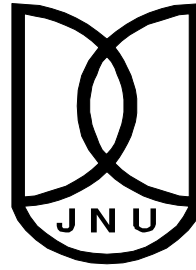


ONLINE INDEXING OF ĀDIPARVA IN MAHĀBHĀRATA

*Dissertation submitted to Jawaharlal Nehru University
in partial fulfillment of the requirements
for the award of the
degree of*

MASTER OF PHILOSOPHY

DIWAKAR MANI



SPECIAL CENTRE FOR SANSKRIT STUDIES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI-110067
INDIA
2008



विशिष्ट संस्कृताध्ययन केन्द्र
जवाहरलाल नेहरू विश्वविद्यालय
नई दिल्ली-११००६७

**SPECIAL CENTRE FOR SANSKRIT STUDIES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI-110067**

July 29, 2008

DECLARATION

I declare that the dissertation entitled “**Online Indexing of Ādiparva in Mahābhārata**” submitted by me for the award of the degree of **Master of Philosophy** is an original research work and has not been previously submitted for any other degree or diploma in any other institution/university.

(Diwakar Mani)



विशिष्ट संस्कृताध्ययन केन्द्र
जवाहरलाल नेहरू विश्वविद्यालय
नई दिल्ली-११००६७

**SPECIAL CENTRE FOR SANSKRIT STUDIES
JAWAHARLAL NEHRU UNIVERSITY
NEW DELHI-110067**

July 29, 2008

C E R T I F I C A T E

This dissertation entitled “**Online Indexing of Ādiparva in Mahābhārata**” submitted by **Diwakar Mani** to **Special Centre for Sanskrit Studies, Jawaharlal Nehru University, New Delhi-110067**, for the award of the degree of **Master of Philosophy**, is an original work and has not been submitted so far, in part or full, for any other degree or diploma of any University. This may be placed before the examiners for evaluation.

Prof. Varyam Singh
(Chairperson)

Dr. Girish Nath Jha
(Supervisor)

To

Beloved

Pitaji & Maa

ACKNOWLEDGEMENT

*ajñānatimirāndhasya jñānāñjanaśalākayā |
cakṣurunmīlitaṁ yena tasmai śrī gurave namaḥ ||*

*I acknowledge with deep sense of gratitude, continuous help and encouragement from my supervisor and the mentor **Dr. Girish Nath Jha**. Despite his busy schedule he has provided ample of time to me. His patience, generosity and excellent guidance have molded my work in this present shape. I am grateful to my guide, who inspired me to pursue research in the field of NLP.*

*I express sincere thanks and gratitude for Chairperson and other faculty members, especially **Shashiprabha Ma'm** and **Hariram Sir** of the Special Center for Sanskrit Studies (SCSS), Jawaharlal Nehru University, New Delhi, for encouraging and allowing this research to complete. I also express my sincere thanks to all the office staff of my centre and those who directly or indirectly helped me for completing this work.*

I duly acknowledge University Grants Commission (UGC) for providing me the financial assistance.

I pay my hearty thanks to Prof John Smith and his website for providing me authentic Unicode text of Mahābhārata.

*My special thanks to my Seniors and friends **Abhay Tiwari ji, Vimalji, Devendraji, Subhasji, Narayanji, Muktanand, Manji, Sachin** and **Surjit** for helping me systematize and structure the ideas for my research. I express my special thanks to **Srimon** for helping me to edit some chapters of my dissertation. Also I am very much thankful to **Diwakar Mishra** for his invaluable support.*

I extend my special appreciation to Vishwesh, Rajnish, Baldev, Archana, Jay, Devalina, Shashi and some other juniors of MA for entering data of this research.

*I am thankful to my dear friends **Mukeshji, Mansingh (BBC), Angadji, Vivek, Ravish, Rahul, Gajju, and Chandu**, who have always been the source of inspiration. **Ved**, is it necessary to acknowledge you?*

*I express my sincere thanks to **Anuja M'am** who has always encouraged me for my betterment.*

*I am deeply indebted to my **Jijaji, Upendra Bhaiya**, my Sisters (**Anita, Kiran and Poonam**) and brother **Sudhakar** for their prayers, affection and immense support. Also it is hard to forget the lovely and unbreakable talks of **Golu and Supriya**. Thank you **Maa, Pitaji** without you I would not be stand at this phase of my life, though this “Thank You” is not enough for you.*

I am solely responsible for any kind of error in this dissertation.

Contents

| | |
|--|------------------|
| ACKNOWLEDGEMENT | i-ii |
| CONTENTS | iii-vii |
| List of abbreviations | viii-ix |
| List of Tables | x |
| List of Snapshots and diagrams | xi |
| Transliteration key used in the dissertation | xii |
| Devanagari input mechanism according to Baraha software | xiii |
| INTRODUCTION | 01-04 |
| Chapter I | 05-24 |
| Computational Linguistics and Indian heritage | |
| 1.1. Recent Trends in Computational Linguistics and importance of Lexical resources | 05-11 |
| 1.1.1. Recent Trends in Computational Linguistics | |
| 1.1.2. Importance of Lexical resources | |
| 1.1.2.1. Statistical Method | |
| 1.1.2.2. Rule Based Method | |
| 1.2. Printed Lexicon vs. Electronic Lexicon | 11-13 |
| 1.2.1. Printed Lexicon | |
| 1.2.2. Electronic Lexicon | |
| 1.3. Electronic Indices vs. Search engines | 13-22 |
| 1.3.1. Electronic Indices | |
| 1.3.1.1. General Information | |
| 1.3.1.2. Computer-Aided Indexing | |
| 1.3.1.3. Automatic Indexing Methods | |
| 1.3.1.3.1. Statistical analysis | |
| 1.3.1.3.2. Syntactic and Semantic Analysis | |
| 1.3.2. Search engines | |
| 1.3.2.1. History of Search engines | |
| 1.3.2.2. How does the Search engine work? | |
| 1.4. Electronic Indices for Sanskrit | 22-24 |
| 1.4.1. Why for Sanskrit Text? | |

Chapter II 25-63 Lexical Resources in Indian Tradition

2.1. Kośa Paramparā 25-44

- 2.1.1. The Nighaṇṭu
- 2.1.2. The Nirukta
- 2.1.3. Vyāḍi
- 2.1.4. Kātya
- 2.1.5. Amarakośa
- 2.1.6. Dhanvantarinighaṇṭu
- 2.1.7. Vaijayantīkośa
- 2.1.8. Maṅkhakośa
- 2.1.9. Kalpadrukośa
- 2.1.10. Kośakalpataru
- 2.1.11. Hemacandra
 - 2.1.11.1. Abhidhānacintāmaṇi
 - 2.1.11.2. Anekārthasaṁgraha
 - 2.1.11.3. Nighaṇṭuśeṣa
 - 2.1.11.4. Deśināmamālā
- 2.1.12. Anekārthasamuccaya
- 2.1.13. Anekārthadhvanimañjarī
- 2.1.14. Paryāyaratnamālā
- 2.1.15. Paryāyamuktāvalī
- 2.1.16. Abhidhānaratnamālā
- 2.1.17. Puruṣottamadeva
 - 2.1.17.1. Trikāṇḍaśeṣa
 - 2.1.17.2. Hārāvalī
 - 2.1.17.3. Varṇadeśanā
 - 2.1.17.4. Dvirūpakośa
 - 2.1.17.5. Ekāṣarakośa
- 2.1.18. Viśvaprakāśa
- 2.1.19. Śabdabhedaprakāśa
- 2.1.20. Nāmamālā
- 2.1.21. Anekārthakośa
- 2.1.22. Ekāṣaranāmamālikā

2.1.23. Madanavinodanighaṇṭu

2.1.24. Rājanighaṇṭu

2.1.25. Pārasīprakāśa

2.1.26. Rājavyavahārakośa

2.2. E-indices and computational lexicography

44-57

2.2.1. E-indices

2.2.1.1 . E-index of Mahābhārata

2.2.1.2 . E-index for Sanskrit Dramas

2.2.2. Computational Lexicography

2.2.2.1. Online Multilingual Amarakośa

2.2.2.2. Online Macdonell Sanskrit-English Dictionary

2.2.2.3. Cologne Digital Sanskrit Lexicon & Capeller's
Sanskrit-English Dictionary

2.2.2.4. Online Monier Williams (MW) Sanskrit-English
Dictionary Advanced Search

2.2.2.5. Online Apte English-Sanskrit Dictionary

2.2.2.6. Online Apte Sanskrit-English Dictionary

2.2.2.7. Spoken Sanskrit Dictionary

2.2.2.8. The Sanskrit Heritage Dictionary

2.2.2.9. Database query to Bloomfield's Vedic Concordance

2.2.2.10. Database query to Lubotsky's Rigvedic
Word Concordance

2.2.2.11. Database query to Lubotsky's Indo-Aryan
inherited lexicon

2.3. Methodology for e-indexing of Mahābhārata (Ādiparva)

57-63

2.3.1. Study and ascertain the most authentic text of
the Mahābhārata and its structure

2.3.2. Sandhi splitting

2.3.3. Creation of a Relational Database System
for Mahābhārata (Ādiparva)

2.3.4. Creating Java servlet engine for search

2.3.5. Creating an interactive page for search
with display

2.3.6. Primary Information

2.3.7. Secondary Information

2.3.8. Tertiary Information

| | |
|--|---------------|
| Chapter III | 64-91 |
| Structure of Mahābhārata and the Ādiparva | |
| 3.1. The Text of Mahābhārata | 64-74 |
| 3.2. Mahābhārata according to BORI | 74-84 |
| 3.2.1. BORI and the Mahābhārata | |
| 3.2.2. Structure of Mahābhārata | |
| 3.2.2.1. Ādiparva | |
| 3.2.2.2. Sabhāparva | |
| 3.2.2.3. Āraṇyakaparva | |
| 3.2.2.4. Virāṭaparva | |
| 3.2.2.5. Udyogaparva | |
| 3.2.2.6. Bhīṣmaparva | |
| 3.2.2.7. Droṇaparva | |
| 3.2.2.8. Karṇaparva | |
| 3.2.2.9. Śalyaparva | |
| 3.2.2.10. Saupatikaparva | |
| 3.2.2.11. Strīparva | |
| 3.2.2.12. Śāntiparva | |
| 3.2.2.13. Anuśāsanaparva | |
| 3.2.2.14. Aśvamedhikaparva | |
| 3.2.2.15. Āśramavāsikaparva | |
| 3.2.2.16. Mausalaparva | |
| 3.2.2.17. Mahāprasthānikaparva | |
| 3.2.2.18. Svargārohaṇaparva | |
| 3.2.2.19. Harivaṃsa (Khilaparva) | |
| 3.3. Ādiparva of Mahābhārata | 84-91 |
| 3.3.1. Introduction to Ādiparva of Mahābhārata | |
| 3.3.2. Structure of Ādiparva | |
| 3.3.3. Computer adaptation of Ādiparvan | |
| Chapter IV | 92-115 |
| Online Indexing of Mahābhārata (Ādiparva) | |
| 4.1. Introduction | 92-92 |
| 4.2. Architecture of the system | 92-92 |
| 4.3. Process of the indexing system | 93-94 |
| 4.4. The front-end of the Mahābhārata Indexer | 94-102 |

| | | |
|--|---|----------------|
| 4.4.1. | Java Server Pages | |
| 4.4.2. | Java Servlet Technology | |
| 4.4.3. | Apache Tomcat 4.0 Web Server | |
| 4.5. | The Back-end | 102-103 |
| 4.6. | Database connectivity | 103-103 |
| 4.7. | How does the indexing system work? | 104-105 |
| 4.7.1. | How to use online version of the system | |
| 4.7.2. | How to use CD version of the system | |
| 4.8. | The snapshots of the Mahābhārata indexer | 106-115 |
| CONCLUSION | | |
| | 116-117 | |
| APPENDICES | | 118-131 |
| BIBLIOGRAPHY | | 132-137 |
| Mahābhārata indexer CD Enclosed | | |

List of Abbreviations

| | |
|----------------|---|
| AI | Artificial Intelligence |
| ALPAC | Automatic Language Processing Advisory Committee |
| ASCII | American Standard Code for Information Interchange |
| BICORD | Bilingual Corpus-based Dictionary |
| BORI | Bhandarkar Oriental Research Institute |
| C-DAC | Centre for Development of Advance Computing |
| CL | Computational Linguistics |
| CLAWS | Constituent Likelihood Automatic Word-tagging System |
| DEC | Digital Equipment Corporation |
| DS | Deep Structure |
| FTP | File Transfer Protocol |
| GIST | Graphics and Intelligence based Script Technology |
| GUI | Graphical User Interface |
| HTML | Hyper Text Markup Language |
| HTTP | Hyper Text Transfer Protocol |
| ICT | Information and Communication Technology |
| IIT | International Institute of Information Technology |
| IIT | Indian Institute of Technology |
| INRIA | Institut National de Recherche en Informatique et Automatique |
| ISCI | Indian Standard Code for Information Interchange |
| JDBC | Java Database Connectivity |
| JNU | Jawaharlal Nehru University |
| JSP | Java Server Pages |
| Jughead | Jonzy's Universal Gopher Hierarchy Excavation And Display |
| LS | Lakṣmaṇa Saṁvat |
| M(A)T | Machine Aided Translation |
| M(A)TS | Machine Aided Translation System |
| Mbh | Mahābhārata |
| MGIHU | Mahatma Gandhi International Hindi University |

| | |
|-----------------|---|
| MIRI | Muktabodha Indological Research Institute |
| Mss. | Manuscripts |
| MS-SQL | Microsoft Structure Query Language |
| MT | Machine Translation |
| NCST | National Centre for Software Technology |
| NLP | Natural Language Processing |
| NS | Nepālī Samvat |
| PDF | Portable Document Form |
| R&D | Research & Development |
| RDBMS | Relational Database Management System |
| SCSS | Special Centre for Sanskrit Studies |
| SIMR | Smooth Injective Map Recognizer |
| SS | Surface Structure |
| STT | Speech to Text |
| TDIL | Technology Development in Indian Languages |
| TIFR | Tata Institute of Fundamental Research |
| TTS | Text to Speech |
| UCSG | Universal Clause Structure Grammar |
| UNL | Universal Networking Language |
| UTF | Unicode Transformation Format |
| Veronica | Very Easy Rodent-Oriented Net-wide Index to Computerized Archives |
| VS | Vikrama Samvat |
| WWW | World Wide Web |

List of Tables

| Table No. | Name of the Table | Page No. |
|-----------|---|----------|
| 1.1 | Most popular search engines | 20 |
| 2.1 | Structure of <i>Amarakośa</i> | 28 |
| 2.2 | Structure of <i>Vaijayantīkośa</i> | 30-31 |
| 2.3 | Structure of <i>Kalpद्रुकोśa</i> | 32 |
| 2.4 | Structure of <i>Kośakalpataru</i> | 33-34 |
| 2.5 | Structure of <i>Abhidhānacintāmaṇi</i> | 35 |
| 2.6 | Data of Ādiparva in Saṁhitā and Split form | 59 |
| 2.7 | Example of unusual sandhi splitting | 59 |
| 2.8 | Database structure of main table (without data) | 60 |
| 2.9 | Database structure of ‘adhyaya’ table (without data) | 61 |
| 2.10 | Database structure of ‘akhyana’ table (without data) | 61 |
| 2.11 | Database structure of main table ‘upaparva’ table (without data) | 61 |
| 2.12 | Database structure of main table ‘parva’ table (without data) | 62 |
| 3.1 | Total number of Mss. used for compilation of first parvan of critical edition of <i>Mbh.</i> | 66 |
| 3.2 | The list of <i>Mbh.</i> Mss. found in the volumes of critical edition of BORI and some other libraries. | 67-74 |
| 3.3 | Structure of <i>Mbh.</i> | 83-84 |
| 3.4 | Structure of Ādiparva of <i>Mbh.</i> | 87 |
| 3.5 | Database structure of main table (with data). | 88-89 |
| 3.6 | Database structure of ‘adhyaya’ table (with data). | 89 |
| 3.7 | Structure of the third table of the database (with data). | 90 |
| 3.8 | Structure of the fourth table of the database (with data). | 90 |
| 3.9 | Structure of the fifth table of the database (with data). | 91 |

List of Snapshots and diagrams

| Fig. No. | Name of Snapshot and diagram | Page No. |
|----------|--|----------|
| D.1 | Architecture of the system | 03 |
| D.2 | Process flow of the system | 03 |
| 2.1 | Home page of Goenttingen's Mahabharata Indexer | 45 |
| 2.2 | Second page of Goenttingen's Mahabharata Indexer | 45 |
| 2.3 | Final page of the Goenttingen's Mahabharata Indexer | 46 |
| 2.4 | Main page of <i>Bhāsa</i> indexer | 46 |
| 2.5 | Second page of <i>Bhāsa</i> indexer | 47 |
| 2.6 | Third page of <i>Bhāsa</i> indexer | 47 |
| 2.7 | Search page of Multilingual Amarakosha | 49 |
| 2.8 | Second page of Multilingual Amarakosha | 50 |
| 2.9 | Third page of Multilingual Amarakosha | 50 |
| 2.10 | Search page of Macdonell's Sanskrit-English dictionary | 51 |
| 2.11 | Result page of Macdonell's Sanskrit-English dictionary | 51 |
| 2.12 | Main page of Cologne Digital Sanskrit Lexicon | 52 |
| 2.13 | Snapshot of online MW Sanskrit-English Advance dictionary | 52 |
| 2.14 | Online Apte's English-Sanskrit dictionary | 53 |
| 2.15 | Search page of online Apte's Sanskrit- English dictionary | 54 |
| 2.16 | Result page of online Apte's Sanskrit- English dictionary | 54 |
| 2.17 | Snapshot of Spoken Sanskrit Dictionary site | 55 |
| 2.18 | Snapshot of Sanskrit Heritage Dictionary | 55 |
| 2.19 | Snapshot of web page of the Bloomfield's Vedic Concordance | 56 |
| 2.20 | Snapshot of webpage of Lubotsky's Rigvedic Word Concordance | 56 |
| 2.21 | Snapshot of web page of Lubotsky's Indo-Aryan inherited lexicon | 57 |
| 4.1 | Multi-tier architecture of the Mahabharata Indexer | 91 |
| 4.2 | Process of Indexing system | 92 |
| 4.3 | Diagram of the interconnection between the database tables | 102 |
| 4.4 | Homepage of the <i>Mahābhārata</i> Indexer | 105 |
| 4.5 | Search page of the <i>Mbh</i> indexer | 106 |
| 4.6 | Third page of the <i>Mbh</i> indexer | 107 |
| 4.7 | Referential page of the <i>Mbh</i> indexer | 108 |
| 4.8 | Cross-referential page of the <i>Mbh</i> indexer | 109 |
| 4.9 | Cross-referential page of the <i>Mbh</i> indexer | 110 |
| 4.10 | Cross-referential page of the <i>Mbh</i> indexer | 111 |
| 4.11 | Cross-referential page of the <i>Mbh</i> indexer | 112 |
| 4.12 | Cross-referential page of the <i>Mbh</i> indexer | 113 |
| 4.13 | Cross-referential page of the <i>Mbh</i> indexer i.e, Sanskrit Wikipedia | 114 |

Transliteration key used in the dissertation

| | | | | | |
|----|---|----|-----------------------|---|----|
| अ | = | a | ढ | = | ḍh |
| आ | = | ā | ण् | = | ṇ |
| इ | = | i | त् | = | t |
| ई | = | ī | थ् | = | th |
| उ | = | u | द् | = | d |
| ऊ | = | ū | ध् | = | dh |
| ऋ | = | r̥ | न् | = | n |
| ॠ | = | r̄ | प् | = | p |
| लृ | = | l̥ | फ् | = | ph |
| ए | = | e | ब् | = | b |
| ऐ | = | ai | भ् | = | bh |
| ओ | = | o | म् | = | m |
| औ | = | au | य् | = | y |
| क् | = | k | र् | = | r |
| ख् | = | kh | ल् | = | l |
| ग् | = | g | व् | = | v |
| घ् | = | gh | श् | = | ś |
| ङ् | = | ṅ | ष् | = | ṣ |
| च् | = | c | स् | = | s |
| छ् | = | ch | ह् | = | h |
| ज् | = | j | क्ष् | = | kṣ |
| झ् | = | jh | त्र् | = | tr |
| ञ् | = | ñ | ज्ञ् | = | jñ |
| ट् | = | ṭ | ऽ | = | ' |
| ठ् | = | ṭh | ˘ (<i>Anusvāra</i>) | = | m̐ |
| ड् | = | ḍ | : (<i>visarga</i>) | = | ḥ |

Devanāgarī input mechanism

[According to Baraha software (<http://www.baraha.com>)]

| VOWELS | | | | |
|------------|-----------|---------|------------|------------|
| a [अ], | aa/A [आ], | i [इ], | ee [ई], | u [उ], |
| oo [ऊ], | Ru [ऋ], | RU [ॠ], | IRu [ऌ], | IRU [ॡ], |
| e [ए], | ai [ऐ], | o [ओ], | au [औ], | aM [अं], |
| aH [अः] | | | | |
| CONSONANTS | | | | |
| k [क], | kh/K [ख], | g [ग], | gh [घ], | ~G [ङ], |
| c [च], | C [छ], | j [ज], | jh/J [झ], | ~J [ञ], |
| T [ट], | Th [ठ], | D [ड], | Dh [ढ], | N [ण], |
| t [त], | th [थ], | d [द], | dh [ध], | n [न], |
| p [प], | ph [फ], | b [ब], | bh [भ], | m [म], |
| y [य], | r [र], | l [ल], | v/w [व], | sh/S [श], |
| Sh;[ष] | s [स], | h [ह], | kSh [क्ष], | tra [त्र], |
| j~J [ज्ञ], | | | | |

Introduction

The scope of the present research is to develop a relational database system for the storage and interactive indexing of the *Ādiparva* of *Mahābhārata* (*Mbh*). Some attempts have been made by University of Goettingen, Germany to develop a static indexer of *Mbh*. Except this, there is no online indexing system for a text of Indian heritage which comprehensively allows for search either in static or dynamic fashion.

