAD - षड्भ्यो जायते षड्जः, यद् व्याडिः-

"कण्ठादुत्तिष्ठते व्यक्तं षड्जः षड्भ्यस्तु जायते ।

कण्ठोरस्तालुनासाभ्यो जिह्वाया दशनादपि ॥१॥

ऋषभो गोरुतसंवादित्वात्, यदाह व्याडिः-

"वायुः समुत्थितो नाभेः कण्ठशीर्षसमाहतः ।

नर्दद्वृषभवद् यस्मात् तेनैष ऋषभः स्मृतः" ॥२॥

गां वाचं धारयति गान्धारः गन्धवहमियर्ति वा ।

यदाह-

वायुः समुत्थितो नाभेः कण्ठशीर्षसमाहतः ।

नानागन्धवहः पुण्यैर्गान्धारस्तेन हेतुना ॥३॥

मध्ये भवो मध्यमः, यदाह-

"तद्वदेवोत्थितो वायुरुरःकण्ठसमाहतः ।

नाभिप्राप्तो महानादो मध्यमस्तेन हेतुना" ॥४॥

पञ्चमस्थानभवत्वात् पञ्चमः, यदाह-

"वायुः समुत्थितो नाभेरुरोहृत्कण्ठमूर्धसु ।

विचरन् पञ्चमस्थानप्राप्त्या पञ्चम उच्यते ॥५॥

घिया वतो धीवतस्तस्याऽयं धैवतः,

दधाति संधत्ते स्वरानिति वा; यदाह-

अधिसंधीयते यस्मात् स्वरांस्तेनैष धैवतः" ।

"पुतपित्त-" ॥ (उणा-२०४) ॥ इति ते निपात्यते

॥६॥ निषीदन्ति स्वरा अत्र निषधो निषादाख्यः, यदाह-

"निषीदन्ति स्वरा अस्मिन् निषादस्तेन हेतुना" ।

308. ABCH - ते मन्द्रमध्यताराः स्युरुरःकण्ठशिरोभवाः । (१४०२)

Dantila - यद् दन्तिलः-

"नृणामुरसि मन्द्रस्तु द्वाविंशतिविधो ध्वनिः ।

स एव कण्ठे मध्यः स्यात् तारः शिरसि गीयते" ॥

309. ABCH - चोक्षं तु निःशोध्यमनवस्करम् ॥१४३६॥

निर्णिक्तं शोधितं मृष्टं धौतं क्षालितमित्यपि । (१४३७)

AMAR – अमरस्तु-"निर्णिक्तं शोधितं मृष्टं निःशोध्यमनवस्करम्" इत्येकार्थानाह ॥

310. ABCH - साधुरम्यमनोज्ञानि पेशलं हृद्यसुन्दरे । (१४४५)

NRKT - सुष्ठु नन्दयतीति नैरुक्ताः ॥

ACVH - लडहो देश्यः, संस्कृतेऽपि ।

GAUD - यद् गौडः- "मनोज्ञं मञ्जुलं मञ्जु लडहं रमणीयं च" इत्यादि ॥

311. ABCH - छदनव्यवधान्तद्धापिधानस्थगनानि च ॥१४७७॥

ACVH - छाद्यते छदनम्, छदेश्चुरादीनां प्रायो णिजन्तत्वाणिचोऽभावः ।

Unknown - छदिरदन्तः इति केचित् ॥

312. ABCH - मनितं प्रतिपन्नं च (१४९६)

Amaratīkā - 'मनूयी बोधने' इत्यस्येति मनितम्, कृति-वृति-नृतीना- मैदित्करणेन

इट्निषेधस्यानित्यत्वज्ञापनात् "वेतोऽपतः" ॥ ४ । ४ । ६२ ॥ इति इट्निषेधोऽत्र न भवतीत्यमरटीका ।

Unknown - "मानण् पूजायाम्" इत्यस्य स्थाने 'मनण्' इति केचित् पठन्ति, तस्य वा ॥

313. ABCH - ब्रज्याटाट्या पर्यटनं (१५०९)

ACVH - "आस्यटि-" ॥ ५ ।३ ।९७॥ इति क्यपि अट्यापि ।

MANU - यन्मनुः- 'तौर्यत्रिकं वृथाट्या च कामजो दशको गणः' ॥

314. ABCH - चर्या त्वीर्यापथस्थितिः ।

Unknown - यदाह- 'ईर्यापथो ध्यानमौनादिकं भिक्षुव्रतम्' ।

MUNI - 'चर्या त्वीर्या विदुर्बुधाः' इति तु मुनिः ॥

315. ABCH - कामं प्रकामं पर्याप्तं निकामेष्टे यथेप्सिते । (१५०५)

ACVH - एतेऽकारान्ता अनन्ययाः ।

SHAS - यत् शाश्वतः- कामे निकामे कामाख्या अव्ययास्तु मकारान्ताः ॥

316. ABCH - आकारस्त्विङ्ग इङ्गितम् ।

KAUT - आकारो मुखरागादिरिति कौटिल्यो व्याख्यत् ॥

317. साकं सत्रा समं सार्द्धममा सह (१५२७)