Mbh, the great epic of India ascribed to *Veda Vyāsa*, can be un-hesitatingly given the honor of being the cultural encyclopedia of India. The *Mbh* is the story of a great war that ended one age and began another. The story has been passed down to us in a classical canon of Sanskrit verses - some 90,192 stanzas (including additional *Harivaṃśa*) long, or some 1.8 million words in total (among the longest epic poems worldwide) divided in 18 *parvans*, 98 *upa-parvans* (excluding the two of *Harivaṃśa*) which are again divided into 1995 chapters. The work holds a significant place in the cultural history of India.

Why indexing system is necessary for Mahābhārata?

The *Mbh* extols its greatness itself in the following words: “*Yadihāsti tadanyatra yannehāsti na tat kvacit*”. The saying “*Vyāsocchiṣṭam jagatsarvam*” also stresses this point. *Mbh* is neither history in the modern sense of the term, nor chronicle. But it stands in incomparable isolation, defying all definitions. It is a veritable encyclopedia comprising heterogeneous material from all branches of knowledge. Taking the core-story of the feud between two branches of a royal family and the circumstances leading to a catastrophic war, several branches of knowledge including philosophy, law, ethics, statecraft, warfare, history and ethnology are embodied in its structure. The *Mbh* is a comment on the human condition with all its richness, complexity and subtlety. The *Mbh* is the source of many compositions like *Abhijñāna Śākuntalam* of *Kālidāsa*, *Naiṣadhīya Caritam* of *Śrīharṣa* etc. It is the text that is most sought for in order to enrich cultural, social and any type of knowledge about Indian civilization.

While on the one hand, the text is very important, on the other, it is so huge that it becomes virtually impossible for someone to search a specific keyword in it. The indices thus prepared will constitute a separate text in itself due to the size of the *Mbh* and will be of tremendous use to the researchers and users.

Uses of the indexing system

An indexing is a basic step of any search engine. A web crawler which browses the www in a methodical, automated manner is called an automatic indexer.¹ The crawler copies the visited web pages and those are indexed to give fast search results. The indexing system of Sanskrit documents can be used in various NLP applications like building Sanskrit WordNet, dictionaries, Sanskrit- Indian Language Machine Translation System (MTS) etc. This work, besides being an essential resource in NL system of Sanskrit, may also be useful for authentic and referential knowledge about Indian heritage. The system can also be very useful for the researches of historical, socio-political and geographical researches by providing the facts from the huge text which cannot be easily read.

Methodology for Mahābhārata indexer

The *Mbh*, an encyclopedia of Indian civilization has always been attracted not only to Indian scholars, but also to the western. The *Mbh* being a popular epic has several versions. The editions selected for this work is the critical edition of Mahābhārata, critically edited by V.S. Sukthankar and published by BORI (1944). It has been digitized by Prof. John Smith in Unicode Devanāgarī Text format.

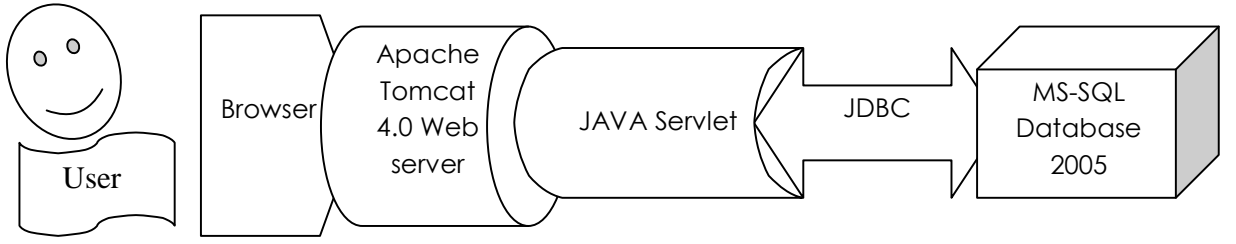
To provide more comprehensive search the text is segmented according to Pāṇinian sandhi rules. After this the text has been adapted to the database system. The original and segmented text has been stored in database tables. The other information of the structure of the text has been stored in different tables and those are connected with each other. The connections of table complete the reference of the searched query and connect all the data with other relative data. The database has five tables having the information of *Parva*, *Upaparva*, *Adhyāya*, *Ākhyāna* and the *Śloka*s respectively. The connections are defined through the table diagram in the database.

Development of the Mahābhārata indexer

A dynamic search engine-cum-indexer has been developed under this research. It is built in the front-end of Apache Tomcat Web server using JSP and Java servlets. It has its data in MS-SQL Server 2005 with Unicode. For connecting the front-end to the database server the MS-JDBC

¹ http://en.wikipedia.org/wiki/Web_crawler (accessed on July 20, 08)

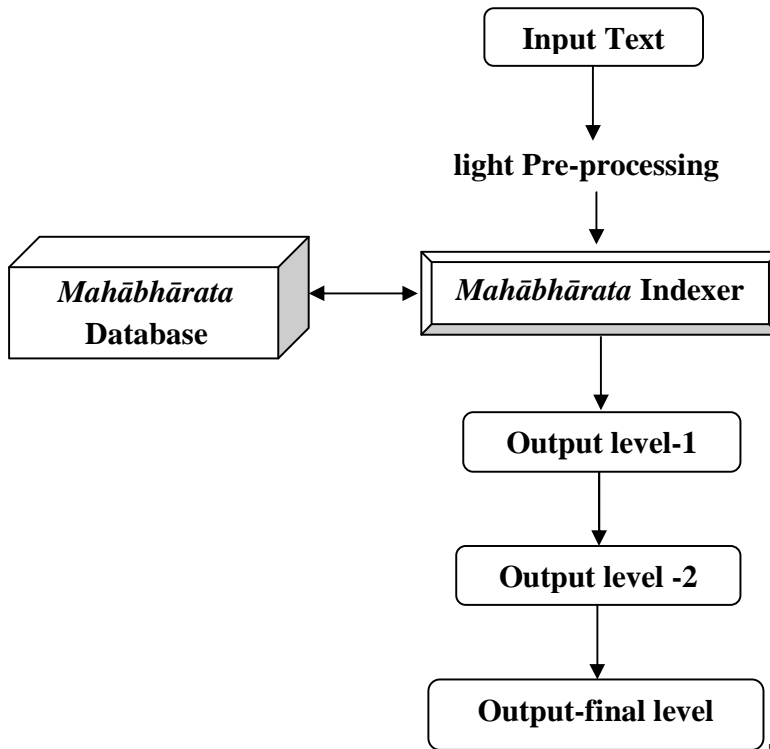
connectivity has been used. The system is available online on <http://Sanskrit.jnu.ac.in/mb> with input and output in Devanāgarī Unicode. The system works as an interactive and multi-dimensional knowledge based indexing system for *Ādiparva* of *Mbh* and it is being extended to all the *parvans* of *Mbh*. The system can be used also as a generic system for all Sanskrit texts of similar structure. The architecture of the system is as follows:



[Fig D.1: Architecture of the system]

Process flow of the system

There are three ways to give input to the system e.g. Direct Search, Alphabet search and Search by the structure of the text in Devanāgarī UTF-8 format.



[Fig D.2: Process flow of the system]

Chapter description

The dissertation has been divided into four chapters. The first chapter '*Computational Linguistics and Indian heritage*' discusses the recent developments in computational linguistics and the importance of lexical resources. It also illustrates lexicographic and search engine methods and techniques. At the end of the chapter, the survey of similar works has been given.

The second chapter '*Lexical Resources in Indian Tradition*' discusses the tradition of Sanskrit lexicographic works. It describes 26 major and the few other *kośas* in first section. The second unit gives the details of some important available online lexical resources. The last unit of the chapter deals with the methodology applied for this research.

The third chapter '*Structure of Mahābhārata and the Ādiparva*' is a detailed study of the structure and contents of the *Mbh* and its first parva- *Ādiparva*. This chapter describes the structure of the *Ādiparva* of *Mbh* and gives its RDBMS adaptation.

The fourth chapter '*Online Indexing of Mahābhārata (Ādiparvan)*' talks about the implementation aspects such as the front end, Java objects, lexical resources and how they work.

In the concluding part of the dissertation, the limitations of the system and its implications for future research have been summarized. The appendices contain the sample data of the *Mbh* (*Ādiparva*) used to develop the indexing-cum-search engine of *Mbh*. A portable CD has also been enclosed with the dissertation which comprises the sample data and a screen-shot of the interface.

Chapter-1

Computational Linguistics and Indian Heritage

1.1. Recent Trends in Computational Linguistics and the importance of Lexical resources

1.1.1. Recent Trends in Computational Linguistics (CL)

Computational Linguistics came into existence in 1949 as a Machine Translation (MT) project when Warren Weaver suggested in his famous memorandum that translation by machine might be possible. However, this area of MT has been moved in the sub-area of CL in the mid 1960s and after that it contributed a lot, both to the area of linguistics and computer science. The major areas of research and development in computational linguistics can be summarized as below.¹

- **Corpora Creation**
- **MT and M(A)T**
- **Speech Processing**
- **Localization**
- **Tools**
- **E-Contents**

Corpora Creation is one of the prime requirements of language processing. A good corpus, properly tagged and annotated, serves the basis of a good and efficient linguistic tool. Corpus data are the raw fuel for many applications and the test bed on which Natural Language Processing (NLP) applications are evaluated. The field of corpus building has become so vast that a whole disciplinary term of corpus linguistics² has emerged. Several corpora have been built over the years in many languages, for example the British National Corpus and the Penn Tree Bank for English, the European Union funded multilingual English-French-Spanish corpus, the CLAWS English text tagger etc. In India, the same effort has been made for Sanskrit, Hindi and other languages at several places, for example; IIIT Hyderabad, IIT Kanpur and Delhi, MGIHU Wardha, C-DAC Noida, SCSS/JNU³ etc.

The history of **MT** can be traced from the pioneers and early systems of the 1950s and 1960s, the impact of the ALPAC report in the mid-1960s, the revival in the 1970s, the appearance of

¹ Jha, Girish N, an article on “Language Technology in India” in CSI Communication, Nov, 2005.

² (See Mitkov, T. 2003.)

³ <http://sanskrit.jnu.ac.in>

commercial and operational systems in the 1980s, research during the 1980s, new developments in research in the 1990s, and the growing use of systems in the past decade. However, reaching a hundred percent correct output of a machine translated text is still to be achieved and research and development in this area is still on to bring about the best that a machine can deliver. In India, the initiative towards work on MT is relatively young. The earliest efforts date from the late 80s and early 90s. AnglaBharati and AnuBharati projects of IIT Kanpur are ambitious efforts towards developing an MT system not just for Hindi and English but for also other Indian languages.⁴ Similar kinds of efforts are also on at IIIT Hyderabad (the Shakti MTS), C-DAC Mumbai (the MaTra MT system), C-DAC Pune & Bangalore (the Mantra project), University of Hyderabad (UCSG-based English-Kannada MT), IIT Bombay (UNL-based MT between English, Hindi and Marathi), AU-KBC Research Centre Chennai (Tamil-Hindi Anusaaraka and English-Tamil MT), Jadavpur University Kolkata (English-Hindi M(A)T for news sentences), Super Infosoft Pvt. Ltd. (Anuvadak English-Hindi software) and at IBM IRL (English-Hindi Statistical MT).⁵ The Technology Development in Indian Languages (TDIL), an initiative of the Department of IT, Ministry of Communications and Information Technology, Government of India, has played an instrumental role by funding these projects.⁶ Google's translator has made a first level effort to translate English text and web pages into Hindi and vice-versa on-the-fly.⁷

Study of speech signals and the processing methods of these signals are called **Speech processing**.⁸ The signals are usually processed in a digital representation whereby speech processing can be seen as an intersection of digital signal processing and natural language processing. There are several sub-categories in the area of Computational Linguistics where research and development is on. The area of speech processing can be divided into the following categories: i.e. i) **speech recognition**- which deals with analysis of the linguistic content of a speech signal, ii) **speaker-recognition**- where the aim is to recognize the identity of the speaker, iii) **speech-synthesis**- the artificial synthesis of speech, which usually means computer generated speech and iv) **speech enhancement**- enhancing the perceptual quality of speech signal by

⁴ For more details:- <http://www.cse.iitk.ac.in/users/langtech/anglabharti.htm> (accessed on May 02, 08)

⁵ Rao, Durgesh, an article on "Machine Translation in India: A brief Survey" at- <http://www.elda.org/en/proj/scalla/SCALLA2001/SCALLA2001Rao.pdf> (accessed on May 07, 08)

⁶ For more details:- <http://tdil.mit.gov.in/> (accessed on May 08, 08)

⁷ http://translate.google.com/translate_t (accessed on June 07, 08)

⁸ http://en.wikipedia.org/wiki/Speech_processing (accessed on June 06, 08)

removing the destructive effects of noise, limited capacity recording equipment, impairments etc.. Besides these, of particular importance are TTS and STT systems that have come in several languages. Efforts in this area in India are made mainly at TDIL, C-DAC centers, Media Lab Asia centers and at TIFR.

Localization (sometimes shortened to 'l10n') is the process of adapting a product or service to a particular language, culture, and a desired local 'look-and-feel'. Localizing the software applications and platforms is a requirement of the present day world so that the great digital divide that has been created due to linguistic hurdles could be linked. The linguistic diversity is considered to be an asset for the world and this diversity is in danger as English and its resources are gaining greater grounds in several areas and thereby limiting the use of the other languages for some specific domains. Several governments have understood the need for a remedy and it is eagerly sought through the localization of software applications and platforms. Many public and private endeavors in India are being put forward in the area of localization, for example, Indix, a localization of graphical user interface of Linux operating system is being developed at C-DAC Mumbai (formerly known as NCST), through TDIL program of Ministry of Information Technology and Telecommunications, Govt. of India. The work of translating OpenOffice.org into many Indian Languages (BharateeyaOO.o) is happening at C-DAC Bangalore. Apart from C-DAC Mumbai and Bangalore, Localization work is going in IndLinux,⁹ Utkarsh,¹⁰ Punlinux,¹¹ GIST group at C-DAC Pune,¹² Yahoo India,¹³ Google India,¹⁴ Microsoft India¹⁵ and many others.

Tools are other areas of computational linguistics where research and development are carried out with given short term goals that serve major needs. Some examples of tools developed and under development process for Indian languages are the ITERM- an Indian script terminal for Unix X windows, word processors, Anusaarka- a language accessor among Indian languages to Hindi, Sanskrit Authoring Systems, Devanāgarī search engines for Unicode etc.

⁹ <http://indlinux.org/> (accessed on May 12, 08)

¹⁰ It works for Gujarati speaking populace- <http://www.utkarsh.org/> (accessed on May 12, 08)

¹¹ It is dedicated to Punjabi language- <http://punlinux.sourceforge.net/> (accessed on May 12, 08)

¹² http://www.cdac.in/html/gist/lang_tools/local_frmwk.asp (accessed on May 12, 08)

¹³ Yahoo has localized itself in seven Indian languages. Visit:- <http://in.yahoo.com/> (accessed on May 12, 08)

¹⁴ Google has localized itself in many Indian languages including Sanskrit. Visit:- <http://www.google.co.in/> and <http://www.google.com/intl/sa/> (accessed on May 12, 08)

¹⁵ For more details:- <http://www.bhashaindia.com/Community/CommunityHome.aspx> (accessed on May 12, 08)

Linguistic e-contents are another widely used area of computational linguistics. There are several such contents already available on the web that can be availed freely. The bilingual electronic dictionaries are available in many languages. Word nets, corpora, speech databases, web based education systems are some examples of linguistic e-contents.

1.1.2. Importance of Lexical resources

The term *computational lexicography* can roughly be interpreted as a subfield of CL concerned with the design and implementation of lexical resources for NLP.¹⁶ Since, the lexicographic work turns out to be very data-intensive, so the support provided by machines is extremely helpful and highly welcome. In this area, computers may be helpful at both ends: lexicographer and user. By means of the computer, the lexicographer shall be supported in acquiring, editing, representing and processing lexicographic data and on the other end, the dictionary user shall have electronic access to relevant data of his interest. Besides this, many NLP systems are in need of lexical resources which make available lexical information required to process spoken or written utterances. The more the applications in academic world and the industry depend on NLP, the higher the need for these lexical resources will be. The processes of building, editing and extending dictionaries as well as research on dictionaries can profit from recent developments and progress in CL. Huge samples of recorded language use are nowadays available in the form of corpora and databases, and computational methods, both linguistic and statistical, of retrieving, annotating and exploiting the data have been developed and refined. Given the dynamics and interactions of these developments, the potential of computational lexicography has scarcely been exhausted yet.

1.1.2.1. Statistical Method

Statistical methods¹⁷ can be used to summarize or describe the collection of data of a lexicon. The patterns in the data of a lexicon may be modeled in a way that accounts for randomness and uncertainty in the observations, and then used to draw inferences about the process being studied. The statistical method is very useful for MT and M(A)T system. This invention learns phrase translation relationships by receiving a parallel aligned corpus with phrases to be learned identified in a source language. Candidate phrases in a target language are generated and an

¹⁶ http://milca.sfs.uni-tuebingen.de/B2/SS_2004/Intro/MiLCA_COLEX_Intro.xhtml (accessed on June 10, 08)

¹⁷ <http://en.wikipedia.org/wiki/Statistics> (accessed on June 09, 08)

inside score is calculated based on word association scores for words inside the source language phrase and candidate phrase. An outside score is calculated based on word association scores for words outside the source language phrase and candidate phrase. The inside and outside scores are combined to obtain a joint score.¹⁸

The ***K-vec***¹⁹ algorithm is a statistical approach for bilingual lexicon extraction and was developed by Pascale Fung in the Computer Science Department at Columbia University, New York. The first step in this method is to extract lexicon candidates by looking for similarities in the distribution of source and target language word. For this purpose, the bilingual text is split into K pieces. Then K -dimensional binary vectors are created for the source and the target language word. If a specific text piece contains the source language word (or the target language word respectively), the corresponding flag in the vector will be set. Some another Statistical approaches and techniques used for automatic lexicon extraction are:

- a) The ***Termight system*** can be seen as a tool for supporting bilingual lexicon creation. It was developed at the AT&T Bell laboratories by Ido Dagan and Ken Church and represents a semi-automatic tool for the identification of technical terms and the support of translation processes.²⁰
- b) ***Char_align*** is a program for aligning parallel texts. It was developed at the AT&T Bell laboratories by Kenneth Ward Church. This approach uses the existence of cognates between cross-language token pairs.²¹
- c) The ***DEC*** by Pim van der Eijk is a method which concentrates on identifying noun phrase correlations from a previously aligned and tagged parallel corpus. The statistical method for finding correlations is based on the assumption that the translation is more frequent in the subset of the target alignments which are aligned to the source language term than in the entire target language text.²²
- d) An approach developed by I. Dan Melamed at the University of Pennsylvania in Philadelphia concentrates on the automatic lexicon evaluation by applying several filters. These filters use external knowledge sources and heuristics. First, all source language and target language words from a sentence alignment were combined into word pairs. Then,

¹⁸ <http://www.freepatentsonline.com/7249012.html> (accessed on June 11, 08)

¹⁹ <http://stp.ling.uu.se/~joerg/diplom/node4.html> (accessed on June 11, 08)

²⁰ <http://portal.acm.org/citation.cfm?id=974367> (accessed on June 11, 08)

²¹ www.aclweb.org/anthology/P93-1001.pdf (accessed on June 14, 08)

²² <http://acl.ldc.upenn.edu/E/E93/E93-1015.pdf> (accessed on June 14, 08)

the filters were applied in cascades to find the N-best translations among the translation candidates. For this purpose Melamed used parts of speech, machine-readable bilingual dictionary, cognate and word alignment filters. The part of speech filter removes every translation candidate with different parts of speech in the source and the target language. If a translation candidate appears in the second filter, the machine-readable bilingual dictionary (MRBD), all pairs with the same source language word and a different target language word, and all pairs with the same target language word and a different source language word which occur in the same sentence pair will be removed. Cognate filters are based on the assumption that there are similarities between the source language word and its translation in related languages. The last filter assumes that in related languages information is expressed with a similar word order. The heuristic in this case is that crossing alignments are not very probable. Some experiments of this method were made with the Canadian Hansards corpus. The precision for the single best translations yielded about 52%.²³

- e) **BICORD** statistical approach of Judith Klavans and Evelyn Tzoukermann concentrate on combining lexicon data from machine-readable dictionaries and bilingual lexica.²⁴
- f) **DK-vec** based on the K-vec method, Pascale Fung and Kathleen McKeown developed a new algorithm for aligning noisy parallel corpora.²⁵
- g) **SIMR**, an approach to map bilingual text correspondences as developed at the University of Pennsylvania in Philadelphia by Dan Melamed. Like char align, it uses cognates between the source and target language texts to align the bilingual corpora at the character level.²⁶
- h) At the IBM research laboratories, a stochastic language system, *Candide*, was developed as a statistical MT system. This system involves considerations of word groups as well as single words.²⁷

²³ <http://stp.ling.uu.se/~joerg/diplom/node4.html> (accessed on June 14, 08)

²⁴ <http://acl.ldc.upenn.edu/C/C90/C90-3031.pdf> (accessed on June 16, 08)

²⁵ <http://arxiv.org/abs/cmp-lg/9409011> (accessed on June 16, 08)

²⁶ <http://arxiv.org/pdf/cmp-lg/9609009> (accessed on June 18, 08)

²⁷ <http://www.cs.cmu.edu/~abberger/mt.html> (accessed on June 18, 08)

1.1.2.2. Rule Based Method²⁸

The rule-based system starts with a rule-base, which contains all of the appropriate knowledge encoded into If-Then rules, and a working memory, which may or may not initially contain any data, assertions or initially known information. The system examines all the rule conditions (IF) and determines a subset, the conflict set, of the rules whose conditions are satisfied based on the working memory. Of this conflict set, one of those rules is triggered (fired). Which one is chosen is based on a conflict resolution strategy. When the rule is fired, any actions specified in its THEN clause are carried out. These actions can modify the working memory, the rule-base itself, or do just about anything else the system programmer decides to include. This loop of firing rules and performing actions continues until one of two conditions is met: there are no more rules whose conditions are satisfied or a rule is fired whose action specifies the program should terminate.