Dvyāśrayamahākāvya (DVYA) - यथाऽस्मदुपज्ञे द्रव्याश्रयमहाकाव्ये-

"पुलिनानि सह क्षोमैः सरांसि नभसा समम् ।

ज्योत्स्न्यऽमाऽह्नाऽमिषन्मेघाः साकं कैलाससानुभिः ॥

318. ABCH - कृतं त्वलम् ॥१५२७॥ भवत्वस्तु च किं तुल्याः (१५२८)

DVYA - यथा द्रव्याश्रये-

"गोत्रेण पुष्करावर्त्त ! किं त्वया गर्जितैः कृतम् ? ।

विद्युताऽलं भवत्वद्भिहंसा ऊचुर्बिलं घनम्" ॥

Bibliography

1. Agirre, Eneko, and Philip Edmonds. Word Sense Disambiguation: Algorithms and Applications. Vol. 33. Dordrecht: Springer, 2007.
2. Apte, Vaman Shivram. Sanskrit-Hindi Kośa. Delhi: Motilal Banarsidass, 2007.
3. Apte, Vaman Shivram. The Practical Sanskrit-English Dictionary, Containing Appendices on Sanskrit Prosody and Important Literary & Geographical Names in the Ancient History of India. Poona: Shiralkar & Co. Book-sellers, 1890.
4. Apte, Vāmana Śivarāma, Paraśurāma Kṛṣṇa Gode, Cintāmaṇa Gaṇeśa Karve, and Kaśinātha Vāsudeva Abhyāṅkara. Revised and Enlarged Edition of Prin. V. S. Apte's The Practical Sanskrit-English Dictionary. Poona: Prasad Prakashan, 1957.
5. Baker, Collin F., Charles J. Fillmore, and John B. Lowe. "The Berkeley FrameNet Project." In Proceedings of the 17th International Conference on Computational Linguistics (COLING 1998), Vol. 1. 1998.
6. Bär, Daniel, Torsten Zesch, and Iryna Gurevych. "Text Reuse Detection Using a Composition of Text Similarity Measures." In Proceedings of COLING 2012, 167–184. Mumbai, 2012.
7. Bär, Daniel, Torsten Zesch, and Iryna Gurevych. "Dkpro Similarity: An Open Source Framework for Text Similarity." In Proceedings of the 51st

- Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 121–126. 2013.
8. Bejcek, Eduard, Václava Kettnerová, and Markéta Lopatková. “Automatic Mapping Lexical Resources: A Lexical Unit as the Keystone.” In Proceedings of the International Conference on Language Resources and Evaluation (LREC), 2826–2832. 2014.
 9. Bevilacqua, Michele, Tommaso Pasini, Alessandro Raganato, and Roberto Navigli. “Recent Trends in Word Sense Disambiguation: A Survey.” In Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI), 4330–4338. International Joint Conference on Artificial Intelligence, Inc., 2021.
 10. Bhānujīdikṣita. The Nāmaliṅgānuśāsana (Amarakośa) of Amarasimha with the Commentary (Vyākhyāsudhā or Rāmāśramī) of Bhānujī Dīkṣita. Edited by Paṇḍita Śivadatta. Bombay: Nirṇaya-Sāgara Press, 1929.
 11. Bhattacharyya, Pushpak. “Indowordnet.” In Proceedings of the International Conference on Language Resources and Evaluation (LREC 2010). Malta, May 2010.
 12. Bonferroni, Cornelio E. “Teoria statistica delle classi e calcolo delle probabilità.” *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 8 (1936): 3–62.

13. CDSD. Cologne Digital Sanskrit Dictionaries. Cologne: Cologne University, 2025. Version 2.8.0. Accessed July 7, 2025. <https://www.sanskritlexicon.uni-koeln.de>.
14. Cochran, William G. "The Comparison of Percentages in Matched Samples." *Biometrika* 37, no. 3/4 (1950): 256–66.
15. Damerau, Fred J. "A Technique for Computer Detection and Correction of Spelling Errors." *Communications of the ACM* 7, no. 3 (1964): 171–176.
16. Devasāgara (AVRD). *Kalikālasarvajñaśrīhemacandrācāryaviracitā Vācakaśrīdevasāgaragaṇikṛtavvyutpattiratnākarakalitā Abhidhānacintāmaṇināmamālā*. Śrī Rāndera Roaḍa Jaina Saṅghaḥ. V.S. 2059.
17. Dietterich, Thomas G. "Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms." *Neural Computation* 10, no. 7 (1998): 1895–1923.
18. Gabrilovich, Evgeniy, and Shaul Markovitch. "Computing Semantic Relatedness Using Wikipedia-Based Explicit Semantic Analysis." In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Vol. 7, 1606–1611. 2007.
19. Goyal, Pawan, Gérard Huet, Amba Kulkarni, Peter Scharf, and Ralph Bunker. "A Distributed Platform for Sanskrit Processing." In *Proceedings of COLING 2012*, edited by Martin Kay and Christian

- Boitet, 1011–1028. Mumbai: The COLING 2012 Organising Committee, December 2012.
20. Halāyudha (ARMH). Halāyudhakośaḥ (Abhidhānaratnamālā). Edited by Jayaśaṅkara Jośī. Vārāṇasī: Sarasvatī Bhavana, V.S. 2014.
21. Hamming, Richard W. “Error Detecting and Error Correcting Codes.” The Bell System Technical Journal 29, no. 2 (1950): 147–160.
22. Hemacandra (ABCH). Abhidhānacintāmaṇi. Abhidhānacintāmaṇināmamālā (svopajña-sārthaśabdānukramaṇikābhyāṃ samalaṅkṛtā). Ahmedabad: Śrī Jaina Sāhitya Vardhaka Sabhā, Śrī Vṛddhi-nemi-amṛta Granthamālā, no. 72. V.S. 2032.
23. Hemacandra (ACPH). Abhidhānacintāmaṇipariśiṣṭa (The Abhidhāna-saṅgraha or A Collection of Sanskrit Ancient Lexicons). Book No. 7. Edited by Paṇḍit Śivadatta and Kāśīnātha Pāṇḍuraṅga Paraba. Bombay: Nirṇayasāgara Press.
24. Hemacandra (ACVH). Abhidhānacintāmaṇisvopajñavṛtti. Abhidhānacintāmaṇināmamālā (svopajña-sārthaśabdānukramaṇikābhyāṃ samalaṅkṛtā). Ahmedabad: Śrī Jaina Sāhitya Vardhaka Sabhā, Śrī Vṛddhi-nemi-amṛta Granthamālā, no. 72. V.S. 2032.
25. Huet, Gérard. “Sanskrit Lexicography, Past and Future.” In *Research on the Language and Script in Buddhist Sutras*, edited by Li Wei. Hangzhou: Hangzhou Buddhist Academy, 2019.