1.2. Printed Lexicon vs. Electronic Lexicon

Before going to compare the similarities and differences it would be better to know why lexicons are given more importance in natural language processing. Lexical knowledge, i.e. knowledge about individual words in the language, is essential for all types of natural language processing. Developers of machine translation systems, which from the beginning have involved large vocabularies, have long recognized the lexicon as a critical system resource. As researchers and developers in other areas of natural language processing move from toy systems to systems which process real texts over broad subject domains, larger and richer lexicons will be needed and the task of lexicon design and development will become a more central aspect of any project.²⁹

1.2.1. Printed Lexicon

A **Lexicon** is a book of alphabetically listed words in a specific language, with definitions, etymologies, pronunciations, and other information. Besides this, a book of alphabetically listed words in one language with their equivalents in another may be a lexicon too.³⁰ According to

²⁸ <http://ai-depot.com/Tutorial/RuleBased.html> (by James Freeman-Hargis) (accessed on June 19, 08)

²⁹ <http://cslu.cse.ogi.edu/HLTsurvey/ch12node6.html> (accessed on May 15, 08)

³⁰ Webster's New World College Dictionary, Fourth Edition, 2002

MSN Encarta dictionary³¹ a reference book that alphabetically lists words and their meanings; or the entire stock of words belonging to a branch of knowledge or known by somebody is called a lexicon.

From linguistic point of view, the lexicon of a language is its vocabulary, including its words and expressions. More formally, it is a language's inventory of lexemes, which is an abstract unit of morphological analysis in linguistics that roughly corresponds to a set of forms taken by a single word. For example, *paṭh*, *paṭhati*, *paṭhan* etc. are the forms of the same lexeme, conventionally written as *Paṭh*. Lexemes are formed according to morpho-syntactic rules and express sememes. In this sense, a lexicon organizes the mental vocabulary in a speaker's mind: First, it organizes the vocabulary of a language according to certain principles (for instance, all verbs of motion may be linked in a lexical network) and second, it contains a generative device producing (new) simple and complex words according to certain lexical rules.³²

A lexeme denotes: Headwords, Morphology, Phonetics, Syntax (sub categorization), Genus/hierarchy/thesaural category, Domain (sport, law etc.), Meaning and Pragmatics.³³ A great deal of effort has been invested over centuries in the development of printed monolingual and bilingual lexicons (also thesauri and encyclopedias) for human use. It would be very helpful for computational linguistics to automate this large body of information. Extensive efforts have been made in this direction. However, conventional dictionaries often do not contain the right sort of information (e.g. detailed sub-categorization information), and/or do not encode what they do hold in a sufficiently explicit, consistent, or even accurate way to be readily automated. By 'lexicons', *made for use by humans* which are about words, the main content of which is divided into articles ('entries') each of which is about a word or group of related words. This formulation of 'lexicon' includes:³⁴

- Standard definitional dictionaries such as Merriam-Webster (1963) or Larousse (1971).
- Bilingual dictionaries (although such a work as an English-Sanskrit/Sanskrit-English dictionary is in fact two lexicons bound in a single volume).
- Thesauruses.

³¹ http://encarta.msn.com/dictionary/_/lexicon.html (accessed on May 15, 08)

³² <http://en.wikipedia.org/wiki/Lexicon> (accessed on May 15, 08)

³³ <http://www.cs.man.ac.uk/~mary/CS3421lectures/node14.html> (accessed on May 15, 08)

³⁴ <http://interglacial.com/~sburke/ma/online.html#lexicons> (accessed on May 15, 08)

- Phonetic dictionaries like rhyming dictionaries or pronouncing dictionaries.
- Orthographic dictionaries like shorthand dictionaries, or crossword puzzle dictionaries. (Although these are unusual in that the entry for a given word consists generally of just the headword itself.)
- More encyclopedic dictionaries like: ethnographic dictionaries (e.g., Franciscan Fathers 1910) or dictionaries of specialized fields of knowledge (e.g., Howe 1994).

1.2.2. Electronic Lexicon

A lexicon in the Internet mainly fulfils the same service as a printed lexicon provides: it explains the meaning of words and gives background information. In many cases Internet lexicons contain more and more recent content than printed ones. As Internet lexicons also profit from link and search functionality, they can be quite comfortable to use.³⁵ An online lexicon tends to contain more terms connected to everyday's life and lifestyle than a printed lexicon. As the production of an online lexicon is more flexible they can be quite more up-to-date than the recent version of a printed lexicon. Very often new events or terms are described on the very day they happened or were created.³⁶

1.3. Electronic Indices vs. Search engines

1.3.1. Electronic Indices

1.3.1.1. General Information

An **Index** is a systematic guide to items contained in, or concepts derived from a collection. These items or derived concepts are represented by entries in a known or stated searchable order, such as alphabetical, chronological, or numerical.³⁷

Indexing is the process of analyzing the informational content of records of knowledge and expressing the informational content in the language of the indexing system. It involves:

- a) selecting indexable concepts in a document; and

³⁵ <http://en.web-fortbildung.de/glossary/Lexicon.php> (accessed on May 16, 08)

³⁶ http://en.web-fortbildung.de/search_spots.php (accessed on May 16, 08)

³⁷ Indexing Concepts and Methods (pg. 8)

- b) expressing these concepts in the language of the indexing system (as index entries); and an order list.

An **indexing system** is the set of prescribed procedures (manual/machine) for organizing the contents of records of knowledge for purposes of retrieval and dissemination.

1.3.1.2. Computer-Aided Indexing

Automatic text analysis became a subject for serious research in the late 1950s and a major focus of activity within the then emerging field of information science. Many streams converged to create this strong new current. First there was the desire to provide faster and more complete access to the scientific and engineering literature and “to control the information explosion.” A second factor was the availability of computers and the recognition that these machines were symbol manipulators capable of processing words as well as numbers. A third influence was the emergence of a new field of study called computational linguistics, that is, the application of computer analysis to the structure and meaning of language, led by such men as Noam Chomsky (1956, 1957) and Zellig Harris (1962). Still a fourth thrust was the research and advances in the fields of artificial intelligence and self-organizing systems (‘thinking machines’). The National Science Foundation (USA) supported these efforts to nurture the development of information science. Studies on automatic indexing blossomed during this period, as did studies on automatic abstracting, machine translation, and various other information-processing and retrieval activities.³⁸

The literature on automatic indexing is voluminous. Two excellent state-of-the-art reviews have been prepared by Stevens (1970) and Sparck-Jones (1974a,b). In addition, there is the series *Annual Review of Information Science and Technology*, as well as a number of relevant books (*Machine Indexing*, 1961; Borko, 1967; Salton, 1971).

1.3.1.3. Automatic Indexing Methods

The differences between word indexing and subject indexing are crucial for evaluating machine-indexing techniques. Word indexing has also been called *derived indexing*, as contrasted with subject indexing, which is *assigned indexing*. Word indexes are derived from the words used by

³⁸ Indexing Concepts and Methods (pg. 113)

the authors of a work, and they guide to words and their contexts. Subject headings are assigned and modifications coined (using by a human indexer) using words (generally those of the author) that best lead to the subject matter of the work and not to extraneous material. Subjects are expressed in words, of course, but subject indexes are quite different from word indexes.

The foundation on which automated language analysis rests is that ideas are communicated by words (written or spoken) and their arrangement. Stemming from this fundamental axiom is the hypothesis that the subject of a document can be derived by a mechanical analysis of the words in a document and by their arrangement in a text. In the current state-of-the-art, fully automatic machine indexing is derived or word indexing. However, subject indexing may be approximated (a) by using some human intervention, as in machine-aided indexing; and (b) possibly by use of advanced linguistic techniques, including automatic syntactic and semantic analysis.

For computational purposes, a word is defined as a sequence of symbols, alphabetic, numeric, or punctuative, separated by spaces on both sides. Assuming that the words are in machine-readable code, the computer can be programmed to process these words in a variety of ways. Programs can count the number of words in a document and calculate the mean number of words in the average sentence of the document. Programs can count the frequency with which a given word appears in a document or in the entire database and can print lists of words, ordered alphabetically or by frequency of occurrence.

Truncation (removal of prefixes and suffixes from the stems of words) makes it possible to count all appearances of the same stem as one word type. For example, programs can combine the counts of the stem *index* with the counts of *indexes*, *indexer*, *indexers*, and *indexing*. Programs can also count the number of times that the word *indexing* is preceded by the word *automatic*, and can search titles and abstracts for a selected word or phrase and obtain a list of all documents in which these terms occur. An almost unlimited number of kinds of counts can be made; the problem is to select only those word counts that prove useful in automating the assignment of index headings or in determining the subject of a document-if, indeed, word counts alone can do this. The solution of this problem has been a major concern of researchers. Three basic methods of automatic indexing have been studied (a) statistical analysis of text; (b) syntactic analysis; and (c) semantic and discourse analysis.³⁹

³⁹ Indexing Concepts and Methods (pg. 114-15)

1.3.1.3.1. Statistical analysis⁴⁰

One hypothesis underlying the statistical method of indexing is that the more times a word is an indicator of the subject matter. Based upon this hypothesis, a computer program lists all of the words in a document; the words are grouped by number of occurrences and arranged alphabetically within each frequency. Function words (such as articles, conjunctions, prepositions, and pronouns) are usually excluded. Words having the same stem can be counted either as the same or as different words.

Counting words is basic to all techniques of machine indexing and was 1st suggested by Luhn (1957). At the simplest and most un-sophisticated level, the computer is programmed to select as index terms all words on the list that have been used more than a specified minimum number of times in the work being indexed. The application of this rule may result in the selection of large numbers of index terms, most of which will provide poor guidance to subjects reported by the author. This result is not surprising because indexing is much more complicated than merely selecting all words used more than a minimum number of times. Certainly, human indexes do not use as a criterion for index selection the number of times a term is used in the text, and authors try to avoid word repetition. Thus, word counts cannot be used as a sole basis for selection. If the cutoff number of words is set too high, for example, 10 to 12 repetitions, then many useful index headings will be eliminated; if it is set too low, for example, 1 or 2 repetitions, many terms useless as subject guides will be included.

Machine Indexing (1961), Stevens (1970), and Borko (1967) describe a number of research studies aimed at providing more effective procedures for automatically selecting index terms. Some of the more pertinent suggestions are:

- i. *Weighting by location:* For example, a word appearing in the title might be assigned a greater weight than a word appearing in the body of the work.
- ii. *Relative frequency weighting:* This is based upon the relation between the number of times the word is used in the document being indexed and the number of times the same word appears in a sample of other documents.
- iii. *Use of noun phrases:* Only nouns and adjective-noun phrases are used as index terms, and these are selected from the title or abstract.

⁴⁰ Indexing Concepts and Methods (pg. 115-16)

- iv. *Use of thesaurus*: In addition to combining words of the same stem, a thesaurus is used to combine synonyms and otherwise related words. In this way, the count of some word types is increased, as is the separation between “good” and “poor” index terms.
- v. *Use of association factors*: By means of statistical association and correlation techniques, the degree of term relatedness, that is, the likelihood that two terms will appear in the same document, is computed and used for selecting index terms.
- vi. *Maximum-depth indexing*: This procedure indexes a document by all of its content words and weights these words, if desired, by the number of occurrences in the document. In this way, the problem of selecting terms is avoided.

1.3.1.3.2. Syntactic and Semantic Analysis⁴¹

All of the forgoing techniques of statistical analysis were developed by the mid-1960s and formed the basis for most automatic indexing studies. Statistical analysis identifies frequently used words and, by hypothesis, significant content words. Syntactical analysis identifies the role of the word in the sentence, that is, its grammatical class (e.g., whether the word *work* is used as a noun or as a verb) and the relation among the words in the sentence (*dog bites man* vs. *man bites dog*). Semantic analysis helps to establish the paradigmatic or class relations among terms so as to associate words with simple concepts.

The linguistic model proposed by Chomsky distinguishes between SS and DS of language. For instance, the two sentences “Mary went home with John” and “Mary and John went home together” have different SSs but the same DS. By means of transformational grammar, a sentence can be changed; it can go through a series of transformations that will exhibit its DS. It is Chomsky’s view that a purely syntactic transformation can provide a semantic interpretation of the sentence. As is probably obvious, this type of sentence parsing, especially in automatic text analysis, is complex. Sparck-Jones concludes her discussions of syntactic indexing by saying,

It must be allowed that syntactic descriptions may be of value in specific contexts, but there is no hard evidence to the effect that they are generally of value However, it may be that the correct way of providing syntactic information in document descriptions has not yet been discovered, and that when it has, it may be worth writing, or attempting to write, programs to provide it automatically [Sparck-Jones, 1974a, 3,9].

⁴¹ Indexing Concepts and Methods (pg. 116-17)

The goal of semantic analysis as applied to document processing is to identify the subjects and content-bearing words of the document or surrogate text. A number of procedures have been studied, including (a) keyword normalization- stripping words of prefixes and suffixes; (b) dictionary or thesaurus reference, in which the extracted word is looked up in a thesaurus; (c) various classification techniques aimed at grouping related words. The value of these techniques of semantic analysis for automatic indexing is, at best, unproven; much more experimentation is needed.

1.3.2. Search engines

A **search engine**⁴² is an information retrieval system designed to help find information stored on a computer system. Search engines help to minimize the time required to find information and the amount of information which must be consulted, similar to other techniques for managing information overload. The most public, visible form of a search engine is a **Web search engine**⁴³ which searches for information on the WWW. Here, information may consist of web pages, images and other types of files. Some search engines also mine data available in newsgroups, databases, or open directories. Unlike Web directories, which are maintained by human editors, search engines operate algorithmically or are a mixture of algorithmic and human input.

1.3.2.1. History of Search engines

The very first tool used for searching on the Internet was *Archie*⁴⁴ which was created in 1990 by Alan Emtage, a student at McGill University in Montreal. The program downloaded the directory listings of all the files located on public anonymous FTP sites, creating a searchable database of file names. The rise of *Gopher*⁴⁵ (created in 1991 by Mark McCahill at the University of Minnesota) led to two new search programs, Veronica and Jughead. Like Archie, they searched the file names and titles stored in Gopher index systems. Veronica provided a keyword search of most Gopher menu titles in the entire Gopher listings. Jughead was a tool for obtaining menu information from specific Gopher servers.

⁴² http://en.wikipedia.org/wiki/Search_engine_%28computing%29 (accessed on June 07, 08)

⁴³ http://en.wikipedia.org/wiki/Search_engine (accessed on June 07, 08)

⁴⁴ "Internet History - Search Engines" (from Search Engine Watch), Universiteit Leiden, Netherlands, September 2001. Web- <http://www.internethistory.leidenuniv.nl/index.php3?c=7> (accessed on June 07, 08)

⁴⁵ http://en.wikipedia.org/wiki/Gopher_%28protocol%29 (accessed on June 07, 08)

The first Web search engine was *Wandex*,⁴⁶ a now-defunct index collected by the World Wide Web Wanderer, a web crawler developed by Matthew Gray at Massachusetts Institute of Technology in 1993. Another very early search engine, *Aliweb*,⁴⁷ also appeared in 1993, and still runs today. *JumpStation*⁴⁸ (released in early 1994) used a crawler to find web pages for searching, but search was limited to the title of web pages only.

One of the first "full text" crawler-based search engines was *WebCrawler*,⁴⁹ which came out in 1994. Unlike its predecessors, it let users search for any word in any webpage, which became the standard for all major search engines since. It was also the first one to be widely known by the public. Also in 1994 *Lycos*⁵⁰ (which started at Carnegie Mellon University) was launched, and became a major commercial endeavor. Soon after, many search engines like *Magellan*, *Excite*, *Infoseek*, *Inktomi*, *Northern Light*, *AltaVista* and *Yahoo!* appeared and vied for popularity.

Around 2000, the *Google*⁵¹ search engine rose to prominence. The company achieved better results for many searches with an innovation called PageRank. This iterative algorithm ranks web pages based on the number and PageRank of other web sites and pages that link there, on the premise that good or desirable pages are linked to more than others. Google also maintained a minimalist interface to its search engine. In contrast, many of its competitors embedded a search engine in a web portal. By 2002, Yahoo was providing search services based on Inktomi's search engine. *Yahoo!* switched to Google's search engine until 2004, when it launched its own search engine based on the combined technologies of its acquisitions.⁵² Microsoft first launched MSN Search (since re-branded Live Search) in the fall of 1998 using search results from Inktomi. In 2004, Microsoft began a transition to its own search technology, powered by its own web crawler (called msnbot).

As of late 2007, Google was by far the most popular Web search engine worldwide.⁵³ A number of country-specific search engine companies have become prominent; for example *Baidu*⁵⁴ is the

⁴⁶ http://en.wikipedia.org/wiki/World_Wide_Web_Wanderer (accessed on June 07, 08)

⁴⁷ <http://en.wikipedia.org/wiki/Aliweb> (accessed on June 07, 08)

⁴⁸ <http://en.wikipedia.org/w/index.php?title=JumpStation&action=edit&redlink=1> (accessed on June 07, 08)

⁴⁹ <http://en.wikipedia.org/wiki/WebCrawler> (accessed on June 07, 08)

⁵⁰ <http://en.wikipedia.org/wiki/Lycos> (accessed on June 08, 08)

⁵¹ http://en.wikipedia.org/wiki/Google_Search (accessed on June 08, 08)

⁵² http://en.wikipedia.org/wiki/Search_engine (accessed on June 08, 08)

⁵³ <http://searchengineland.com/070921-105613.php> (accessed on June 10, 08)

⁵⁴ <http://en.wikipedia.org/wiki/Baidu> (accessed on June 10, 08)

most popular search engine in China and *guruji.com*⁵⁵ in India. The table mentioned below shows the current market share of some of the famous search engines:

| Most popular search engines worldwide (Dec. 2007) ⁵⁶ | | |
|---|----------------------|-----------------------|
| Company | Millions of searches | Relative market share |
| Google | 28,454 | 46.47% |
| Yahoo! | 10,505 | 17.16% |
| Baidu | 8,428 | 13.76% |
| Microsoft | 7,880 | 12.87% |
| NHN | 2,882 | 4.71% |
| eBay | 2,428 | 3.9% |
| Time Warner (includes AOL) | 1,062 | 1.6% |
| Ask.com and related | 728 | 1.1% |
| Yandex | 566 | 0.9% |
| Alibaba.com | 531 | 0.8% |
| Total | 61,221 | 100.0% |

[Table 1.1: Most popular search engines]

1.3.2.2. How does the Search engine work?

Web search engines work by storing information about many web pages, which they retrieve from the WWW itself. These pages are retrieved by a *Web crawler* (sometimes also known as a spider), an automated Web browser which follows every link it sees. Exclusions can be made by the use of robots.txt. The contents of each page are then analyzed to determine how it should be indexed (for example, words are extracted from the titles, headings, or special fields called Meta tags). Data about web pages are stored in an index database for use in later queries. Some search engines, such as Google, store all or part of the source page (referred to as a cache) as well as information about the web pages, whereas others, such as AltaVista, store every word of every page they find. This cached page always holds the actual search text since it is the one that was actually indexed, so it can be very useful when the content of the current page has been updated

⁵⁵ <http://en.wikipedia.org/wiki/Guruji.com> (accessed on June 10, 08)

⁵⁶ <http://www.comscore.com/press/release.asp?press=2018> (accessed on June 10, 08)

and the search terms are no longer in it. This problem might be considered to be a mild form of linkrot, and Google's handling of it increases usability by satisfying user expectations that the search terms will be on the returned webpage. This satisfies the principle of least astonishment since the user normally expects the search terms to be on the returned pages. Increased search relevance makes these cached pages very useful, even beyond the fact that they may contain data that may no longer be available elsewhere.

When a user enters a query into a search engine (typically by using key words), the engine examines its index and provides a listing of best-matching web pages according to its criteria, usually with a short summary containing the document's title and sometimes parts of the text. Most search engines support the use of the Boolean operators AND, OR and NOT to further specify the search query. Some search engines provide an advanced feature called proximity search which allows users to define the distance between keywords.

The usefulness of a search engine depends on the relevance of the **result set** it gives back. While there may be millions of webpages that include a particular word or phrase, some pages may be more relevant, popular, or authoritative than others. Most search engines employ methods to rank the results to provide the "best" results first. How a search engine decides which pages are the best matches, and what order the results should be shown in, varies widely from one engine to another. The methods also change over time as Internet usage changes and new techniques evolve.