26. Jaro, Matthew A. "Advances in Record-Linkage Methodology as Applied to Matching the 1985 Census of Tampa, Florida." *Journal of the American Statistical Association* 84, no. 406 (1989): 414–420.
27. Jiang, Jay J., and David W. Conrath. "Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy." arXiv preprint [cmp-lg/9709008](https://arxiv.org/abs/cmp-lg/9709008), 1997.
28. Jinadevasūri (ACSJ). *Abhidhānacintāmaṇipariśiṣṭa* (The Abhidhāna-saṅgraha or A Collection of Sanskrit Ancient Lexicons). Book No. 11. Editor Paṇḍit Śivadatta and Kāśīnātha Pāṇḍuraṅga Paraba. Bombay: Nirṇayasāgara Press. 1896.
29. Johansson, Richard, and Luis Nieto Piña. "Embedding a Semantic Network in a Word Space." In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1428–1433. 2015.
30. Joshi, Salil, Arindam Chatterjee, Arun Karthikeyan Karra, and Pushpak Bhattacharyya. "Eating Your Own Cooking: Automatically Linking WordNet Synsets of Two Languages." In *Proceedings of COLING 2012: Demonstration Papers*, 239–246. 2012.
31. Keśavasvāmin (AKKS). *The Nāmaliṅgānuśāsanam* (Amarakosha of Amarasimha) With the Commentary (Amarakoshoddghāṭana) of Keśavasvāmin. Editor Krishnaji Govind Oka. Poona: Law Printing Press. 1913.

32. Kilgarriff, Adam. "Word Senses." In *Word Sense Disambiguation*, 29–46. 2006.
33. Kouvara, Eleni, Montserrat González, Julia Gross, and Roser Saurí. "Determining Differences of Granularity between Cross-Dictionary Linked Senses." In *Proceedings of the 19th EURALEX International Congress*, 2021.
34. Kuiper, Koenraad, Romain Fromont, and David Gerhard. "Polysemy and Word Frequency." *Journal of Research Design and Statistics in Linguistics and Communication Science* 4, no. 2 (2018): 144–155.
35. Kulkarni, Malhar, Chaitali Dangarikar, Irawati Kulkarni, Abhishek Nanda, and Pushpak Bhattacharyya. "Introducing Sanskrit WordNet." In *Proceedings of the 5th Global WordNet Conference (GWC 2010)*, 287–294. Mumbai: Narosa, 2010.
36. Landauer, Thomas K., Peter W. Foltz, and Darrell Laham. "An Introduction to Latent Semantic Analysis." *Discourse Processes* 25, no. 2–3 (1998): 259–284.
37. Levenshtein, Vladimir I. "Binary Codes Capable of Correcting Deletions, Insertions, and Reversals." *Soviet Physics Doklady* 10, no. 8 (1966): 707–710.
38. Maheśvara. *Amarakośa with Commentary of Maheśvara*. Pune: Varada Books. 1990.

39. Maheśvara. Śabdaprabheda. Edited by Vijayaśrīcandrasūri and Vinayasāgara. Surat: Rāndera Roḍa Jaina Saṅghaḥ. 2010.
40. Mallinātha. Amarakośa with the Unpublished South Indian Commentaries – Amarapadavivṛti of Liṅgayasūrin, Amarapadapārijāta of Mallinātha, and the Amarapadavivarāṇa of Appayārya. Vols. 1–2. Edited by A. A. Ramanathan. Adyar: The Adyar Library and Research Center. Adyar: 1971.
41. McNemar, Quinn. “Note on the Sampling Error of the Difference Between Correlated Proportions or Percentages.” *Psychometrika* 12, no. 2 (1947): 153–57.
42. “Dictionary.” Merriam-Webster.com. 2025. <https://www.merriam-webster.com> (accessed August 7, 2025).
43. “Thesaurus.” Merriam-Webster.com. 2025. <https://www.merriam-webster.com> (accessed August 7, 2025).
44. Miller, George A. “WordNet: A Lexical Database for English.” *Communications of the ACM* 38, no. 11 (1995): 39–41.
45. Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine J. Miller. “Introduction to WordNet: An On-Line Lexical Database.” *International Journal of Lexicography* 3, no. 4 (1990): 235–244.