The majorities of search engines are run by private companies using proprietary algorithms and closed databases, though some are open source.⁵⁷ A recent enhancement to search engine technology is the addition of geocoding and geoparsing to the processing of the ingested documents being indexed, to enable searching within a specified locality (or region). Geoparsing attempts to match any found references to locations and places to a geospatial frame of reference, such as a street address, gazetteer locations, or to an area (such as a polygonal boundary for a municipality). Through this geoparsing process, latitudes and longitudes are assigned to the found places, and these latitudes and longitudes are indexed for later spatial query and retrieval. This can enhance the search process tremendously by allowing a user to search for documents

⁵⁷ http://en.wikipedia.org/wiki/List_of_search_engines#Open_source_search_engines (accessed on June 10, 08)

within a given map extent, or conversely, plot the location of documents matching a given keyword to analyze incidence and clustering, or any combination of the two.⁵⁸

1.4. Electronic Indices for Sanskrit

1.4.1. Why for Sanskrit Text?

India has seen amazing strides in ICT applications for Indian Languages in general and for Sanskrit in particular. Since MT from Sanskrit to other Indian Languages is often the desired goal, traditional Sanskrit lexicography and indexing (*anukramaṇī*) has attracted a lot of attention of ICT and CL community. While several attempts are being made to build word-nets and indices on traditional Indian epistemological and logical principles, there have been no serious efforts for generating a cross-referential indexing system to create lexical resource for *Mahābhārata*, the ocean of Indian heritage.⁵⁹

The *Mahābhārata* extols its greatness itself in the following words: “*Yadihāsti tadanyatra yannehāsti na tat kvacit*”. The saying “*Vyāsocchiṣṭam jagatsarvam*” also stresses this point. *Mahābhārata* is neither a history in the modern sense of the term, nor a chronicle. But it stands in incomparable isolation, defying all definitions. It is a veritable encyclopedia comprising of heterogeneous material from all branches of knowledge. Taking the core-story of the feud between two branches of a royal family and the circumstances leading to a catastrophic war, several branches of knowledge including philosophy, law, ethics, statecraft, warfare, history and ethnology are embodied in its structure. The *Mahābhārata* is a comment on the human condition with all its richness, complexity and subtlety. The *Mahābhārata* is the source of many compositions like *Kālidāsa’s Abhijñāna Śākuntalam*, *Śrīharṣa’s Naiṣadhīya Caritam* etc. It is the text that is most sought for in order to enrich cultural, social and any type of knowledge about Indian civilization. But the huge size of encyclopedic epic makes it virtually very difficult for someone to search a specific keyword in it. The indices thus prepared will constitute separate text in itself due to the size of the *Mahābhārata* and will be of tremendous use to the researchers and users.

⁵⁸ http://en.wikipedia.org/wiki/Local_search_%28Internet%29 (accessed on June 12, 08)

⁵⁹ Jha, Girish N., in proceeding of ASIALEX 2005.

1.4.2. Previous Works

The history of textual indexing is very rich in India. *Śaunaka*, a great scholar of Vedas, made a Vedic index named *Sarvānukramaṇī*. Book Indexing of *Mahābhārata* is an ongoing project at Sanskrit evam Prācyā Vidyāsansthāna, Kurukshetra University. They did first two Parvas (i.e., *Ādiparva* and *Sabhāparva*) indexing of *Mahābhārata* which is based on BORI's critical edition of *Mahābhārata*. A western scholar S.Sörensen created an index of "Names in the Mahābhārata" published from Motilal Banarasidas.

For Sanskrit Informatics, it is necessary to make indices available online. Unfortunately, the area of Electronic Indexing for Indian Heritage has not attracted required attention of computational linguists. Some efforts made in this area, directly and indirectly, are listed here:

1. The important work in the area of *Online Indexing of Indian heritage* has been done in University of Goettingen, Germany. It is only work which is directly related with my ongoing research. It includes a *word indexing* of complete *Mahabharata* in Roman transliteration. The database they have used is in text file. The work is in two styles, one is in alphabetical order and another is in *parva* style. It is available online at http://www.sub.uni-goettingen.de/ebene_1/fiindolo/gretil/1_sanskr/2_epic/mbh/sas/mahabharata.htm
2. Another work has been developed by the Indology Department, University of Wuerzburg, Germany. The name of this work is "Multimedia Database to Sanskrit drama". This work is mainly focused on *word indexing* of Bhāsa's (a great Sanskrit dramatist) drama but also includes *Mudrārākṣasa* of Viśākhadatta. It is available online at <http://www.indologie.uni-wuerzburg.de/bhasa/index.html>
3. Electronic text of the Critical Edition of the *Mahābhārata* is available in downloadable text format of several commonly-used encodings, such as- *Unicode Devanagari*, *Unicode Roman*, *ISCII*, *ASCII*, *Norman* etc. at the home page of Professor John Smith, Cambridge University. It is available at <http://bombay.indology.info/index.html>
4. Maharshi University of Management has created PDF files of all 18 Parvas of *Mahābhārata*. The site is <http://is1.mum.edu/vedicreserve/itihas.htm>
5. Kishari Mohan Ganguli has translated the *Mahābhārata* in English which is available at <http://www.sacred-texts.com/hin/maha/index>

6. Girish Nath Jha, SCSS, JNU has created 'Online Multilingual Amarakosh' using RDBMS techniques. This is available at <http://sanskrit.jnu.ac.in/amara/index.jsp>
7. Linguistics Research Centre of The University of Texas has an online *Ṛgveda* in Romanized transliteration format. This online text is based on the van Nooten & Holland electronic version. This can be seen at- <http://www.utexas.edu/cola/centers/lrc/RV/>
8. <http://sanskritdocuments.org/> has created online versions of Sanskrit documents including Vedas, Mahābhārata and many more in ITX, HTML, PS, XDVNG, GIF and PDF format.
9. "Digital Library in India" provides access to the documents of many languages and areas through the internet with its 21 participating centres for the Govt. of India. It can be accessed at- <http://dli.iiit.ac.in/>
10. A pioneering work has been done by MIRI where the catalogue and a complete collection of photographic facsimiles of the 1144 paper transcripts of Sanskrit manuscripts including some manuscripts of *Mahābhārata*, is now made available. It is a collaborative project of MIRI, French Institute of Pondicherry (IFP) and French School of Asian Studies (EFEO). Among these paper transcripts of Sanskrit manuscripts, one can find all the available records of the specific manuscripts through an indexing system. It can be seen at- http://muktalib.org/access_page.htm
11. A project to translate the full epic into English prose, translated by various translators, began to appear in 2005 from the Clay Sanskrit Library, published by New York University Press. The translation is based not on the *Critical Edition* but on the version known to the commentator *Nīlakaṇṭha*. Currently *Sabhā*, *Āraṇyaka*, *Virāṭa*, *Udyoga*, *Bhīṣma*, *Droṇa*, *Karṇa* and *Śalyaparva* are available. The translation of *Ādiparva* is in the projected list. It is accessible at- <http://www.claysanskritlibrary.org/>
12. The Central Secretariat Library, Department of Culture, has sponsored the *Mahābhārata* Database project for making the great epic available to researchers and common readers. The project proposes to create a database of *Mbh* ślokas in ISCII format, such that these can be transliterated into any Indian language script using C-DAC's GIST technology and stored in a CD-ROM for use by research scholars.⁶⁰

⁶⁰ Annual Report (2000-2001), C-DAC.

Chapter-2

Lexical Resources in Indian Tradition

2.1. Kośa Paramparā

2.1.1. The Nighaṇṭu - The oldest Sanskrit lexicon known so far is the *Nighaṇṭu* which is a vocabulary of Vedic words. The derivation and original meaning of the word *nighaṇṭu* cannot be determined with reasonable certainty. The *Nirukta*, however, which is a commentary on the *Nighaṇṭu* gives the derivation of the word *nighaṇṭu* as follows:

samāmnāyaḥ samāmnātāḥ |... tam imam samāmnāyam nighaṇṭavaḥ iti ācakṣate |

The author of the *Nirukta* gives this derivation as that of *Aupamanyava*.¹ The *Nighaṇṭu*, as transmitted to us, consists of five chapters. The first three chapters are called the *Naighaṇṭuka-kāṇḍa*, the fourth is called the *Naigamakāṇḍa* and the fifth is *Daivatakāṇḍa*. The first *kāṇḍa* deals with synonyms, the second with homonyms and the third gives the names of the deities. The first chapter deals with physical things like earth, air, water and natural objects like clouds, dawn, day and night, etc., the second chapter treats with the physical aspects of human beings like his limbs such as finger, arm and objects and qualities associated with men such as wealth, prosperity, anger, battle etc., and the third chapter deals with abstract qualities such as heaviness, lightness etc.²

The *Nighaṇṭu* is not the work of a single author. From the derivation of the work as given by *Aupamanyava* it seems probable that it is a compilation of different persons who might have been the sages of older days. Most probably it is a work of generations of ancient Vedic scholars. One tradition ascribed the authorship of the *Nighaṇṭu* to the sage *Kāśyapa* relying on the verses found in the *Mahābhārata*.³ This, however, appears to be improbable.

One can, therefore, be justified in saying that the *Nighaṇṭu*, as it is presented to us marks the beginning of the lexicographical literature in Sanskrit. The *Nighaṇṭu* is a mere vocabulary of Vedic words. It does not give the meanings of words and in this sense it cannot be called a work of lexicography in the modern sense of the term. Even so, it must be admitted that the words in the *Nighaṇṭu* are arranged according to specific groups and this arrangement is generally found

¹ *Chandobhyaḥ samāhṛtya samāhṛtya samāmnātāḥ | te nighaṇṭavaḥ eva santaḥ nigamanāt nighaṇṭavaḥ ucyante iti Aupamanyavaḥ-Nirukta.*

² L. Sarup, *The Nighaṇṭu* (1920), Introduction p. 13.

³ The verses occur in the Mokṣaparvan (ch. 342, vv. 86-7) and read as under:

“vṛṣo hi bhagavān dharmāḥ khyāto lokeṣu Bhārata
Nighaṇṭuka padākhyāne viddhi mām vṛṣamuttamam |
Kāpīr-varāṇāḥ śreṣṭhaśca dharmāśca vṛṣa ucyate
tasmāt vṛṣākapim prāha kāśyapo mām Prajāpatiḥ ||”

in later lexicons. It is mainly for this reason that the *Nighaṇṭu* can be regarded as the starting point in the compilation of later Sanskrit works on lexicography.

2.1.2. The Nirukta- The *Nirukta* is a running commentary on the *Nighaṇṭu* is ascribed by *Yāska* (800-700BC). Unlike other commentaries which merely explain the words or passages occurring in the text this commentary gives not only the meanings of the words occurring in the *Nighaṇṭu* but also purports to give the references to the terms as they are used in the Vedic literature. *Yāska*, the author of the commentary, quotes Vedic passages and tries to give the derivation of the words found in the *Nighaṇṭu*. The *Nirukta*, as its subject-matter reveals, is not merely a commentary but contains an extensive original discourse in the form of etymological discussion of the words. Incidentally, the author remarks on the nature and utility of the study of the Vedas, the cosmological functions of the Vedic gods and so forth. As *Yāska*'s main object in writing a commentary on the *Nighaṇṭu* is to give the etymology of every word occurring in the text, he tries to derive every word from its original root.

The *Nirukta* is divided into twelve chapters, each chapter being again sub-divided into *pādas* which range from three to seven in different sections. There are two additional chapters, viz. chapters XIII and XIV, which are known as *pariśiṣṭas* or appendices. It is doubtful whether these two chapters formed a part of the original work. The twelve chapters of the *Nirukta* correspond to the three divisions of the *Nighaṇṭu*. Thus, chapters I-III correspond to the first three chapters of the *Nighaṇṭu* which are known as the *Naighaṇṭukakāṇḍa*. Chapters IV-VI deal with the fourth chapter of the *Nighaṇṭu* known as the *Naigamakāṇḍa* and chapters VII-XII deal with the last chapter of the *Nighaṇṭu* called the *Daivatakāṇḍa*. The first chapter gives the classification of words into four categories viz. a) noun, b) verb, c) preposition and d) particle. The author then discusses the meaning and nature of the last two categories. He explains prepositions as words which bring into prominence the subordinate meaning of nouns and verbs. The particles are divided into three groups- i) comparatives, ii) conjunctives and iii) expletives. The terms are defined by giving a list of the particles of each group, explaining their meanings and illustrating their usage by quotations from the Vedic literature. The author further discusses whether the Vedic-mantras have any meaning at all. The rest of the book is devoted to the exposition of the *Nighaṇṭu*, except the first three *pādas* of the seventh chapter which contain discussions about the nature, number, and classification of deities.

Apart from being a commentary on a Vedic glossary the *Nirukta* has an importance of its own. It is a work of great value not only to philologists but also to the students of Sanskrit language. Traces of later lexicography are to be found in the *Nirukta* as also in the *Nighaṇṭu*, for we find *Yāska* referring to many words from the Vedic literature.

2.1.3. VYĀḌI (before 500 AD)- *Vyāḍi* appears to be a renowned lexicographer as he is very often quoted by several reputed authors like *Hemacandra* and others. Unfortunately, no work of *Vyāḍi* has come down to us and hence we are unable to form any definite opinion regarding his lexicon or his personality or age. From the quotations of *Vyāḍi* that are found in the works of later authors like *Hemacandra* or in the commentaries of *Rāyamukuṭa* or *Maheśvara* on the *Amarakośa*, it appears that *Vyāḍi*'s lexicon was arranged in synonymous groups and also consisted of a chapter on homonyms. Lengthy quotations from *Vyāḍi*, found in *Hemacandra*'s commentary to his own work *Abhidhānacintāmaṇi*, show that *Vyāḍi*'s lexicon must have been a voluminous one.

2.1.4. KĀTYA (before 500 AD)- *Kātya* is also one of the ancient lexicographers who preceded *Amara* and flourished before 500 AD. Like *Vyāḍi*, *Kātya* as well is known only by citations and the original lexicon is now lost to us. *Puruṣottama* in his *Trikāṇḍaśeṣa* has identified *Kātya*, the lexicographer with *Kātyāyana* and *Vararuci*. But this identification appears to be doubtful. *Kātya*'s lexicon seems to have contained both synonyms and homonyms and its name appears to be *Nāmamālā*. *Kātya* does not, like *Amara*, put down the synonyms together but often strives to give accurate meanings by means of descriptive clauses.⁴

Kātya as a lexicographer is quoted by *Kṣīrasvāmin* in his commentary on the *Amarakośa*, by *Hemacandra* in his *Abhidhānacintāmaṇīṭikā*, by *Keśava* in *Kalpadrakośa* (composed in 1660 AD), by *Rāyamukuṭa* and *Bhānujī* in their commentaries on the *Amarakośa*. He is also referred to by *Maṅkha* as one of his authorities in his *Anekārthakośa*.

2.1.5. AMARAKOŚA- The *Amarakośa* ascribed by *Amarasimha* (before 6th century AD), as far as the Sanskrit lexicography is concerned, is regarded as a work of paramount authority.

⁴ *Kalpadrakośa*, Introduction, p. xiii.

Numerous commentators on different Sanskrit works have frequently quoted *Amara*'s lexicon in support of the explanation of a particular word given by them while commenting on any Sanskrit text. In India, the *Amarakośa* has the widest circulation and is regarded as a work of unquestionable authority in all the different schools and sects. The popularity of this work can also be determined by the fact that Dr. Aufrecht records not less than forty commentaries on it in his '*Catalogus Catalogorum*.'

This lexicon is popularly known by the name *Nāmaliṅgānuśāsanam*, meaning thereby a work which deals with vocables and their genders. It is also known as *Trikāṇḍa* and is divided into three *kāṇḍas*, each *kāṇḍa* again being sub-divided into sections called the *vargas*. The table of contents is given below:

| Svargādikāṇḍa | Bhūmyādikāṇḍa | Sāmānyakāṇḍa |
|---------------------|-------------------|-------------------------|
| 1. Svargavarga | 1. Bhūmivarga | 1. Viśeṣyanighnavarga |
| 2. Vyomādivarga | 2. Pūravarga | 2. Saṁkīrṇavarga |
| 3. Kālavarga | 3. Śailavarga | 3. Nānārthavarga |
| 4. Dhīvarga | 4. Vanaṣaddivarga | 4. Avyayavarga |
| 5. Śailādivarga | 5. Siṁhādivarga | 5. Liṅgādisaṁgrahavarga |
| 6. Nātyavarga | 6. Nṛvarga | |
| 7. Pātālabhogīvarga | 7. Brahmavarga | |
| 8. Narakavarga | 8. Kṣatriyavarga | |
| 9. Vārivarga | 9. Vaiśyavarga | |
| | 10. Śūdravarga | |

[Table 2.1: Structure of *Amarakośa*]

The *Amarakośa* is arranged in a metrical form in *anuṣṭubh* metre. A major part of it deals with synonyms and only one small section viz., the *nānārthavarga* is devoted to homonyms and is arranged after the final consonants. Indeclinables are treated in one chapter while the last section is devoted to the general rules for determining the genders. The arrangement of the work is faulty and one finds it extremely difficult to trace a particular word in the *kośa* without the help of an index. The genders of the words are expressed by the inflexional endings in some cases, while at times they are recorded with words such as *strī*, *pūṁs* etc. indicative of genders.

The *Amarakośa*, though the most popular of all the Sanskrit lexicons, is not the first of its kind. Leaving aside the question of Vedic glossaries such as the *Nighaṇṭu* and a few others, it is found that there existed a number of lexicons prior to the *Amarakośa*. The author of the text himself states that he compiled his lexicon after consulting other treatises.⁵ In the opinion of *Sarvānanda*, *Amara* seems to have consulted the works of *Vyāḍi*, *Vararuci*, *Trikāṇḍa* and *Utpalinī*.⁶

2.1.6. DHANVANTARINIGHAṆṬU- The *Dhanvantarinighaṇṭu* is a glossary of materia medica and is attributed to *Dhanvantari*⁷ (before AD 500). According to one recension it consisted of seven chapters while according to another it contained nine. There also appears to be a third recension, which regards the text as consisting of ten chapters. Another name of this text is *Dravyāvalīnighaṇṭu*. As the text of the *Dravyāvalī* known to us agrees with that of the *Dhanvantarinighaṇṭu*, one may presume the two works to be identical, although they slightly differ from one another in the different portions of the text.

The work is divided into the 7 vargas or sections (*Guḍūcyādi-Śatapuspādi-Candanādi-Karavīrādi-Kāñcanādi-Āmrādi- Miśrakādivarga*). The text deals with the vocabulary of medicinal herbs and plants. Synonyms for herbs and plants together with their properties are described by the author.

2.1.7. VAIJAYANTĪKOŚA- This *kośa* of *Yādavaprakāśa* (before 1100 AD) is a voluminous lexicon consisting of two broad divisions, one containing synonyms and the other containing homonyms. The former (synonyms) is divided into 5 *kāṇḍas* and the latter (homonyms) is divided into 3 *kāṇḍas*. Each of these sections is again subdivided into several chapters. The table of contents is given below:

⁵ *Samāhṛtyānyatantrāṇi saṁkṣiptaiḥ pratisaṁskṛtaiḥ* |

Sampūrṇamucyate vargaiḥ Nāmaliṅgānuśāsanam || *Amarakośa*, 1.1.2.||

⁶ While commenting on *Amara*, 1.1.2, *Sarvānanda* remarks: “*Anyatrāṇi Vyāḍivararuciprabhṛtinām tantrāṇi, samāhṛtya ekīkṛtya ata eva sampūrṇamidam, yataḥ Trikāṇḍotpalinyāḍini nāmamātr tantrāṇi, Vyāḍivararucyādi praṇītāni tu liṅgamātrāṇi*” etc.

⁷ In *Dhanvantarinighaṇṭu* (ASS), Introduction p.2, the editor Vaidya Narayan Shastri Purandare seems to be of the opinion that the work must have been composed by one of the disciples of *Dhanvantari*, who is supposed to be the most ancient authority on Indian medicine, during his life-time (*Ayam granthaḥ Dhanvantaryantevāsināmā kenacidapi chātṛeṇa grathita iti jīvasvāyurvedācāryeṣu Dhanvantarisvasya granthasya viracanā jātetyuditavacanāt suniścitameva*).

| | KĀṇḌA | ADHYĀYA |
|----------|-----------------|---|
| Synonyms | Svargakāṇḍa | 1. Ādidevādhyāya 2. Lokapālādhyāya 3. Yakṣādhyāya |
| | Antarīkṣakāṇḍa | 4. Jyotirādhyāya 5. Meghādhyāya 6. Khagādhyāya 7. Śabdādhyāya |
| | Bhūmikāṇḍa | 8. Deśādhyāya 9. Śailādhyāya 10. Vanādhyāya 11. Paśusaṁgrahādhyāya 12. Manuṣyādhyāya 13. Brāhmaṇādhyāya 14. Kṣatriyādhyāya 15. Vaiśyādhyāya 16. Śūdrādhyāya |
| | Pātālakāṇḍa | 17. Sarīsrpādhyāya 18. Jalādhyāya 19. Purādhyāya 20. Bhūtādhyāya |
| | Sāmānyakāṇḍa | 21. Paṇādhyāya 22. Dharmakarmādhyāya 23. Guṇādhyāya 24. Arthavallīṅgādhyāya |
| Homonyms | Dvayakṣarakāṇḍa | 25. Puṁllīṅgādhyāya 26. Strīlīṅgādhyāya 27. Napuṁsakalīṅgādhyāya 28. Abhidheyavallīṅgādhyāya 29. Nānālīṅgādhyāya |
| | Tryakṣarakāṇḍa | 30. Puṁllīṅgādhyāya 31. Strīlīṅgādhyāya |

| | | |
|--|-----------|--|
| | | 32. Napuṃsakaliṅgādhyāya 33. Abhidheyavalliṅgādhyāya 34. Nānāliṅgādhyāya |
| | Śeṣakāṇḍa | 35. Puṃliṅgādhyāya 36. Strīliṅgādhyāya 37. Napuṃsakaliṅgādhyāya 38. Abhidheyavalliṅgādhyāya (also called Arthavalliṅgādhyāya) 39. Nānāliṅgādhyāya 40. Paryāyasaṃyojanyāyapradarśanādhyāya 41. Anekārthāvyayādhyāya 42. Avyayaparyāyādhyāya 43. Liṅgasaṃgrahādhyāya |

[Table 2.2: Structure of *Vaijayantīkośa*]

The distinguishing feature of the *Vaijayantīkośa* is that it contains numerous words from the Vedic literature, and is considerably more voluminous. Except the *Nighaṇṭu*, there is *no* other lexicon dealing with Vedic words. On account of this feature the *Vaijayantīkośa* has been looked upon as a work of considerable merit and authority.