46. Miller, George A., Claudia Leacock, Randee Teng, and Ross T. Bunker. "A Semantic Concordance." In Proceedings of the Human Language Technology Conference, 1993.
47. Monge, Alvaro, and Charles Elkan. "An Efficient Domain-Independent Algorithm for Detecting Approximately Duplicate Database Records." In Proceedings of the ACM-SIGMOD Workshop on Research Issues on Knowledge Discovery and Data Mining, 1997.
48. Nair, Sivaja S. The Knowledge Structure in Amarakośa. Hyderabad: University of Hyderabad, 2011.
49. Navigli, Roberto, and Simone Paolo Ponzetto. "BabelNet: The Automatic Construction, Evaluation and Application of a Wide-Coverage Multilingual Semantic Network." Artificial Intelligence (2012): 217–250.
50. Pradeep, Anagha, Radhika Mamidi, and Pavankumar Satuluri. "Context and WSD: Analysing Google Translate's Sanskrit to English Output of Bhagavadgītā Verses for Word Meaning." In Proceedings of the 7th International Sanskrit Computational Linguistics Symposium, 14–26. 2024.
51. Rāyamukuṭa (APCR). Padacandrikā – A Commentary on the Liṅgānuśāsanam of Amara. Edited by Kali Kumar Dutta. Calcutta: Sanskrit College, 1966.
52. Resnik, Philip. "Using Information Content to Evaluate Semantic Similarity in a Taxonomy." arXiv preprint cmp-lg/9511007, 1995.

53. Rijsbergen, CJ Van. *Information Retrieval*. 2nd ed. London: Butterworths, 1979.
54. Salgado, Ana, Sina Ahmadi, Alberto Simões, John McCrae, and Rute Costa. "Challenges of Word Sense Alignment." In *Proceedings of the LREC 2020 7th Workshop on Linked Data in Linguistics (LDL-2020)*, 45–51. Marseille: European Language Resources Association (ELRA), 2020.
55. Sarvānanda (TSS). Nāmaliṅgānuśāsanam śrīmadamarasiṃhaviracitaṃ Vandyaghaṭīyasarvānandapraṇītayā Tīkāsarvasvavyākhyayā sametaṃ. Edited by T. Ganpati Shastri. Anantaśayana Saṃskṛtagranthāvali, no. 38. Anantaśayana, 1914.
56. Saurí, Roser, Louis Mahon, Irene Russo, and Mironas Bitinis. "Cross-dictionary linking at sense level with a double-layer classifier." In *2nd Conference on Language, Data and Knowledge (LDK 2019)*, pp. 20-1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2019.
57. Schuler, Kipper Karin. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. PhD diss., University of Pennsylvania, 2005.
58. Scozzafava, Federico, Marco Maru, Fabrizio Brignone, Giovanni Torrisi, and Roberto Navigli. "Personalized PageRank with syntagmatic information for multilingual word sense disambiguation." In *Proceedings*

of the 58th annual meeting of the association for computational linguistics: system demonstrations, pp. 37-46. 2020.

59. Śaktidharaśarman (ARSS). Amarakośa Arthāt Nāmaliṅgānuśāsanam Nāma Koṣaḥ. Lucknow: Navalkishor Press Book Depot.
60. Śāśvata (ANKS). Śāśvatakośaḥ – The Anekārthasamucchaya of Śāśvata. Edited by Narayan Nathaji Kulkarni. Poona: Oriental Book Agency. 1929.
61. Speer, Robyn, Joshua Chin, and Catherine Havasi. “ConceptNet 5.5: An Open Multilingual Graph of General Knowledge.” In Proceedings of the AAAI Conference on Artificial Intelligence 31, no. 1. 2017.
62. Stokoe, Christopher. “Differentiating Homonymy and Polysemy in Information Retrieval.” In Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, 403–410. 2005.
63. Śrīvallabha (HSDS). Jinadevasūri’s Haimanāmamālāśiloṇcha with a Commentary by Śrīvallabha. L. D. Series 46. Ahmedabad: L. D. Institute of Indology.
64. Vācaspati (VCSP). Śabdārṇava of Vācaspati. In Journal of the American Oriental Society 85, no. 4 (1965): 524–543; and 88, no. 2 (1968): 345–347.

65. Viśnudattaśarman (ANVS). The Namalingānuśāsana viz Amarakosha of Amarasimha with the Commentary – Namacandrikā. Bombay: Shri Venkateshwar Steam Press, 1929.
66. Vogel, Claus. Indian Lexicography. Delhi: Motilal Banarsidass, 2015.
67. Vossen, Piek. "A multilingual database with lexical semantic networks." Dordrecht: Kluwer Academic Publishers. doi 10 (1998): 978-94.
68. Vyāḍi (VYDI). Fragments from Three Lost Kośas: II. Vyāḍi's Utpalinī. Edited by Robert Birwé. Journal of the American Oriental Society 87, no. 1 (1967): 39–52.
69. Webster, Noah. Webster's Unabridged Dictionary of the English Language. 1900. Kikwansha.
70. Wilson, H. H. A Dictionary in Sanscrit and English. 2nd ed. London: Parbury, Allen & Co., 1832.
71. Winkler, William E. "String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage." The Educational Resource Information Center (ERIC), 1990.
72. Wise, Michael J. "Yap3: Improved Detection of Similarities in Computer Program and Other Texts." In Proceedings of the Twenty-Seventh SIGCSE Technical Symposium on Computer Science Education, 130–134. 1996.

73. Yādavaprakāśa. The Vaijayantī of Yādavaprakāśa. Edited by Gustav Oppert. Madras: Madras Sanskrit and Vernacular Text Publication Society, 1893.
74. Yates, William. A Dictionary in Sanscrit and English, Designed for the Use of Private Students and of Indian Colleges and Schools. Calcutta: Baptist Mission Press, 1846.
75. Zipf, George K. Human Behaviour and the Principle of Least Effort. Cambridge, MA: Addison-Wesley, 1949.