2.1.8. MAṆKHA KOŚA (Anekārthakośa)- This lexicon, popularly known after its author *Maṅkha* (about 1140 AD), is a homonymous dictionary consisting of about 1007 verses. The words in the *Maṅkhakośa* are arranged in an alphabetical order according to the final consonants and the words are further arranged according to the number of syllables. The indeclinables are placed last. The lexicon has no division into chapters or *vargas*. The verses are written continuously without a break. The author of this lexicon says that he composed his lexicon by consulting the works of *Bhāguri*, *Kātya*, *Halāyudha*, *Amarasimha* and some others.⁸

⁸ *bhāgurikātyahalāyudhadurgāmarasimhaśāśvatādikṛtān |
kośānnirīkṣya nipuṇam dhanvantarinirmitam nighaṇṭum ca |
liṅgānuśāsanāni ca vicārya lakṣyaṇi mahākavīnām ca |
kurute'nekārthānām śabdānām maṅkhakaḥ kośam ||*

2.1.9. KALPDRUKOŚA- Ascribed by Keśava (AD 1660), this is one of the biggest lexicons, containing about 4000 verses. It is divided into three main divisions known as *skandhas* which are further divided into 27 sub-divisions known as *prakāṇḍas*.⁹ The table of *skandhas* and *prakāṇḍas* is given below:

| Skandhas of Kalpdrukośa | | | |
|--------------------------------|--|---|---|
| | Bhūmiskandha | Bhuvāṣkandha | Svargaskandha |
| | 1. Deśa 2. Pūḥ 3. Narādi 4. Ṛṣigotrābrahma 5. Kṣatriya 6. Vaiśya 7. Śudrādi 8. Viśeṣyanighna 9. Saṁkīrṇa 10. Vanauśadhi 11. Sarabhādi 12. Parvatādi 13. Samudrādi 14. Pātālādi 15. Sarpādi 16. Narakādi 17. Jina | 1. Sādhāraṇa 2. Bhūsthadeva 3. Nabhasthadeva 4. Kāla 5. Nāṭya | 1. Sūryādi 2. Brahma 3. Avyaya 4. Stryādiliṅga |

[Table 2.3: Structure of *Kalpdrukośa*]

It can be seen from this exhaustive list of sub-divisions that the author has tried his best to make each group of synonyms as complete as possible. The method adopted by the author is not very systematic and it is very difficult to find a particular word from this large collection of synonyms

⁹ Cf. *Asminkośe trayāḥ skandhāḥ prakāṇḍaḥ saptavimśatiḥ, ślokāścātuḥsahasrāṇi pañcāśatyā śatadvayam*. -Tanjore Des. Cat. Of Mss., no. 4739.

without the help of an index of words. A large number of abbreviations are used for indicating genders and wherever any member of a compound can be replaced by its synonym the author does not give the compound in its original form but only the word of the compound.

2.1.10. KOŚAKALPATARU- Among the important lexicons that were composed after 15th century, mention must be made of *Kośakalpataru* ascribed to *Viśvanātha* (17th century). This lexicon is one of the largest of its kind and contains more than 5000 verses. It combines both the synonymous and the homonymous characters. The first synonymous part is modeled after the *Amarakośa* and is accordingly divided and sub-divided into major and minor portions such as the *Kāṇḍas* and the *Vargas*. The second homonymous portion is arranged according to the number of letters under each head such as *kavarga*, *Cavarga*, *Ṭavarga*, *Tavarga* and so on. Besides these two main divisions the lexicon contains two portions, one dealing with the genders of vocables and the other with the indeclinables. The table of contents is given below:

| Synonymous | KĀṆḌA | VARGA |
|------------|-------------|--|
| | Svargakāṇḍa | 1. Paramātmā 2. Svarga 3. Vyoma 4. Kāla 5. Dhī 6. Nāṭya |
| Synonymous | Bhūkāṇḍa | 7. Pṛthvī 8. Sura 9. Śaila 10. Vanauṣadhi 11. Simha 12. Nṛ 13. Brahma 14. Kṣatra 15. Vaiśya 16. Śudra |

| | | |
|----------------------|--|--|
| | Pātālakāṇḍa | 17. Pātāla 18. Nāraka 19. Viśeṣyaṇighna 20. Prakīrṇa 21. Dhātu |
| Homonymous | Ka-Ca-Ṭa-Ta-Pa-Antaḥstha-Ūṣmānta Varga | |
| Genders | Līṅgavarga | 1. Puṁlīṅgādhikāra 2. Strīlīṅgādhikāra 3. Napuṁsakalīṅgādhikāra 4. Strīpuṁsaliṅgādhikāra 5. Puṁnapuṁsakalīṅgādhikāra 6. Strīnapuṁsakalīṅgādhikāra |
| Indeclinables | Avyayasāmānyaskandha | |

[Table 2.4: Structure of *Kośakalpataru*]

The detailed analysis of the contents will introduce the voluminous nature of the lexicon of *Viśvanātha*. The work is composed in a variety of meters and occasionally contains prose passages also.

2.1.11. HEMACANDRA (AD 1088-1175)- This celebrated Jain monk of the 12th century is a remarkable figure in the history of Jain and Sanskrit literature. He was a versatile writer and wrote on several branches of Sanskrit and Prakrit literature. He authored four lexicographical works viz., i) *Abhidhānacintāmaṇi* (with its commentary), ii) *Anekārthasaṁgraha*, iii) *Nighaṇṭuśeṣa*, and iv) *Deśnāmamālā*.

2.1.11.1. ABHIDHĀNACINTĀMAṆI- It is a large lexicon containing about 1542 verses written in different meters. The work begins with a description of the *ruḍha*, *yaugika* and *miśra* terms and a note on the component parts of compound words as to which of them are

commutable and which are not. For gender the reader is referred to the author's *Liṅgānuśāsana*.¹⁰ Contents of this lexicon are given here in tabular form:

| | Kāṇḍa | Contents |
|----------|---|---|
| Synonyms | <i>Devādhideva</i> | the Jain gods and the religious terms |
| | <i>Deva</i> | Brahmanical and Buddhistic gods and terms connected with them |
| | <i>Martya</i> | human beings in their different relations |
| | <i>Bhūmi</i> | lower animals |
| | <i>Naraka</i> | the beings of the nether world |
| | <i>Sāmānyakāṇḍa</i> | the abstract notions, adjectives and the indeclinables. |
| Homonyms | It is arranged according to the number of syllables in each word. | |

[Table 2.5: Structure of *Abhidhānacintāmaṇi*]

2.1.11.2. ANEKĀRTHASAMGRAHA- It is a dictionary of homonyms which consists of about 1829 stanzas. It is divided into six *kāṇḍas* (*Eka-dvi-trī-catuḥ-pañca-ṣaṭsvarakāṇḍa*) with an additional supplementary *kāṇḍa* of the indeclinables. In each *kāṇḍa*, the words are arranged in a two-fold alphabetical order, firstly according to the initial letters.¹¹ The words are mentioned without their genders which are to be learnt from *Liṅgānuśāsana*, another work by this author.¹²

2.1.11.3. NIGHANṬUŚEṢA- It is a supplement of *Abhidhānacintāmaṇi* and is a glossary of medicinal plants and herbs. It is a synonymic dictionary and is divided into six *kāṇḍas* (*Vṛkṣa-gulma-latā-śāka-tṛṇa-dhānyakāṇḍa*) and consists of 396 stanzas.

2.1.11.4. DEŚĪNĀMAMĀLĀ- It is a lexicon of *deśī* words. It is composed in *Prākṛta* *Āryās* and gives *Prākṛta* words along with their equivalents in *Prākṛta*. The whole work is divided into 8 divisions called the *vargas*, which are- i) *svavararga* ii) words beginning with gutturals iii) words beginning with palatals iv) words beginning with linguals v) words beginning with dentals

¹⁰ *Kalpadrakośa*, Introduction, xxxi.

¹¹ *Akārādi krameṇādāvatra kādikramo'ntataḥ* |

¹² *Liṅgānuśāsane'smābhiḥ varṇito liṅganirṇayaḥ* |
Ato na grathitaḥ sūtre granthagauravabhīrūbhiḥ ||

vi) words beginning with labials vii) words beginning with liquids *ra* and *la* and viii) words beginning with *sa* and *ha*. Each division has a supplement dealing with words having more than one meaning. The words are arranged according to their meanings and the number of syllables.

According to *Hemacandra*, *deśī* words are those words which are in use since times immemorial (*tamhā aṇādi pāla payatṭabhāṣā viśesao deśī*). Accordingly, this lexicon ought to include only such words as have been used in ancient *Prākṛta* literature, and which cannot be derived from their Sanskrit prototypes. But the rule laid down by the author himself does not seem to have been followed by him in his treatise and we frequently find that the author has included many words as *deśī* which could be traced to their Sanskrit originals.

The importance of this lexicon to the *Prākṛta* literature and language is invaluable. It has the same importance as the *Amarakośa* which is regarded as the basis of Sanskrit lexicography. It contains about 4000 *Prākṛta* words together with their *Prākṛta* equivalents and a systematic study of these words will be a valuable contribution to the *Prākṛta* language.

2.1.12. ANEKĀRTHASAMUCCAYA- This lexicon, popularly known as the *Śāśvatakośa* after its author *Śāśvata* (about the 6th century AD), is a dictionary of homonyms. It is not a complete *kośa* and is not divided into different divisions. The words are arranged in full verses, half verses, and even in quarter verses. There is neither any alphabetical arrangement of the words nor are they arranged according to syllables as is generally found in many lexicons. The arrangement being defective, this lexicon is relatively inconvenient for ready reference. The work consists of 807 verses¹³ and is divided into six sections, the last two of which deal with indeclinables.¹⁴

The *Anekārthasamuccaya*, though a small work in extent, seems to have been the basis of later lexicographers. *Śāśvata*, has very often been quoted by *Kṣīrasvāmin*, by *Vararuci* in *Gaṇaratna-mahodadhī*, by *Mallinātha*, and others. This shows that *Śāśvata*'s lexicon must have obtained a high degree of popularity in later years.

2.1.13. ANEKĀRTHADHIVANIMĀÑJARĪ- This lexicon of *Mahākṣapaṇaka* (before 925 AD), is a homonymous dictionary consisting of 3 chapters (*ślokaḍhikāra- ardhaślokaḍhikāra-*

¹³ *Śāśvatakośa* edited by K.G.Oak, Poona, 1918.

¹⁴ Ibid, p.59: 'athāvyayā nibadhyante Ślokaḍhena savistaram' |

padādhikāra). Some manuscripts, however, add one more chapter viz. *ekākṣarādhikāra*. The work is also known by different names in different manuscripts. Some site its name as *Anekārthadhvanimañjarī* while others as *Anekārthamañjarī* and as *Anekārthapadamañjarī*.

2.1.14. PARYĀYARATNAMĀLĀ- This work of *Mādhavakara*'s (700 AD) is a synonymous medical dictionary of botanical terms and contains the names of plants and herbs which were generally used by physicians for medical purposes in old days. Here, the arrangement of the words is neither systematic, nor is the work divided into sections. The printed edition of the work¹⁵ contains 1754 lines consisting of synonyms (13-1474), homonyms (1475-1641) and *māna* or measures (1642-1754). These are further subdivided into full stanzas, half stanzas, quarter stanzas, and half quarter stanzas; the homonyms are presented according to the number of meanings attached to every word viz. i) words having a single additional meaning, ii) words having two meanings each and iii) words having many meanings.

Paryāyaratnamālā have attained much popularity and is frequently quoted in the works and commentaries of *Sarvānanda*, *Rāyamukūṭa*, *Medinī*, *Bhānujī* and others. The lexicon contains the names of a large number of drugs and plants many of which are not traceable in other glossaries; e.g. *Akṣuka*, *Āṣṭa*, *itakaṭa*, *uccatā*, *ṛkṣagandhā*, *kapāṭacakra*, *kumārajīva*, etc.

2.1.15. PARYĀYAMUKTĀVALĪ- It is a medical glossary written by *Haricarana Sena* on the model of the *Paryāyaratnamālā* of *Mādhavakara*. In the introductory stanzas the author acknowledges his debt to *Mādhavakara*.¹⁶ The whole work is written in a metrical form and is divided into 23 sections (*Sugandhi-madhyagandha-hīnagandha-sāraja-ratna-dhātūpadhātu-madhura-amlā-uttamaśāka-tikkaśāka-puṣpa-līlāphala-kanda-mahāvṛkṣa-madhyamavṛkṣa-hrasvavṛkṣa-latā-śimbīśūkādhanya-trṇadhānya-kṛtānna-pānīya-āvaśyajka-bhautikādivarga*).

2.1.16. ABHIDHĀNARATNAMĀLĀ- This lexicon of *Halāyudha* (950 AD) is a vocabulary of small extent containing about 900 stanzas and is divided into five *kāṇḍas* (*svar-bhūmi-pātāla-*

¹⁵ Edited by Dr. Tarapada Chawdhary, Patna, 1946.

¹⁶ *Pūrvam lokahitāya Mādhavakarābhikhyo bhiṣak kevalam koṣānveṣaṇatātparaḥ pravītatāyurvedaratnākarāt | mālāṁ ratnamayīm cakāra sa yathālabham samuddhṛtya yān sāsmbābhiḥ kamanīyabhaktiracanādvāranyathā grathyate ||*

sāmānya-anekārthakāṇḍa). The first four of these deal with synonyms while the last is devoted to homonyms and the indeclinables. The genders are indicated by giving the declensional forms. The work does not treat the genders so strictly as the *Amarakośa* although in other respects it generally follows the latter, and is composed of a variety of matters.

2.1.17. PURUṢOTTAMADEVA (between 1050 and 1200 AD) - A well known commentator on Pāṇini's *Aṣṭādhyāyī* flourished in Bengal and wrote about five lexical works: i) *Trikāṇḍaśeṣa* or *Amaraviveka* ii) *Hārāvalī* iii) *Varṇadeśanā* iv) *Dvirūpakōśa* and v) *Ekāṣarakośa*. The earliest reference to his work is found in *Sarvānanda*'s commentary on the *Amarakośa*, composed in AD 1159.

2.1.17.1. TRIKĀṆḌAŚEṢA OR AMARAŚEṢA- It is a supplement to the *Amarakośa* and contains words which are omitted in it. Corresponding to the three sections of the *Amarakośa*, this work is also divided into as many *vargas* as in the original. In the introductory portion, the author states that his aim in writing the lexicon was to give only such vocables as were omitted by *Amara* but were found in common use.¹⁷

2.1.17.2. HĀRĀVALĪ- It is a small work consisting of about 270 stanzas. It consists almost exclusively of uncommon words and is divided into synonyms and homonyms. The former is again subdivided into three sections having full verses, half verses and quarter verses. The latter is also divided into similar sections and gives the different meanings attached to the words. In the end, the author states that the composition of *Hārāvalī* was the result of the consultation of several lexicons.¹⁸

2.1.17.3. VARṆADEŚANĀ- As the name of the work indicates, it is a treatise on the proper spelling of nouns with certain cognate consonants such as *kha* and *kṣa*, *ha* and *gha*, *ha* and *ḍa*, *ja* and *ya* etc.

¹⁷ *Alaukitatvādamaraḥ svakośe, Na yāni nāmāni samullilekha |*

Vilokya tairapyadhUnā pracāram, ayaṁ prayatnaḥ Puruṣottamasya ||

¹⁸ *Śabdārṇavautpalinī saṁsārāvarataityapi, kośā Vācaspativyāḍivikramādityanirmitāḥ |*
ādāya sārāmeteṣāṁ anyeṣāṁ ca viśeṣataḥ, Hārāvalī nibaddheyaṁ mayā dvādaśa vatsaraiḥ ||

2.1.17.4. DVIRŪPAKOŚA- It is a small work on lexicography containing about 75 stanzas. It is a vocabulary of words which are spelt in two different ways but which are similar in sound, e.g. the words *āṣādhā* and *āsāḍha*, *śasvara* and *sasvara*, *kuśala* and *kuṣala* etc.

2.1.17.5. EKĀKṢARAKOŚA- As its name indicates, this is a lexicon which contains words of one syllable having different meanings attached to such single letters. In many Sanskrit lexicons, we find a number of words with one syllable having various meanings attributed to them. For instance, in this lexicon the meaning attached to the first four letters of *Devanāgarī* alphabet will be found as: ‘*ā- Vāsudeva*’, ‘*ā- Pītāmbara*’, ‘*i- Kāma*’, ‘*i- Lakṣmī*’.

2.1.18. VIŚVAPRAKĀŚA- This lexicon, which is ascribed to *Maheśvara* (AD 1111), is a dictionary of homonymous words arranged according to the final consonants. The words are further arranged in sub-groups according to syllables which range from one to seven, e.g. *kaikam*, *kadvikam*, *katrikam* and so on. Although the work is arranged after the final letters there is no alphabetical order. The words are arranged mostly at random thereby rendering the reference to words very difficult, although not impossible.

2.1.19. ŚABDABHEDAPRAKĀŚA- It is another vocabulary ascribed to *Maheśvara*, the author of previous lexicon. It is a glossary of nouns which, though identical in meaning, differ more or less in their orthography, and is divided in four parts viz. i) *Nirdeśa*, with *śabdabheda*, ii) *bakārabheda*, iii) *ūṣmabheda*, and iv) *liṅgabheda*. This is a supplement of author’s larger work *Viśvaprakāśa*.

On the *Śabdabhedaprakāśa*, *Jñānavimalagaṇi*, a *Kharataragaccha* line of Jain priests wrote a commentary naming ‘*Śabdabhedaprakāśaṭīkā*’ in 1598 AD. This commentary aims at giving the derivation of every word in the text and also tries to give the etymology of the words wherever possible.

2.1.20. NĀMAMĀLĀ- This lexicon of *Dhanañjaya* (about 1123 AD) is a vocabulary of synonyms. According to India Office Manuscript¹⁹ there is only one *pariccheda* viz., the

¹⁹ India Office Cat., no. 1014.

synonyms and contains 205 verses. There are, however, other manuscripts²⁰ of the same work which consist of two or even three *paricchadas*. Some divide the work into two *paricchadas* - synonyms (containing 205 verses) and homonyms (containing 50 verses).

2.1.21. ANEKĀRTHAKOŚA (Nānārthasaṁgraha)- This lexicon of *Ajayapāla* (before AD 1140) is a lexicon small in extent but of considerable authority. It is a work which deals with various meanings attached to a single word. It contains about 1730 words. The words are arranged after the initial letters without regard to the number of syllables and are divided into chapters according to the extent of meanings, in full, half verses and so on. The indeclinables are placed at the end of each chapter instead of being treated separately as in some lexicons.

2.1.22. EKĀKṢARANĀMAMĀLIKĀ- In the Sanskrit language every word has a meaning and even words having one syllable only were looked upon as having some significance. Different meanings were attached to these monosyllabic words, sometimes the same letter having different meanings. In course of time, the single letters of the alphabet attained prominence and lexicons purely devoted to monosyllables and hence known as the *ekākṣarakośa* were gradually composed. The *ekākṣaranāmamālikā* attributed to *Sudhākalaśa* (about 1350 AD) belongs to the category of the homonymous glossaries of monosyllables. It deals with the letters of the alphabet, giving at the same time the meanings attached to them. It is a very small lexicon consisting of only 50 verses approximately.

2.1.23. MADANAVINODANIGHAṆṬU- This huge lexicon of *Madanapāla* (AD 1375) is also known by its shorter title *Madanavinoda*. It is a famous dictionary of drugs (material medica) containing about 2250 verses and is divided into fourteen sections (*Bhayādi-Śugrādi-Karpurādi-Dhātu-Vanaspati-Phala-Śāka-Drava-Madhura-Dhānya-Anna-Māṁsa-Miśra-ka-Praśastivarga*). This lexicon not only gives synonyms for certain medicinal drugs but it also deals with their qualities and properties as in the *Rājanighaṇṭu*. The author describes the different kinds of food in the chapter called *Annavarga*. Different dishes of flesh of various animals and their effect on health are described in *Māṁsavarga*.

²⁰ *ibid*, no. 1015.

2.1.24. RĀJANIGHANṬU- Authored by *Narahari* (after AD 1375), this lexicon, also known as *Abhidhānacintāmaṇi* or *Nighaṇṭurāja*, is a medical glossary consisting of the names of various herbs and their medical properties. It is almost entirely limited to *Materia Medica*, or the synonyms and properties of various vegetable and mineral products considered to possess medicinal value. It is divided into 23 chapters (*Anūpādi-bhūmi-guḍūcyādi- Śātāvahādi-parpaṭādi-pippalyādi-mūlakādi-Śalmalyādi-prabhadrādi-karavīrādi-āmṛādi-candanādi-suvarṇādi-pānīyādi-kṣīrādi-śalyādi-māṃsādi-manuṣyādi-simhādi-rogādi-satvādi-miśrakādi-ekārthādi-dvayarthādivarga*).