SI No. 61119

Notification No. 1

हैदराबाद विश्वविद्यालय

UNIVERSITY OF HYDERABAD

(A Central University Established by an Act of Parliament)

P.O. CENTRAL UNIVERSITY, GACHIBOWLI, HYDERABAD - 500 046 (INDIA)

SEMESTER GRADE TRANSCRIPT**REGULAR EXAMINATION**

REG. NO. 21HSPH01

NAME OF THE STUDENT PATEL DHAVALKUMAR KIRITKUMAR

MONTH AND YEAR OF FEB 2022

SEMESTER 1

COURSE Ph.D. SANSKRIT STUDIES

PARENT'S NAME PATEL KIRITKUMAR JOITARAM / PATEL SUNANDABEN KIRITKUMAR

COURSE NO	TITLE OF THE COURSE	CREDITS	RESULTS
SK812	TOPICS IN NATURAL LANGUAGE PROCESSING	4	PASS
SK830	DISSERTATION RELATED READINGS	4	PASS

SEMESTER GRADE POINT AVERAGE (SGPA) :10.0

(In words) :TEN POINT ZERO

**DEPUTY REGISTRAR
(ACAD & EXAMS)**

Date Of Result Notification : Mar 1, 2022

Date Of Download : Jan 24, 2026

This is system generated certificate issued by O/o The Controller of Examinations, University of Hyderabad

Website: <https://www.uohyd.ac.in> and <https://www.acad.uohyd.ac.in>Phone : 040-23132120, E-mail: ce@uohyd.ernet.in and ceuh@uohyd.ac.in



SI No. 75137

Notification No. 2

हैदराबाद विश्वविद्यालय

UNIVERSITY OF HYDERABAD

(A Central University Established by an Act of Parliament)

P.O. CENTRAL UNIVERSITY, GACHIBOWLI, HYDERABAD - 500 046 (INDIA)

SEMESTER GRADE TRANSCRIPT**REGULAR EXAMINATION**

REG. NO. 21HSPH01

NAME OF THE STUDENT PATEL DHAVALKUMAR KIRITKUMAR

MONTH AND YEAR OF JUN 2022

SEMESTER 2

COURSE Ph.D. SANSKRIT STUDIES

PARENT'S NAME PATEL KIRITKUMAR JOITARAM / PATEL SUNANDABEN KIRITKUMAR

COURSE NO	TITLE OF THE COURSE	CREDITS	RESULTS
EN875	RESEARCH AND PUBLICATION ETHICS	2	PASS
SK801	RESEARCH METHODOLOGY	4	PASS

SEMESTER GRADE POINT AVERAGE (SGPA) :10.0

(In words) :TEN POINT ZERO

**DEPUTY REGISTRAR
(ACAD & EXAMS)**

Date Of Result Notification : Jul 28, 2022

Date Of Download : Jan 15, 2026

This is system generated certificate issued by O/o The Controller of Examinations, University of Hyderabad

Website: <https://www.uohyd.ac.in> and <https://www.acad.uohyd.ac.in>Phone : 040-23132120, E-mail: ce@uohyd.ernet.in and ceuh@uohyd.ac.in

UGC - CARE LISTED

ISSN: 0974 - 8946

अनुसन्धान-प्रकाशन-विभागीया त्रैमासिकी शोध-पत्रिका

शोध-प्रभा

(A Refereed & Peer-Reviewed Quarterly Research Journal)

47 वर्षे प्रथमोऽङ्कः (जनवरी-मार्च) 2022 ई.

प्रधानसम्पादकः

प्रो. मुरलीमनोहरपाठकः

कुलपतिः

सम्पादकः

प्रो. शिवशङ्करमिश्रः

सहसम्पादकः

डॉ. ज्ञानधरपाठकः



श्रीलालबहादुरशास्त्रीराष्ट्रीयसंस्कृतविश्वविद्यालयः

(केन्द्रीयविश्वविद्यालयः)

नवदेहली-16

विषयानुक्रमः

संस्कृतविभागः

1. दीक्षणीयेष्टिस्वरूपम् 1-8
-विद्यावाचस्पतिः डॉ. सुन्दरनारायणझाः
2. उत्सवप्रियाः खलु मनुष्याः 9-13
-डॉ. राधेश्यामगंगवारः
3. कर्मत्वविचारः 14-23
-डॉ. बद्रीनारायणगौतमः

हिन्दी विभाग

4. प्राचीन भारत में संघतन्त्र, राजतन्त्र तथा लोकतन्त्र 24-42
(संस्कृत शास्त्र परम्परा के सन्दर्भ में)
-डॉ० महेश दत्त शर्मा
5. विश्वकर्मा-सूक्त में ईश्वरास्तित्व की सिद्धि 43-54
(न्याय-दर्शन के परिप्रेक्ष्य में)
-डॉ. अनीता राजपाल
6. कालिदास के ग्रन्थों में विज्ञान एवं चिकित्सा शिक्षा के सन्दर्भ 55-60
-डॉ. अमित कुमार जायसवाल
7. वैदिक वाङ्मय में वास्तुशास्त्र 61-68
-श्री जीवन जोशी

English Section

8. Sāṅkhyatattvapradīpa of Tulasīrāma - 69-104
an unpublished Sāṅkhya work
-Dr. Dhaval Patel, IAS



Sāṅkhyatattvapradīpa of Tulasīrāma - an unpublished Sāṅkhya work

Dr. Dhaval Patel, IAS

Introduction

During my work of preparing a catalogue of the manuscripts in Chunilal Gandhi Vidyabhavan, Sarvajanika Education Society, Surat, I came across this rare manuscript. The present work is an unpublished work of Sāṅkhya philosophy. The name of the work is Sāṅkhyatattvapradīpa. Its author is one Tulasīrāma. This work is different from its namesake Sāṅkhyatattvapradīpikā of Keśava. My search in New Catalogus Catalogorum and National Manuscript Mission¹ database did not yield any other copy of the manuscript of the present work. So it seems improbable that there are any more copies of this work, at least in India. The only extant manuscript of the work is also the author's copy. It seems that the work did not get much fan following, and therefore there are no copies of this work available anywhere else. As this work is based on a single manuscript, there are places where it is absolutely impossible to find out the correct reading. This is a common problem with solitary manuscripts.