2.1.25. PĀRASĪPRAKĀŚA- Under the patronage of Akbar, *Vihārī Kṛṣṇadāsa* (between AD 1556 and 1605) composed a bilingual glossary dealing with Persian and Sanskrit words. Like the *Pārasīprakāśa* of *Vedāṅgarāya* the present work also deals with Persian equivalents for Sanskrit expressions but differs from the former in its nature. The *Pārasīprakāśa* of *Vedāṅgarāya* deals mainly with astronomical and astrological terms whereas the work of *Vihārī Kṛṣṇadāsa* gives the Sanskrit equivalents of Persian terms in general.

2.1.26. RĀJAVYAVAHĀRAKOŚA- It is a lexicon which was composed for Śivājī by his minister Raghunātha Nārāyaṇa Hanumante (about AD 1676-77). Unlike many other Sanskrit lexicons, it contains many Persian and Arabic terms along with their Sanskrit equivalents and is divided into ten sections (*Rāja-kāryasthāna-bhogya-śāstra-caturaṅga-sāmanta-durga-lekhana-janapada-paṇyavarga*). The author states that at the time of Chhatrapati Shivaji, many of the Sanskrit words became obsolete and Persian and Arabic words came in vogue. For the purpose to do away with the foreign words and to revive the usage of old Sanskrit terms, the present lexicon was composed.

Except the above mentioned lexicons, there are some other important works which have a great role to play in the enrichment of Sanskrit lexical resources. *Bhūriprayoga* of *Padmanābhadatta* (last quarter of 14th century) is a lexicon of synonymous and homonymous words and is professed to be a supplement to the *Amarakośa*. *Abhidhānatantra* of *Jaṭādhara* (before 1431 AD) is a vocabulary of synonymous and homonymous words. *Anekārthadhvanimañjarī* of

Gadasimha (before 1431 AD) is a vocabulary of words having different meanings. *Śrutaśabdasamuccaya* of *Someśvara* (before 1550 AD) is a vocabulary of homonymous words belonging to Vedic literature.²¹ *Śīghrabodhinīnāmamālā* of *Puṇḍarīka Viṭṭhala* (latter half of the 16th century) has four parts which contains 472 stanzas. The first gives the names of the different parts of the body, and also of all things and attributes of women; the second deals with inanimate objects such as mountains etc; the third is devoted to the enumeration of celestial objects; and the fourth to words of several meanings, synonyms, prepositions, names of arts, science etc. *Pañcavargasaṁgrahanāmamālā* of *Śubhaśīla* (between AD 1450 and 1500) is a small lexicon which imitates the *Abhidhānacintāmaṇi* of *Hemacandra* in style, division and general form. *Uṇādināmamālā* of *Śubhaśīla* consists of words having *unādi* suffixes. *Pathyāpathyavibodhanighaṇṭu* of *Kaiyadeva* (before 1648 AD) is a dictionary of material medica and hygiene. It gives a list of numerous articles of food, herbs, plants, vegetables etc, and deals with their medicinal properties stating at the same time which substances are congenial to health. *Pañcatattvaprakāśa* of *Veṇīdatta* (1644) is a metrical glossary containing vocables of five elements (*prthvī-jala-tejas-vāyu-ākāśa*) in 335 verses. *Uṇādinighaṇṭu* of *Veṅkaṭeśvara* (1684-1712 AD) is a small work of about 720 verses in five chapters. It brings together all the words that have been dealt with in *Uṇādisūtras*. *Śabdacandrikā* of *Cakrapāṇidatta* (1060 AD) is a medicinal dictionary with vegetables and mineral substances and also contains a chapter on compounds, both in medicine and dietetics and is divided into nine *vargas*. *Śabdapradīpa* of *Sureśvara* (1075 AD) is a glossary of botanical terms which gives the names of different plants and also mentions their medical properties. It is one of the oldest important botanical glossaries enumerating various plants in contemporary Bhārata. *Dharaṇīkośa* (also called *Anekārthasāra*) ascribed to *Dharaṇīdhara* (before AD 1159), is a vocabulary of homonymous words and arranged after the final letters and number of syllables like the *Medinīkośa*. *Dvirūpakośa* is a small lexicon by *Śrīharṣa* (2nd half of the 12th century) which deals with words having two forms slightly different from one another. Such words may differ in the *mātrās* or syllables or in gender; e.g., *amarṣa* and *āmarṣa*, *agastyah* and *agastiḥ* etc. *Nānārthāṇṇavasamkṣepa* is a very extensive lexicon by *Keśava* (12th or 13th century) which deals with homonymous words consisting of 5800 verses and of 6 *kāṇḍas* according to the number of syllables in a word. The

²¹ Cf. *Vaidikavyavahāreṇa nāmaliṅgapramāṇataḥ | arthaprakāśadīpābhāprabhāpratyayavedanam ||*

words are arranged in an alphabetical order and agree with the *Vaijayantīkośa* in many respects especially in its treatment of Vedic words which are generally not included in the classical lexicons.²² ***Hṛdayadīpaka*** of *Vopadeva* (about 1250 AD) is a medical work being a collection of approved recipes, for the treatment of several acute and chronic diseases. It also contains a glossary of medical technical terms and consists of 176 verses and is divided into 8 *vargas* (*Catuṣpāda-tripāda-dvipāda-ekapāda-dvināma-ekanāma-nānārtha-miśrakavaraga*).

Avyayasamgrahanighaṇṭu of *Śākalyamallabhaṭṭa* (about 1330 AD) is a small lexicon of *avyayas* which contains only 50 stanzas. ***Ekāṣararatnamālā*** is a lexicon of *Mādhava* (about 1350 AD) which deals with single letters of the alphabet and the meanings attached to each letter. ***Anekārthatilaka*** (also called *Nānārthatilaka*) ascribed to *Mahīpa* (before AD 1434) is a dictionary dealing with homonyms which is divided into four sections (*Ekāṣara-Dyakṣara-Tryakṣara-Samkīrṇavarṇakāṇḍa*). ***Śabdaratnākara*** of *Vāmanabhaṭṭa Bāṇa* (about 1400 AD) is a dictionary of synonyms, homonyms and the indeclinables which contains 1050 verses in three sections. ***Śabdacandrikā*** is another small synonymous lexicon by *Vāmanabhaṭṭa Bāṇa* which consists of about 100 stanzas and is divided into 5 sections called *adhikāras* (*svargādīlokapāla-antarikṣa-bhūmi-samudra-manuṣyādhikāra*). ***Nānārtharatnamālā*** of *Irugappa Daṇḍādhinātha* (latter half of the 14th century) is a vocabulary of homonymous words and is divided into 6 *kāṇḍas*. This vocabulary records words of various meanings arranged according to their syllables and final letters.²³ ***Rūpamañjarīmālā*** is a small thesaurus by *Rūpacandra* (AD 1588) consisting 120 verses only, giving synonyms for vocables and is divided into 09 *vargas*. The last *sāmānyavarga* deals with homonymous words and is styled as the *anekārthavarga*. The last but not least, ***Śabdamuktāmahārṇava***, a modern vocabulary prepared for Colebrooke by a Pandit named *Tārāmaṇi* (about AD 1785) is a voluminous dictionary arranged alphabetically and further rearranged so that the words beginning with the same letter are again arranged according to the number of syllables and also according to the final letter of the words.

Since the last two centuries, there have been some important efforts in the direction of building bilingual Sanskrit dictionaries. These works have enriched the lexical resources of Sanskrit

²² *Kalpद्रुकोśa*, Introduction, xxxviii.

²³ Narang, Satyapal (1998), *Saṁskṛta Kośa-sāstra ke Vividha Āyāma*, pg. 159-160.

enormously e.g. Monier Williams' Sanskrit-English Dictionary, Apte's English-Sanskrit Dictionary, MacDonell's Sanskrit-English Dictionary, Wilson's Sanskrit-English Dictionary, Boehtlingk's Sanskrit-German Dictionary, Boehtlingk & Roth's Sanskrit-German Dictionary, Cappeller's Sanskrit-English Dictionary, etc.

2.2. E-indices and computational lexicography

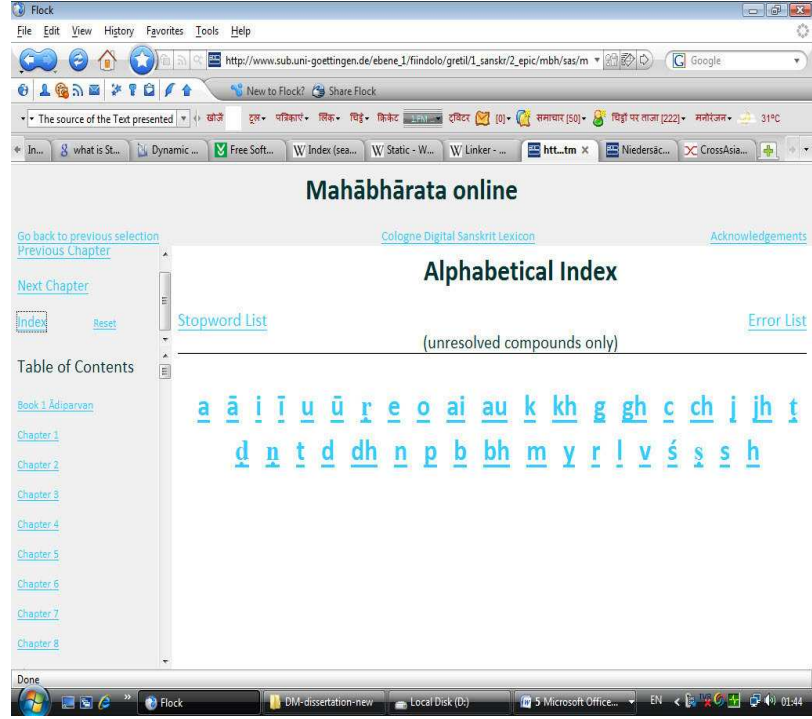
2.2.1. E-indices

The E-index is a computer program which collects, parses, and stores data to facilitate fast and accurate information retrieval. Index design incorporates interdisciplinary concepts from linguistics, cognitive psychology, mathematics, informatics, physics and computer science. An alternate name for the process in the context of e-index designed to catalogue any document on the Internet is **Web indexing**. The e-index comes in two modes, the first static indexing system shows query on the basis of compiled data while the second dynamic indexing system works on running data basis. It means a static indexing system stored the data in a simple text file which gives the result if the query exists in file. Whereas a dynamic indexing system is based on RDBMS technology in which data are stored in several tables and all the tables are adjoined with each other.

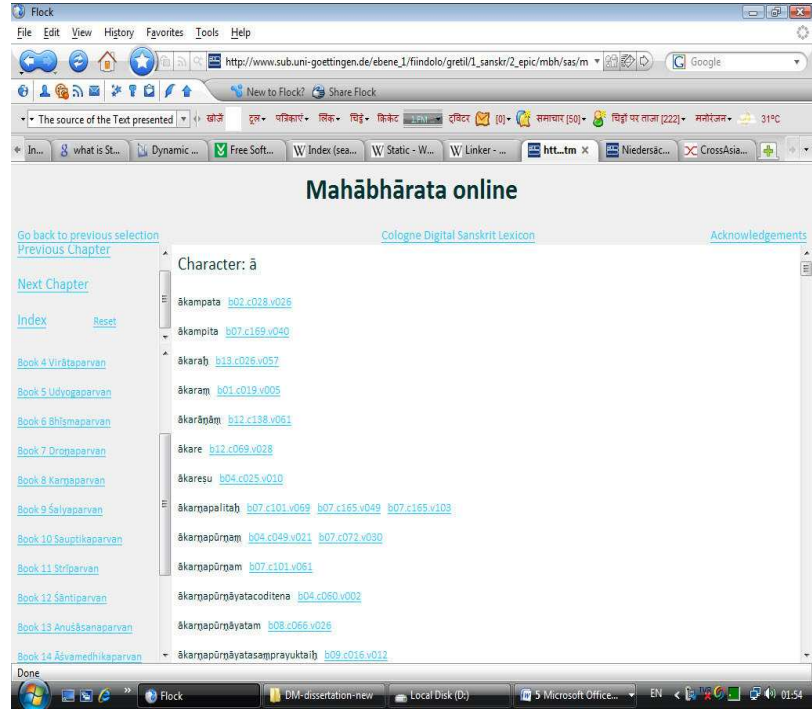
2.2.1.1. E-index of Mahābhārata²⁴

An electronic index for Mahābhārata is developed by University of Goettingen, Germany. The source of the Text presented here is from Mahābhārata Electronic text (C) Bhandarkar Oriental Research Institute, Pune, India, 1999 on the basis of the text entered by Muneo Tokunaga et al., revised by John Smith, Cambridge. Some snap shots of this work are presented below to see how this indexing system of Mahābhārata works:

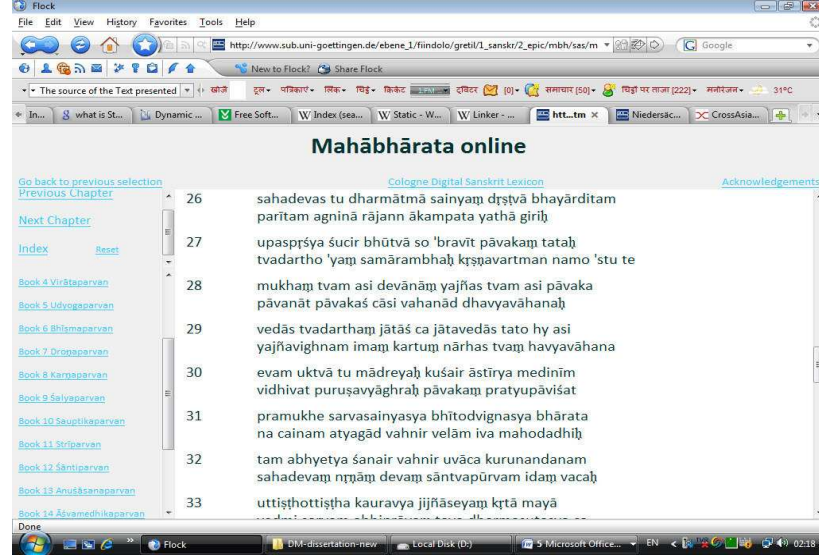
²⁴ http://www.sub.uni-goettingen.de/ebene_1/fiindolo/gretil/1_sanskr/2_epic/mbh/sas/mahabharata.htm (accessed on June 10, 08)



(Fig-2.1: Home page of Mahābhārata Indexer)



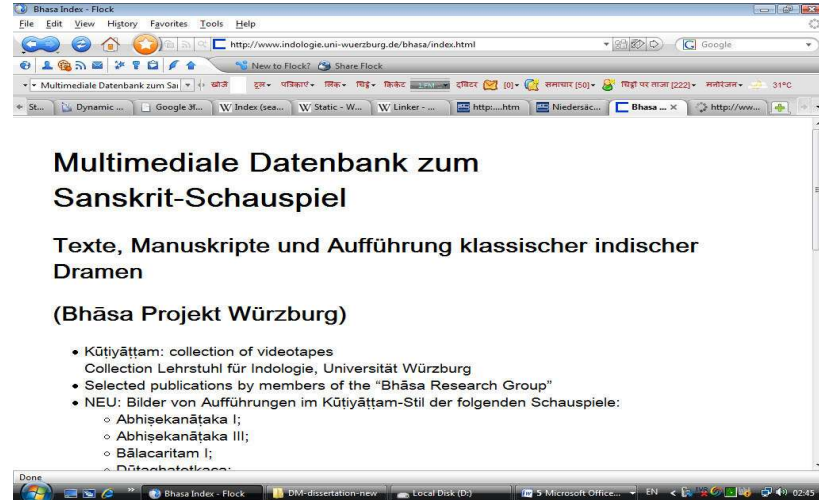
(Fig-2.2: It shows all the words which starts with 'ā')



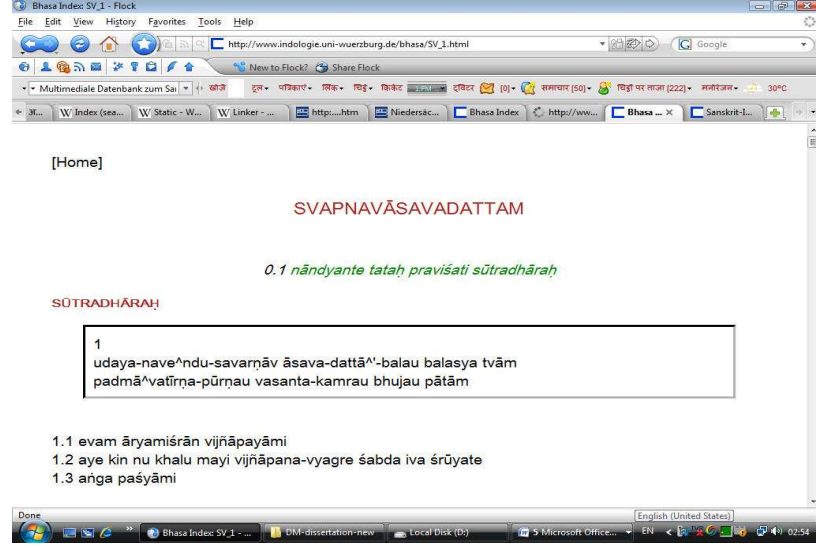
(Fig-2.3: After clicking a specific word, it shows full śloka with its reference)

2.2.1.2. E-index for Sanskrit Dramas²⁵

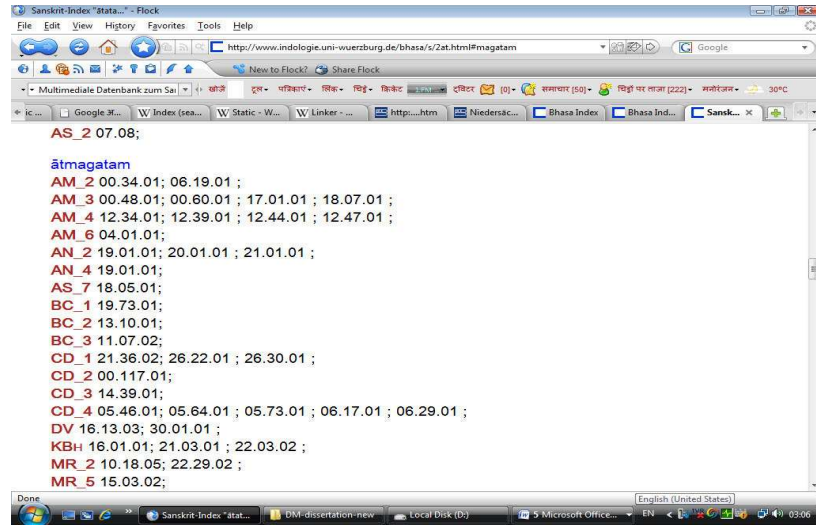
The work 'Multimediale Datenbank zum Sanskrit-Schauspiel' (Multimedia Database to Sanskrit drama) which mainly focused on word indexing of *Bhāsa*'s (a great Sanskrit dramatist) drama and also includes *Mudrārākṣasa* of *Viśākhadatta* and *Abhijñāna Śākuntalam* of *Kālidāsa* has been developed by Indology Department, University of Wuerzburg, Germany. Some snap shots of this work are presented below:

(Fig-2.4: Main page of *Bhāsa* indexer)

²⁵ <http://www.indologie.uni-wuerzburg.de/bhasa/index.html> (accessed on June 15, 08)



(Fig-2.5: clicking any one *nāṭaka* among *nāṭakacakras*, the text comes in hyperlinked format)



(Fig-2.6: clicking any hyperlinked word of a specific *Nāṭaka*, the indexer shows its occurrences in all available *Nāṭakas* with hyperlinked mode)

2.2.2. Computational Lexicography

As automatic natural language processing (NLP) moves out of the era of toy pilot projects and begins to grapple with real-life language in all its complexity, it needs access to quantities of information about individual lexical items. The only plausible source of such information lies in machine-readable versions of ordinary published dictionaries; although they are designed for

other purposes and are far from ideal for computer use, they represent a repository of resources that the computational linguistics research community is in no position to match.²⁶

Computational lexicography emerged as a separate discipline within computational linguistics with the appearance of machine-readable dictionaries, starting with the creation of the machine-readable tapes of the *Merriam-Webster Seventh Collegiate Dictionary* and the *Merriam-Webster New Pocket Dictionary* in the 1960s by John Olney et al. at System Development Corporation. Today, computational lexicography is best known through the creation and applications of WordNet.²⁷

Computational lexicography has contributed to the understanding of the content and limitations of print dictionaries for computational purposes. Many computational linguists were dissatisfied with the print dictionaries as a resource for computational linguistics because they lacked sufficient syntactic and semantic information for computer programs. The work on computational lexicography quickly led to efforts in two additional directions. First, the collaborative activities between computational linguists and lexicographers led to an understanding of the role that corpora played in creating dictionaries. Most computational lexicologists moved on to build a large corpora to gather the basic data that lexicographers had used to create dictionaries. The advent of markup languages led to the creation of tagged corpora that could be more easily analyzed to create computational linguistic systems. Part-of-speech tagged corpora and semantically tagged corpora were created in order to test and develop POS taggers and word semantic disambiguation technology. The second direction was toward the creation of Lexical Knowledge Bases (LKBs). A Lexical Knowledge Base was deemed to be what a dictionary should be for computational linguistic purposes, especially for computational lexical semantic purposes. It was to have the same information as in a print dictionary, but totally explicated as to the meanings of the words and the appropriate links between senses. Many began creating the resources they wished dictionaries were, if they had been created for use in computational analysis. WordNet can be considered to be such a development, as can the newer efforts at describing syntactic and semantic information such as the FrameNet work of Fillmore. Outside of computational linguistics, the Ontology work of artificial intelligence can be seen as an evolutionary effort to build a lexical knowledge base for AI applications.