Manuscript details

The details of the manuscript² are as follows:

Manuscript collection - Śrī Śivadatta Śukla alias

Prakāśānandanātha collection

Accession Number - 1167

Title - Sāṅkhyatattvapradīpa

Author - Tulasīrāma

Size of folios - 11"x4.8"

Size of text - 9"x3.8"

Lines per page - 13

Letters per line - 38

ISCLS 2024

**Proceedings of the 7th International Sanskrit
Computational Linguistics Symposium**

15–17 February, 2024
Auroville,
Puducherry, India

Contents

Word Sense Alignment of Sanskrit Lexica	1
<i>Dhaval Patel, Amba Kulkarni</i>	
Context and WSD: Analysing Google Translate's Sanskrit to English Output of Bhagavadgītā Verses for Word Meaning	14
<i>Anagha Pradeep, Radhika Mamidi, Pavankumar Satuluri</i>	
Linguistically Mapping Aśoka: A Dialectometric Approach to the Major Rock and Major Pillar Edicts	27
<i>Patrick Zeithuber</i>	
Hoisting the colors of Sanskrit	39
<i>Gérard Huet</i>	
Using TEI for digital Sanskrit editions containing commentaries:	
A study of Kālidāsa's Raghuvamśa with Mallinātha's Sañjivani	52
<i>Tanuja P Ajotikar, Ketaki Kaduskar, Peter M Scharf</i>	
Inter Sentential Discourse Relations	67
<i>Saeed Vaze, Amba Kulkarni</i>	
A fast prakriyā generator	84
<i>Arun K Prasad</i>	
Anuprāsa Identifier and Classifier: A computational tool to analyze Sanskrit figure of sound	102
<i>Amruta Vilas Barbadikar, Amba Kulkarni</i>	
START: Sanskrit Teaching, Annotation, and Research Tool – Bridging Tradition and Technology in Scholarly Exploration	113
<i>Anil Kumar, Amba Kulkarni, Nakka Shailaj</i>	
The Śabdabrahman exercise platform	125
<i>Peter M Scharf, Harsha Pamidipalli</i>	

Word Sense Alignment of Sanskrit Lexica

Dhaval Patel

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

Amba Kulkarni

Department of Sanskrit Studies,
University of Hyderabad
ambakulkarni@uohyd.ac.in

Abstract

Word sense alignment is a field of study in which lexical resources or texts are aligned at the level of word sense rather than the word. The present paper tries to evaluate the possibility of mechanically aligning Sanskrit lexica at the level of word sense computationally.

1 Introduction

Sanskrit, an ancient Indian language, has been a medium of transmission of knowledge in various fields of study for centuries. Compilation of word lists in Sanskrit commenced at an early date as it was found necessary to access the old literature such as Vedic literature, while the language was undergoing some transformations with meaning shifts. The lexical resources known as kośas were developed. They are of two types – (1) Samānārthaka kośas and (2) Anekārthaka kośas. Samānārthaka kośas enlist the synonyms together. The synonyms are arranged following some theme, semantic criterion, or ontological classification scheme. For example, in the most famous samānārthaka kośa viz. Amarakośa, the words are arranged in three kāṇḍas and further within the kāṇḍas, the headwords are arranged based on either semantic or ontological properties. Anekārthaka kośas enlist different meanings of a given word. The words may or may not have any alphabetic arrangement. Both kinds of kośas were meant to be memorized, applied to texts, and cited as and when the usage of the said word in the literature was to be justified in a commentary. Therefore, the kośas were almost invariably in a verse form. Vogel (2015) has given a comprehensive coverage of these Sanskrit kośas and commentarial literature thereon.

Because of the influence of Western lexicography, a few Sanskrit-Sanskrit dictionaries like Vācaspatyam and Śabdakalpadrūma were also compiled on the lines with the Western methodology of arranging headwords alphabetically and in prose form. Several bilingual Sanskrit dictionaries such as Sanskrit-English, Sanskrit-French, and Sanskrit-German were created starting from the early 19th century. Almost all of the major dictionaries that are free from copyright are available on the Cologne Digital Sanskrit Dictionaries website (CDSD, 2023). This digitized data has various levels of markup. In the recent years Huet (2019) has developed a digital Sanskrit-French dictionary where the lexical items are directly linked to the inflectional and derivational morphology.

Some of these dictionaries, in addition to providing the meaning of Sanskrit words in the target language, also provide citations from Sanskrit texts. The citations in different dictionaries vary. These citations play an important role in understanding the context in which the sense is being used. Aligning the senses of different dictionaries would provide us with more than one example sentence for each sense to understand the context and the semantic criterion that decides the sense of the word in a given usage. Further, with the availability of word embeddings for words in several languages such as Hindi, English, French, German, etc. if the senses in Sanskrit bilingual dictionaries are aligned, one can take advantage of the

**The proceedings of the
Computational Sanskrit and Digital Humanities section**

under the
19th World Sanskrit Conference

Kathmandu, Nepal
June 26 – 30, 2025

Edited by
AMBA KULKARNI & OLIVER HELLWIG

Table of Contents

<i>An introduction to computational identification and classification of Upamā alaṅkāra</i> Bhakti Jadhav, Himanshu Dutta, Shruti Kanitkar, Malhar Kulkarni and Pushpak Bhattacharyya	1
<i>Aesthetics of Sanskrit Poetry from the Perspective of Computational Linguistics: A Case Study Analysis on Śikṣāṣṭaka</i> Jivnesh Sandhan, Amruta Barbadikar, Malay Maity, Pavankumar Satuluri, Tushar Sandhan, Ravi M. Gupta, Pawan Goyal and Laxmidhar Behera	15
<i>Itaretara Dvandva: A challenge for Dependency Tree semantics</i> Amba Kulkarni and Vasudha Neelamana	37
<i>A Case Study of Handwritten Text Recognition from Pre-Colonial era Sanskrit Manuscripts</i> Kartik Chincholikar, Shagun Dwivedi, Kaushik Gopalan and Tarinee Awasthi	52
<i>Towards Accent-Aware Vedic Sanskrit Optical Character Recognition Based on Transformer Models</i> Yuzuki Tsukagoshi, Ryo Kuroiwa and Ikki Ohmukai	70
<i>Vedavani: A Benchmark Corpus for ASR on Vedic Sanskrit Poetry</i> Sujeet Kumar, Pretam Ray, Abhinay Beerukuri and Shrey Kamoji	81
<i>Compound Type Identification in Sanskrit</i> Sriram Krishnan, Pavankumar Satuluri, Amruta Barbadikar, T S Prasanna Venkatesh and Amba Kulkarni	90
<i>IKML: A Markup Language for Collaborative Semantic Annotation of Indic Texts</i> Chaitanya S Lakkundi, Gopalakrishnan Rajaraman and Sai Rama Krishna Susarla	109
<i>Challenges in Processing Vedic Sanskrit: Towards creating a normalized dataset for the Ṛgveda-saṃhitā</i> Sriram Krishnan, Sepuri Gayathri and Amba Kulkarni	131
<i>Pāṇḍitya: Visualizing Sanskrit Intellectual Networks</i> Tyler Neill	148
<i>Anveshana: A New Benchmark Dataset for Cross-Lingual Information Retrieval on English Queries and Sanskrit Documents</i> Manoj Balaji Jagadeeshan, Prince Raj and Pawan Goyal	161
<i>Concordance of Sanskrit Synonyms</i> Dhaval Patel	181

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 Nāmalīṅgānuśāsanam. It also presents the genders of the words explicitly. Amarakośa is 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 Amarasīṃha, Hemacandra thought it fit to compose two separate works for synonymic and homonymic lexica. Therefore, in the present comparison, 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 Abhidhānacintāmaṇiśiloṇcha

Abhidhānacintāmaṇi also has two supplementary treatises Abhidhānacintāmaṇipariśiṣṭa by Hemacandra (ACPH) himself (also known as Śeṣasaṁgrahanāmamālā), and Abhidhānacintāmaṇiśiloṇcha 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



dhaval patel <drdhaval2785@gmail.com>

साङ्ख्यकारिकायां सङ्गति-तन्त्रयुक्त्योः प्रायोगिकविचारः

प्राची प्रज्ञा <prachiprajnaenb@gmail.com>

16 December 2025 at 10:19

To: Dhaval Patel <drdhaval2785@gmail.com>

To
Dr. Dhaval Patel
Department of Sanskrit Studies
University of Hyderabad

Dear Sir,

I am happy to inform you that your research paper "साङ्ख्यकारिकायां सङ्गति-तन्त्रयुक्त्योः प्रायोगिकविचारः (Usage of Sangati and Tantrayukti in Sāṅkhyakārikā)" has been accepted for publication in the 22nd issue of शास्त्रमञ्जूषा of the online journal प्राची प्रज्ञा, to be published in June 2026.

Thank you.

Regards,

Prof. Sugyan Kumar Mahanty

Editor in Chief,

प्राची प्रज्ञा / Prachi Prajna, ISSN 2348 - 8417

The Refereed & Peer Reviewed E-Journal in Sanskrit

प्राचीमनु प्रदिशं प्रेहि विद्वान् (शुक्लयजुर्वेदः/वामासं / 7. 66)

[Quoted text hidden]

साङ्ख्यकारिकायां सङ्गति-तन्त्रयुक्तयोः प्रायोगिकविचारः

Usage of Saṅgati and Tantrayukti in Sāṅkhyakārikā

5

Dhaval Patel

धवल पटेल

Ph.D. Student,

Department of Sanskrit Studies

10 University of Hyderabad

drdhaval2785@gmail.com

9909704306

J.S.R. Prasad

जे.एस.आर. प्रसाद

Professor,

Department of Sanskrit Studies

University of Hyderabad

jsrprasad@uohyd.ac.in

15 **Abstract**

Indian knowledge system has various tools of research methodology. Saṅgati and Tantrayukti are such tools of research methodology. They are the devices by which a treatise is able to coherently put forth its central doctrine. In the present paper, the authors try to define and elucidate the concept of various Saṅgatis and Tantrayuktis. Authors also examine occurrences of Saṅgatis and Tantrayuktis in Sāṅkhyakārikā and establish that these tools find widespread application in the same.