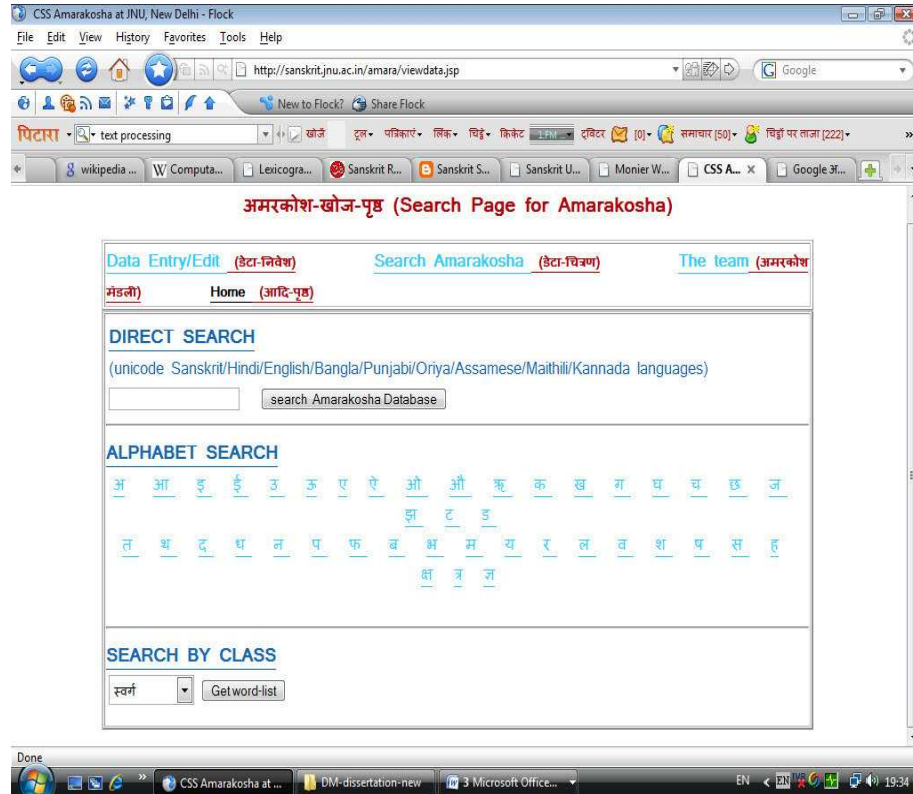
²⁶ <http://www.aclweb.org/anthology-new/J/J90/J90-2004.pdf> (accessed on June 16, 08)

²⁷ http://en.wikipedia.org/wiki/Computational_lexicology (accessed on June 16, 08)

This sub-unit focuses on computational lexicography for Sanskrit Language such as e-lexicon (*kośas*), e-dictionary (mono, bi, or multilingual).

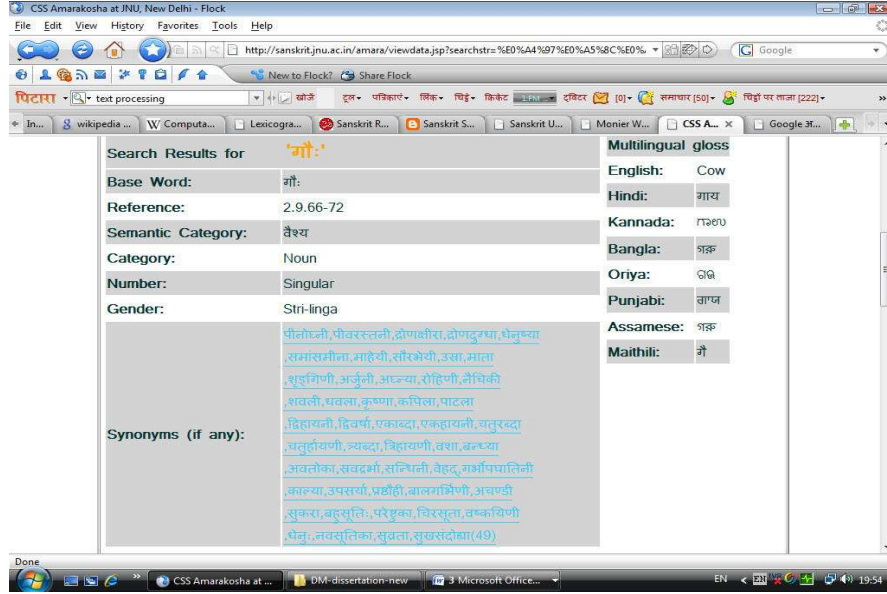
2.2.2.1. Online Multilingual Amarakośa²⁸

The system is based on *Amarakośa*, the Sanskrit thesaurus ascribed to *Amarasinha* (before 6th CE) has been developed by Dr. Girish Nath Jha and his team at SCSS/JNU. This ongoing RDBMS based system uses Java-JSP as front-end and SQL server as back-end. The multilingual (Sanskrit, Hindi, Kannada, Punjabi, Bangla, Oriya, Assamese, Maithili and English languages) data is stored in Unicode. The system facilitates storing upto 50 synonyms with category, gender, number information and detailed glosses, with cross-referencing among synonyms, search capability in the supported Indian languages and ontology display. The snapshots of this work are presented below:



(Fig-2.7: A user can search his query in three ways; either by selecting from a drop-down box or by clicking a letter from alphabetical list or by typing directly.)

²⁸ The live version can be accessed at- <http://sanskrit.jnu.ac.in/amara/index.jsp> (accessed on July 15, 08)



(Fig-2.8: the query for “गौः” gives all the references including category, synonyms, ontology and meanings in existing languages)

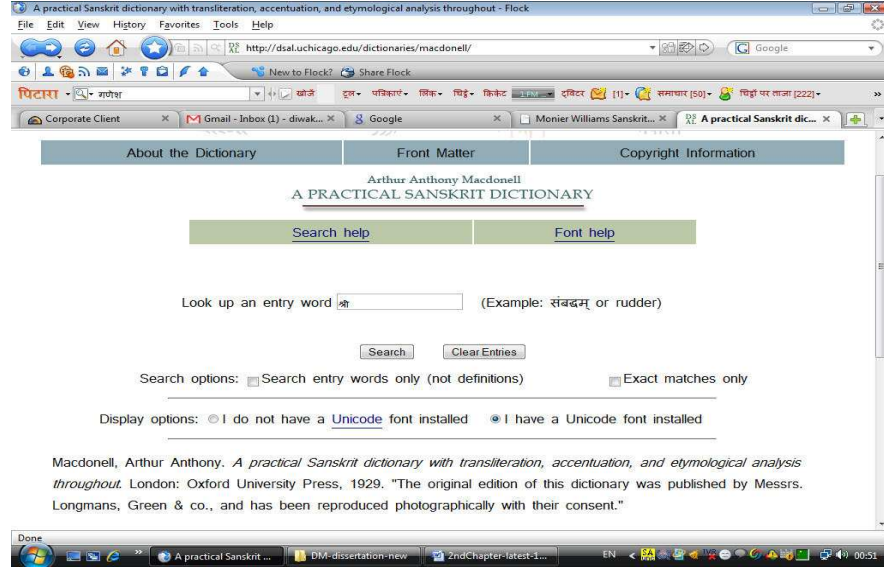


(Fig-2.9: clicking on synonyms on previous page gives all possible results which are shown above)

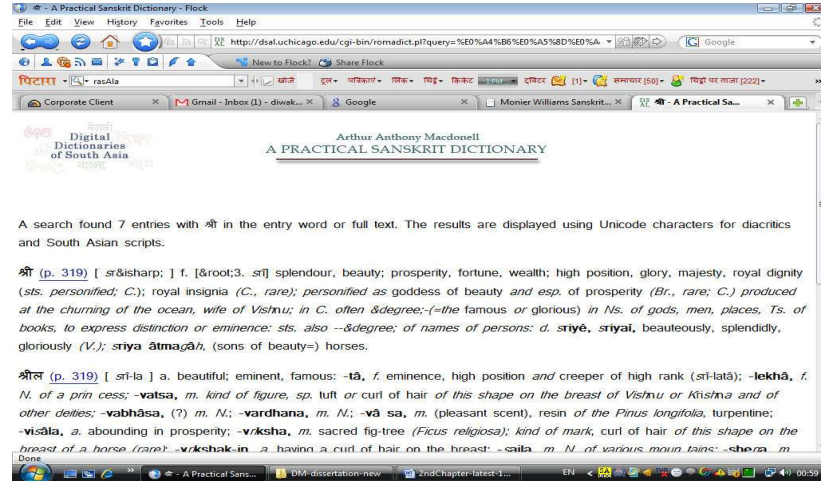
2.2.2.2. Online Macdonell Sanskrit-English Dictionary²⁹

This online dictionary is based on ‘A practical Sanskrit dictionary’ by Arthur Anthony Macdonell with transliteration, accentuation, and etymological analysis throughout. The system includes both Devanagari and Roman alphabets. The snapshots of this system are shown below:

²⁹ It is available at- <http://dsal.uchicago.edu/dictionaries/macdonell/> (accessed on July 17, 08)



(Fig-2.10: The user can put his query for both Sanskrit and English words. The input for Sanskrit may include both Unicode Devanāgarī and ITRANS.)



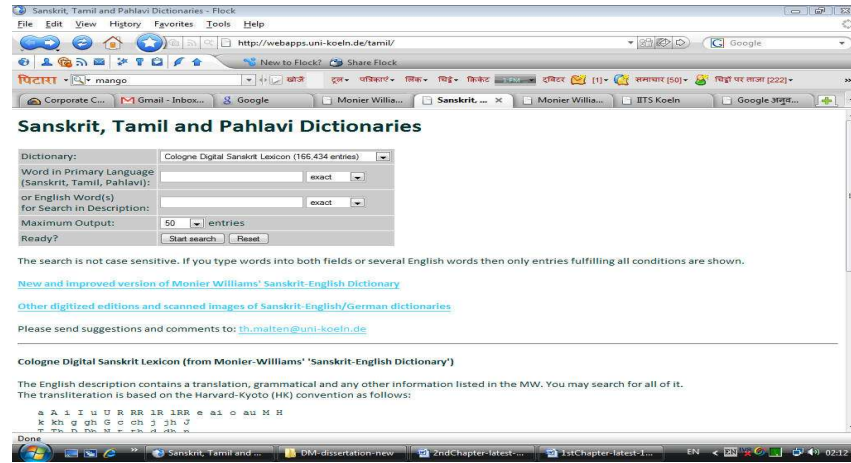
(Fig-2.11: For example, giving 'श्री' as an input, system gives all the occurrences)

2.2.2.3. Cologne Digital Sanskrit Lexicon & Capeller's Sanskrit-English Dictionary³⁰

The Cologne Digital Sanskrit Lexicon developed at the Institute of Indology and Tamil Studies, Cologne University, is based on Monier Williams' Sanskrit-English dictionary and contains 166,434 entries while the online Capeller's dictionary is based on Capeller's 1891 Sanskrit-English Dictionary which includes 37,413 entries. In both the digital dictionaries, the English description contains a translation, and grammatical and other information listed in their

³⁰ <http://webapps.uni-koeln.de/tamil/> (accessed on July 18, 08)

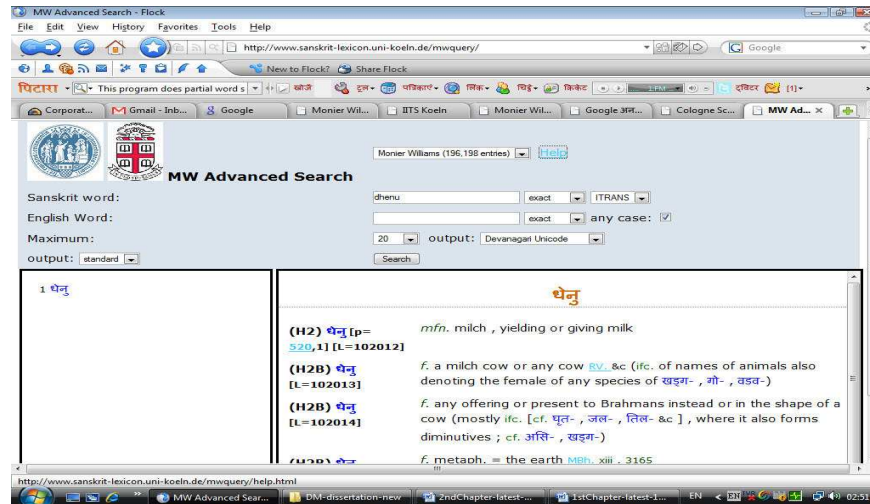
respective printed form. One may search it in entirety from both the languages. For the Sanskrit query, the transliteration is based on the Harvard-Kyoto (HK) convention. The snapshot is given below:



(Fig-2.12: Main page of the dictionaries which includes two other dictionaries of the Tamil and Pahlavi languages.)

2.2.2.4. Online Monier Williams (MW) Sanskrit-English Dictionary Advanced Search³¹

This MW dictionary with advanced search options is prepared by the Institute of Indology and Tamil Studies, Cologne University and contains a total of 196,198 entries. The program does partial word searches for Sanskrit head words or English words in the MW dictionary. It takes queries for Sanskrit words in ITRANS, HK and SLP1 and gives output in several schemes. A snapshot is given here:

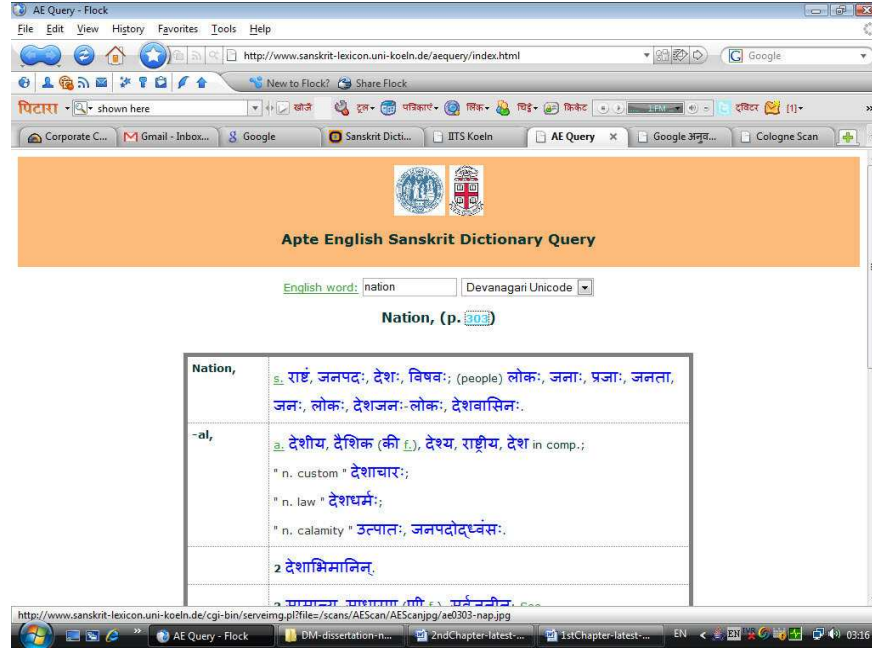


(Fig-2.13: Home page of the MW Sanskrit-English Dictionary Advanced Search)

³¹ <http://www.sanskrit-lexicon.uni-koeln.de/mwquery/> (accessed on July 18, 08)

2.2.2.5. Online Apte English-Sanskrit Dictionary³²

This program, also built by the Institute of Indology and Tamil Studies, Cologne University depends on Vaman Shivaram Apte's English-Sanskrit Dictionary. The output can be seen in Unicode Devanagari, Roman Unicode, HK, SLP1 and ITRANS. One can also see a scanned copy of the original dictionary by clicking the page numbers which comes in parenthesis with the English word. A snapshot of this Apte's dictionary is shown here:



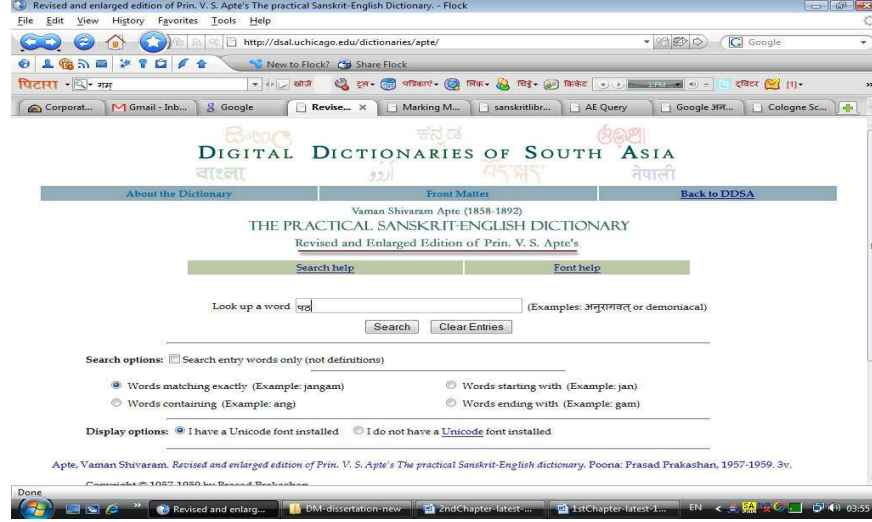
(Fig-2.14: Home page of Apte's Dictionary)

2.2.2.6. Online Apte Sanskrit-English Dictionary³³

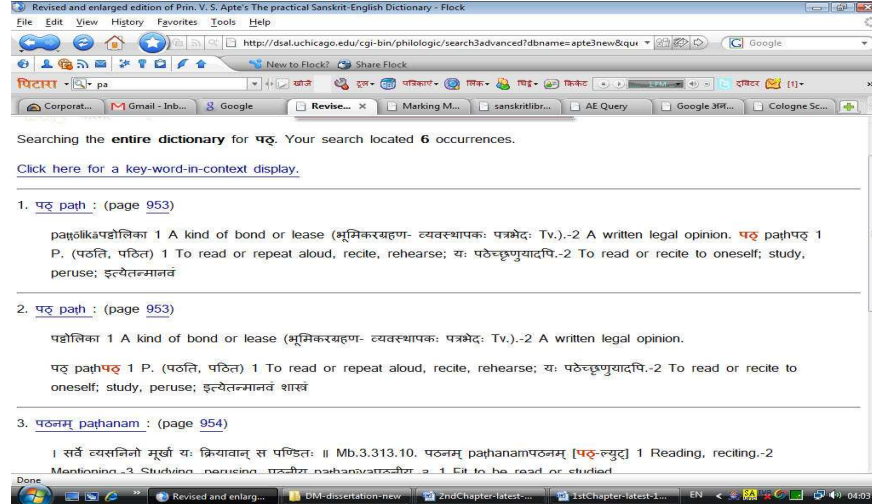
This online dictionary is based on the revised and enlarged edition of Vaman Shivaram Apte's (1858-1892) *The Practical Sanskrit-English Dictionary*. This program, which takes input in Devanagari and Roman alphabets, is developed by the 'South Asia Language and Area Center' at the University of Chicago, Columbia University, and the Triangle South Asia Consortium in North Carolina under a project entitled '*Digital Dictionary of South Asia*'. The snapshots of this dictionary are given here:

³² <http://www.sanskrit-lexicon.uni-koeln.de/aequery/index.html> (accessed on July 19, 08)

³³ <http://dsal.uchicago.edu/dictionaries/apte/> (accessed on July 19, 08)



(Fig-2.15: Home page of Apte's Sanskrit-English Dictionary)



(Fig-2.16: Giving 'पठ' as an input, the system gives all the occurrences)

2.2.2.7. Spoken Sanskrit Dictionary³⁴

This is a new bilingual online hypertext dictionary for Sanskrit-English and English-Sanskrit, designed to focus on spoken Sanskrit, which is alive as a common language. This Sanskrit dictionary is still under construction, and the correction process has not yet been finished. Obvious errors may be corrected by using the **Edit-** link to the right of each dictionary entry. The snapshot can be seen here:

³⁴ It can be accessed at- <http://spokensanskrit.de/> (accessed on July 20, 08)



(Fig-2.17: Home page of the Spoken Sanskrit Dictionary)

2.2.2.8. The Sanskrit Heritage Dictionary³⁵

This is an online searchable hypertext Sanskrit-French dictionary with normal search, search without diacritic marks, and search for inflected forms. The system developed by Gerard Huet requires Devanagari fonts and open type fonts for Roman transliteration with diacritics. A snapshot of this system can be seen below:

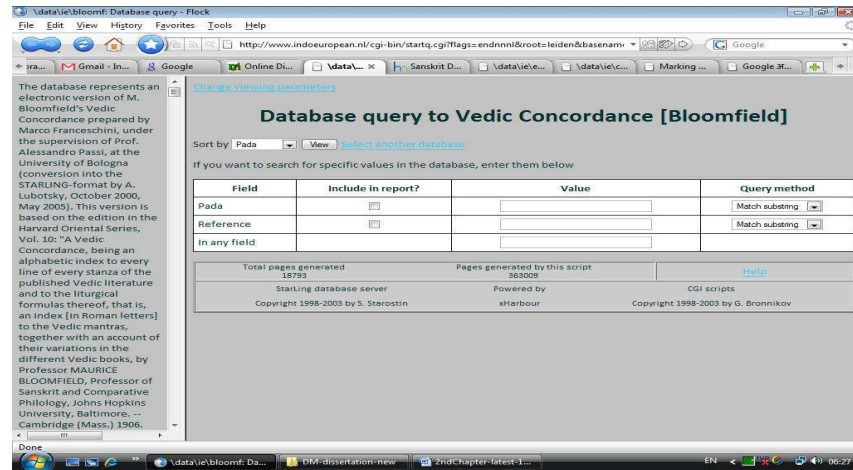


(Fig-2.18: Web page of Sanskrit-French Dictionary)