कुञ्चीशब्दाः

सङ्गतिः, तन्त्रयुक्तिः, साङ्ख्यकारिका, प्रसङ्गः, उपोद्धातः, हेतुता, अवसरः, निर्वाहकैक्यम्, कार्यैक्यम्, अध्याससङ्गतिः, आक्षेपसङ्गतिः, उदाहरणसङ्गतिः, प्रत्युदाहरणसङ्गतिः, तन्त्राधिकरणम्, योगः, हेत्वर्थः, पदार्थः, प्रदेशः, उद्देशः, निर्देशः, वाक्यशेषः, प्रयोजनम्, उपदेशः, अपदेशः, अतिदेशः, अर्थापत्तिः, निर्णयः, प्रसङ्गः, एकान्तः, अनेकान्तः, अपवर्गः, विपर्ययः, पूर्वपक्षः, अनुमतम्, विधानम्, व्याख्यानम्, संशयः, अतीतावेक्षा, स्वसंज्ञा, ऊह्यम्, समुच्चयः, निदर्शनम्, निर्वचनम्, नियोगः, प्रत्युत्सारः, सम्भवः

Justification for Similarity Index > 10%

Dr. Dhavalkumar Kirtikumar Patel's PhD thesis (21hsph01) reported a similarity index of 16%. Of these, the following two resources, which are co-authored by him, constitute 13% of the similarity index.

Following papers from the conference proceedings forms a substantial part of the thesis.

1. Dhaval K Patel and Amba Kulkarni. 2024. Word Sense Alignment of Sanskrit Lexica. In *Proceedings of the 7th International Sanskrit Computational Linguistics Symposium*, pages 1–13, Auroville, Puducherry, India. Association for Computational Linguistics.
2. Dhaval Patel. 2025. Concordance of Sanskrit Synonyms. In *Computational Sanskrit and Digital Humanities - World Sanskrit Conference 2025*, pages 181–188, Kathmandu, Nepal. Association for Computational Linguistics.

Amba P
25.1.2026

Supervisor
Department of Sanskrit Studies
School of Humanities
University of Hyderabad
P.O. Central University
Hyderabad-500 046.TS.

Word Sense Alignment of Sanskrit Lexica

by Dhavalkumar Patel

Submission date: 27-Jan-2026 03:58PM (UTC+0530)

Submission ID: 2864797114

File name: thesis_draft_library.pdf (1.45M)

Word count: 35645

Character count: 147929


पुस्तकालयाध्यक्ष
LIBRARIAN
इन्दिरा गांधी स्मृतिग्रन्थालय लाइब्रेरी
INDIRA GANDHI MEMORIAL LIBRARY
हैदराबाद विश्वविद्यालय
UNIVERSITY OF HYDERABAD
हैदराबाद-500 046, तेलंगणा
Hyderabad-500 046, Telangana

Word Sense Alignment of Sanskrit Lexica

ORIGINALITY REPORT

16%

SIMILARITY INDEX

16%

INTERNET SOURCES

2%

PUBLICATIONS

2%

STUDENT PAPERS

PRIMARY SOURCES

1 aclanthology.org
Internet Source

2 www.aclweb.org
Internet Source

3 dx.doi.org
Internet Source

4 core.ac.uk
Internet Source

5 Submitted to The Robert Gordon University
Student Paper

6 link.springer.com
Internet Source

7 dspace.uohyd.ac.in
Internet Source

8 Submitted to King's College
Student Paper

9 m.moam.info
Internet Source

13%

Supervisor
Department of Sanskrit Studies
School of Humanities
University of Hyderabad
P.O. Central University
Hyderabad-500 046.TS.

1%

<1%

<1%

<1%

<1%

<1%

<1%

<1%

- | | | |
|-------|---|------|
| 10 | <p>"Text, Speech, and Dialogue", Springer
Science and Business Media LLC, 2026
Publication</p> | <1 % |
| <hr/> | | |
| 11 | <p>Khaoula Belila, Okba Kazar, Mohammed
Charaf Eddine Meftah. "Automated Word
Sense Disambiguation Using WordNet
Ontology", International Journal of
Organizational and Collective Intelligence,
2022
Publication</p> | <1 % |
| <hr/> | | |
| 12 | <p>Submitted to Stanton University
Student Paper</p> | <1 % |
| <hr/> | | |
| 13 | <p>brightideas.houstontx.gov
Internet Source</p> | <1 % |
| <hr/> | | |
| 14 | <p>Philipp Sorg, Philipp Cimiano. "Chapter 4 An
Experimental Comparison of Explicit Semantic
Analysis Implementations for Cross-Language
Retrieval", Springer Science and Business
Media LLC, 2010
Publication</p> | <1 % |
| <hr/> | | |
| 15 | <p>www.coursehero.com
Internet Source</p> | <1 % |
| <hr/> | | |
| 16 | <p>Submitted to RMIT University
Student Paper</p> | <1 % |
| <hr/> | | |
| 17 | <p>Submitted to University of Hyderabad,
Hyderabad</p> | <1 % |

Student Paper		
18	itis2012.fis.unm.si Internet Source	<1 %
19	dokumen.pub Internet Source	<1 %
20	journal.equinoxpub.com Internet Source	<1 %
21	Abdo Ababor Abafogi. "Normalized Statistical Algorithm for Afaan Oromo Word Sense Disambiguation", International Journal of Intelligent Systems and Applications, 2021 Publication	<1 %
22	Submitted to Al Akhawayn University in Ifrane Student Paper	<1 %
23	Submitted to Heriot-Watt University Student Paper	<1 %
24	Submitted to TU Delft Student Paper	<1 %
25	Submitted to University of St Andrews Student Paper	<1 %

Exclude quotes ☐

Exclude bibliography ☐

Exclude matches ☐ < 14 words