³⁵ It is available at- <http://sanskrit.inria.fr/DICO/index.html> (accessed on July 15, 08)

2.2.2.9. Database query to Bloomfield's Vedic Concordance

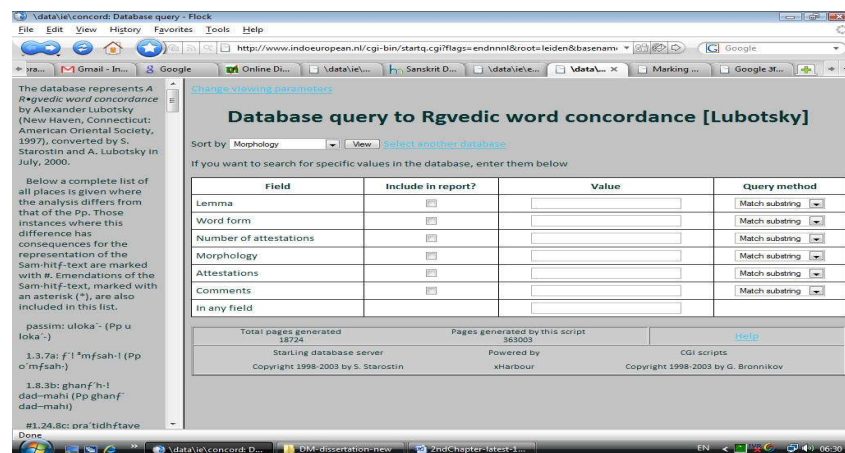
This is an electronic version of M. Bloomfield's *Vedic Concordance* prepared by Marco Franceschini, under the supervision of Prof. Alessandro Passi, at the University of Bologna (conversion into the STARLING-format by A. Lubotsky, October 2000, May 2005). A snapshot of this system can be seen below:



(Fig-2.19: Web page of the Bloomfield's Vedic Concordance)

2.2.2.10. Database query to Lubotsky's Rigvedic Word Concordance

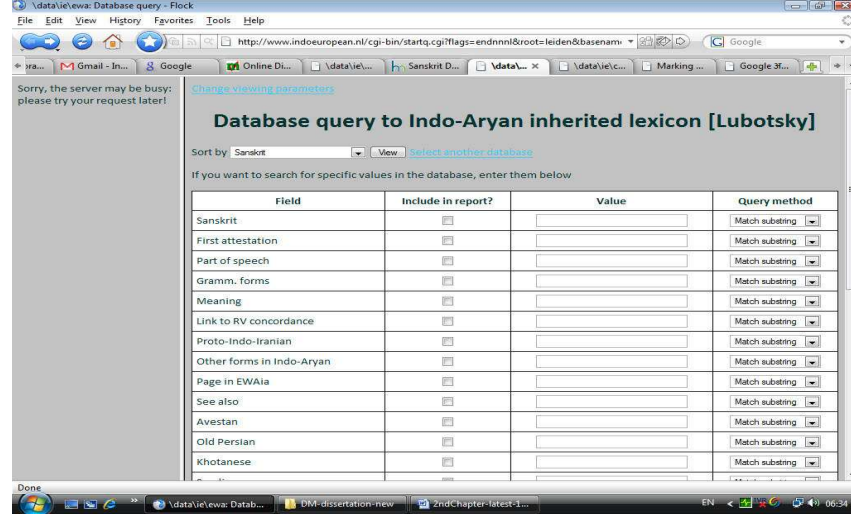
The database access to *A Rigvedic word concordance* by Alexander Lubotsky (New Haven, Connecticut: American Oriental Society, 1997) was converted by S. Starostin and A. Lubotsky in July, 2000. A snapshot of this system can be seen below:



(Fig-2.20: Web page of Lubotsky's Rigvedic Word Concordance)

2.2.2.11. Database query to Lubotsky's Indo-Aryan inherited lexicon

The database of Indo-Aryan inherited lexicon is based on M. Mayrhofer, *Etymologisches Wörterbuch des Altindoarischen* (Heidelberg 1986-1996). A snapshot of this system can be seen below:



(Fig-2.21: Web page of Lubotsky's Indo-Aryan inherited lexicon)

2.3. Methodology for e-indexing of Mahābhārata (Ādiparva)

2.3.1. Study and ascertain the most authentic text of the Mahābhārata and its structure

The *Mahābhārata* was selected for online indexing because it incorporates several branches of knowledge including philosophy, law, ethics, statecraft, warfare, history, and ethnology etc. So the text is not only very important but is also very large. Therefore it becomes impossible for someone to search a specific keyword in it. The indices thus prepared will constitute separate text in itself due to the size of *Mahābhārata* and will be of tremendous use to the researchers and users.

After studying the critical editions of the *Mahābhārata*, the primary and basic task was to determine the most appropriate and accepted structure and the text to be used for indexing purposes. The *Mahābhārata* has come down to us in two recensions, the northern and the southern, which are further subdivided into versions according to the scripts in which the text is written. The northern recensions are available in *Śāradā*, *Nepālī*, *Maithilī*, *Bānglā* and

Devanāgarī scripts. The southern recensions are in *Telugu*, *Grantha* and *Malayālama* scripts. Among the complete printed editions of *Mahābhārata*, the following three are considered to be of importance - the Kolkata, the Bombay and the Kumbhakonam editions. But none of these can be taken to represent any of the two recensions referred above.

At first, the search for the most authentic work on complete *Mahābhārata* was carried out and the Critical Edition of *Mahābhārata* of Bhandarkar Oriental Research Institute, Poona was selected as this work is recognized as the most authentic edition of *Mahābhārata* and has been compiled on the basis of comparative study of several different versions.

Traditionally, *Mahābhārata* is known as “*Śatasāhasrī Samhitā*” containing one lack śloka which is several times more than books like The Bible and Odyssey. Since it was not possible to do online indexing of such a vast work at M.Phil. level, it was decided to confine the study to the first Parva of *Mahābhārata* known as *Ādiparva* that contains 19 sub-parvans, 09 Ākhyanas, 225 Adhyāyas in 7197 śloka.

The next step was to get the electronic data which was found on the web in the form of a soft copy of *Mahābhārata* at the site of Prof. John Smith.³⁶

2.3.2. Sandhi Splitting

Since Sanskrit Sandhi has not yet been completely solved for computing purposes, the *saṁhitā* text of *Ādiparva* was manually Sandhi-split as it was necessary to store the text in both original as well as in split forms (*anvaya* form). As a lexical resource, this was to be the most important part. As Sanskrit is a nearly word-order free language, and also *Mbh.* is a huge collection of *laukika* Sanskrit words, this corpus can serve to be an e-corpus of Sanskrit which can be used as a lexical resource in any system for various purposes. The original text will also be important as an e-corpus of Sanskrit. Here are some examples of split forms of the original śloka in *Ādiparva* of *Mahābhārata*:

³⁶ <http://bombay.indology.info/index.html> (accessed first on May 10, 2006 and last accessed on July 21, 08)

| Samhitā | Split form |
|--|--|
| तमाश्रममनुप्राप्तं नैमिषारण्यवासिनः । चित्राः श्रोतुं कथास्तत्र परिववृस्तपस्विनः ॥1/1/1/0/3॥ | तमः आश्रममनुप्राप्तमनैमिष अरण्यवासिनः । चित्राः श्रोतुमकथाः तत्र परिववृः तपस्विनः ॥ |
| बाह्मीकं दमनं शैब्यं शर्यातिमजितं जितम ॥ विश्वामित्रममित्रघ्नमम्बरीषं महाबलम ॥1/1/1/0/167॥ | बाह्मीकमदमनमशैब्यमशर्यातिमअजितमजितम ॥ विश्वामित्रमअमित्रघ्नमअम्बरीषममहाबलम ॥ |
| युवां वर्णान्विकुरुथो विश्वरूपांस्तेऽधिक्षियन्ति भुवनानि विश्वा । ते भानवोऽप्यनुसृताश्चरन्ति देवा मनुष्याः क्षितिमाचरन्ति ॥1/3/3/0/68॥ | युवामवर्णानविकुरुथः विश्वरूपानस्ते अधिक्षियन्ति भुवनानि विश्वा । ते भानवः अपि अनुसृताः चरन्ति देवा मनुष्याः क्षितिमआचरन्ति ॥ |

[Table 2.6: Data of Ādiparva in Samhitā and Split form]

Working on Sandhi splitting of the text, sometimes the strict Pāṇinian technique was not applied for purposes of storage and search. Here are some examples where Sanskrit grammar rules were not exactly followed:

| Śloka Samhita | Śloka Pada |
|------------------|----------------------------------|
| नमस्कृत्य | नमः कृत्य ³⁷ |
| नैमिषारण्यवासिनः | नैमिष अरण्य वासिनः ³⁸ |
| सर्वतस्तमसावृते | सर्वतः तमस आवृते ³⁹ |
| तमुवाच | तम उवाच ⁴⁰ |

[Table 2.7: Example of unusual sandhi splitting]

2.3.3. Creation of a Relational Database System for Mahābhārata (Ādiparva)

³⁷ “नमस्कृत्य” is a compound of “नमः” and “कृत्य”. Here, “कृत्य” is separated as an independent word while *ktvā* changes into *lyap* only in the case of a prefix or compound.

³⁸ Here, “नैमिष” and “अरण्य” are used as independent word without suffix which is not acceptable in a sentence.

³⁹ Here, “तमः” can be regarded as a usable word (*pada*) while as an independent word it should be “तमसा” to retain its meaning.

⁴⁰ It is not the case of sandhi here. Neither “तम” nor “उवाच” is a substring of “तमुवाच” from the view point of Unicode characters. That is why they are separated.

Using the RDBMS techniques, the most prominent part and the core of this R&D was to design and develop a database driven knowledge base according to the structure of the *Mahābhārata*. Firstly, systematic tabular formats were created for the structure of *Mahābhārata*. The *Mahābhārata* has *parvans*, each *parvans* is further divided into sub-parvans, some sub-parvans also have *Ākhyānas*, all these sub-parvans are further divided into *adhyāyas* and these *adhyāyas* contain *ślokas*. After fixing the structure of *Mahābhārata*, the actual database design was done at the MS-SQL Server2005 in Unicode Devanāgarī. Some sample structures follow –

| Shloka_Id | Adhyaya_Id | Shloka_Nm | ShlokaSamhita | ShlokaPada |
|-----------|------------|-----------|---------------|------------|
| | | | | |
| | | | | |

[Table 2.8: Database structure of main table]

The above ‘**shloka**’ table contains *Ādiparva*’s *śloka* in the original form and also in the split form. The original form is for display and the split form is for the purpose of searching so that any word contained in the *śloka* in the modified form can be searched. The columns *Shloka_Id* and *Shloka_Nm* are different entities here. *Shloka_Nm* denotes the occurrence of a specific *śloka* in a particular *adhyāya* which is used for display while *Shloka_Id* is the unique identification number of each stanza to distinguish a *śloka* with a *Shloka_Nm* of one *adhyāya* from the *ślokas* with the same *Shloka_Nm* of other *adhyāyas*. Here, *Adhyāya_Id* which relates the *ślokas* to their respective *adhyāyas* of *Ādiparvan* is adjoined to ‘**adhyaya**’ table for further references.

As the search result, the *śloka* should be given with its detailed reference. If the components of reference, namely, *parva*, *adhyāya*, *ākhyāna* names are given with each *śloka* entry then it will be repeated and it will encumber the database unnecessarily. This can be avoided by giving it once and the detail could be recalled through relations. Also it is advantageous to give an ID number for each *parvan*, *ākhyāna* etc. for saving space and a fast processing of the search query. For the multi-layered structures, the repetition of super-categories could also be avoided by giving them in a separate table of sub-categories. Thus, only immediate super-categories of the *śloka* or other data are required to be stored in the table. More tables and their relations make the

database well structured and processing faster. A smaller number of columns in a table also ensure a faster search.

| Id | UP_Id | Akh_Id | Adh_Nm |
|----|-------|--------|--------|
| | | | |
| | | | |

[Table 2.9: Database structure of ‘adhyaya’ table]

The ‘**adhyaya**’ table contains four columns namely, *Id*, *UP_Id*, *Akh_Id* and *Adh_Nm*. It is a bridge table between the ‘**shloka**’ and ‘**akhyana/upaparva**’ tables because the column *Id* takes the information from the column *Adhyaya_Id* of ‘**shloka**’ table and passes the information through the adjoining the columns of *UP_Id* and *Akh_Id* to the table ‘**upaparva**’ and ‘**akhyana**’ respectively for the next level reference.

| Id | Akhyana |
|----|---------|
| | |
| | |

[Table 2.10: Database structure of ‘akhyana’ table]

The ‘**akhyana**’ table defines the *Ākhyānas* of *Ādiparvan* and contains two columns viz. *Id* and *Akhyana*. The column *Id* shows unique identification number of each *ākhyāna* of *Ādiparva* and the column *Akhyana* gives their name in Unicode Devanāgarī for display purposes.

| Id | Parva_Id | Upaparva |
|----|----------|----------|
| | | |
| | | |

[Table 2.11: Database structure of main table ‘upaparva’ table]

This ‘**upaparva**’ table contains three columns viz. *Id*, *Parva_Id* and *Upaparva* in which *Id* column is adjoined with the column *UP_Id* of ‘**adhyaya**’ table to take the previous information for connecting with further information. The *Parva_Id* column of this table goes to the *Id* column of next table i.e., ‘**parva**’ table for completing the reference of the user’s query.

| Id | Parva |
|----|-------|
| | |
| | |

[Table 2.12: Database structure of main table ‘parva’ table]

The table ‘**parva**’ has *Id* and *Parva* columns where the column *Id* is adjoined with *Parva_Id* of ‘**upaparva**’ table for giving the last information to complete the reference of user’s query.

2.3.4. Creating Java servlet engine for search

Creating a search engine for *Mahābhārata* was done using Java-JSP on Apache-Tomcat platform. Servlets are Java objects used to extend the capabilities of servers that host applications accessed via a request-response programming model. Although servlets can respond to any type of request, they are commonly used to extend the applications hosted by Web servers. For such applications, Java Servlet technology defines HTTP-specific servlet classes.⁴¹ Building web pages on the fly is useful (and commonly done) for a number of reasons. In this case, the indexer will dynamically create HTML output based on the user queries.

2.3.5. Creating an interactive page for search with display

The front end of the system (GUI) is live at <http://sanskrit.jnu.ac.in>. The main JSP page allows the user to give input in Devanāgarī UTF-8 format using HTML text box component. Upon clicking the button labeled “Search” (“खोजें”) it hands over the request to Java servlet engine for querying the database and matching specific key-words in the Database. After matching the key-

⁴¹ http://java.sun.com/j2ee/tutorial/1_3-fcs/doc/Servlets.html (accessed on July 14, 08)

word, the output is returned with its references and is displayed to the user in Devanāgarī UTF-8 format.

2.3.6. Primary Information

This system in its online mode will take a string (word, sentence or a partial string of positive length) of *Ādiparva* of *Mahābhārata* as an input and will give all possible results regarding the queries of a user as an output. For example, user wants to know about *Satyavati* (सत्यवती) the system will search the database. If the search term is found, all its occurrences will be returned in hyperlinked mode at the first page.

2.3.7. Secondary Information

At the second stage, the users on clicking any one hyperlinked word for further information will get a complete reference.

2.3.8. Tertiary Information

After getting the referential result the next step would be the cross-linking of words to different Sanskrit dictionaries (especially to those having Unicode compatibility), like- Online Multilingual Amarakosha site,⁴² Spoken Sanskrit dictionary site (by Klaus Glashoff, Germany),⁴³ Online Apte Sanskrit-English Dictionary,⁴⁴ Online Macdonell Sanskrit-English Dictionary,⁴⁵ Sanskrit Wikipedia⁴⁶ and others. Thus, anyone wanting to know further information on the searched word can get it by clicking the link. On clicking this word for specific further searches, the system will take the user to the corresponding sites for further details on the searched word. For example, if user wants to know the grammatical information about ‘अम्बिकया’, the system will first do *subanta* analysis for grammatical information (at <http://sanskrit.jnu.ac.in>) and then take the user to a specific external dictionary site for lexical information.

⁴² <http://sanskrit.jnu.ac.in/amara/index.jsp>

⁴³ <http://spokensanskrit.de/>

⁴⁴ <http://dsal.uchicago.edu/dictionaries/apte/>

⁴⁵ <http://dsal.uchicago.edu/dictionaries/macdonell/>

⁴⁶ <http://sa.wikipedia.org/wiki/Special:AllPages/>

Chapter-3

Structure of Mahābhārata and Ādiparva

3.1. The Text of Mahābhārata

Determining the text of *Mahābhārata* (*Mbh*) has been a problem for manuscriptologists, critics and historians. The methodology varies, depending upon the discipline and the purpose. The *Mbh*, as any other text of ancient Indian tradition, is complex to fix as far as its authoritative version is concerned. In the words of Sukthankar (*Ādiparvan*, Fascicule I, 1927) – in the *Mahābhārata* manuscript tradition, perhaps as much as in any literary tradition, the textual critic is faced with a perplexing profusion of versions as well as an amazing mixture of versions. Diverse elements have been working from the earliest times in favor of the development of different types, on the one hand; and on the other hand, there were other elements that operated against the evolution of sharply differentiated types. To understand the phenomenon of this luxuriant growth and indiscriminate fusion of versions, one must appreciate clearly certain historical details, certain special factors in the transmission of the *Mbh*. Notwithstanding the fact that we know so little that is certain and definite about the early history of the text, it can be confidently assumed that after its conception, the great epic was for centuries handed down (in differing forms and sizes) from poet to poet merely by word of mouth. It is moreover extremely probable that even after the text had been written down, large portions of it, especially such portions as were popular, continued to be committed to memory, by itinerant raconteurs (story tellers or narrators) for purposes of recitation. It is further easy to believe that no great care was lavished on the text by these custodians of the tradition to guard it against partial corruption and elaboration or against arbitrary emendation and normalization: to reproduce the received text with any great precision would be neither attempted by these bards nor required of them. It was inevitable that the versatile oral tradition should in one form or another react on the written tradition and *vice versa*. One important and necessary consequence of such antecedents as these is the impossibility of retracing all extant versions to any fixed and authentic archetype; since some of the modern editions could not but be descendants of fluctuating oral versions reduced to writing in some distant past, independently of each other, at different epochs and in different circumstances. In other words, even in its early phases the *Mbh* text tradition must have been not uniform and singular, but multiple and polygenous. To complicate matters further there appears to have followed a period in which there was a free comparison of manuscripts and extensive mutual borrowings, operations which in the course of indiscriminate crossing and re-crossing have completely confused the *differentiae* and produced a perfect wilderness of hybrid types.

These are, at least in part, still mere surmises. But the assumption of some such complicated derangements, beyond the normal vicissitudes of transmission, is necessary, to account for the strange vagaries of the *Mbh* manuscript tradition : to explain why in the best manuscripts one comes across at every step readings that are manifestly inferior and additional lines that are incontestably spurious : to elucidate how textual alterations, especially interpolations, starting from the most inconspicuous source of editorial revision, could extend over large areas with comparative ease and rapidity.

This state of things, if true, would make it impossible to apply to the *Mbh* the special canons of textual criticism which are derived from a study of classical (Greek and Latin) texts and which depend ultimately upon there being a more or less complete concatenation of copies and exemplars reaching finally back to a single authentic (written) archetype. The conflation of codices may, moreover, have been carried to such an extreme that we may even have to renounce all pretensions to disentangle completely, by means of purely objective criteria, their intricate mutual relationships. Prof. Sukthankar sounds a pessimistic note when he says - It would, therefore, be well not to ignore entirely the possibility that a wholly satisfactory restoration of the text to its pristine form (*śatasāhasrī saṁhitā* form) may be a task now beyond the powers of criticism.

It is by no means easy to answer the question how many manuscripts of the *Mbh* there are in existence. There are two factors responsible for this problem; firstly, because, no complete list of these manuscripts has been compiled; and, secondly, because the expression “*Mbh* manuscript”, as generally used, is ambiguous in the extreme; it may apply to a small manuscript of the *Bhagavadgītā* alone, as well as to a complete manuscript of the *Mbh*, in several volumes, containing all the eighteen parvans. Moreover, the parvans are mostly handed down separately, or in groups of few parvans at a time, at least in the oldest manuscripts now preserved. Therefore, in taking stock of *Mbh* manuscripts, it is best to take as unit of measurement a manuscript of a *single* parvan.

V.S. Sukthankar in his Prolegomena (Volume I, *Ādiparva*) states a very approximate computation, about 235 manuscripts of the *Ādiparva*, where, he counted only those who come within knowledge from catalogues of private and public libraries accessible to him. Further he states about a large number of manuscripts in private hands, of which nobody knows next to

nothing. The below structured tabular form shows the categorization of the 235 manuscripts of the Ādiparva:¹

| Parva | Version | No. of total Mss./ Mss. used by BORI for compilation of first two adhyāyas of Ādiparva of the Critical Edition |
|---------------------------|-------------|--|
| Ādiparva (Total 235 Mss.) | Devanagari | 107/30 [*] |
| | Bengali | 32/04 |
| | Nepali | 05/00 |
| | Sharada | 03/00 |
| | Maithili | 01/01 |
| | Nandinagari | 01/00 |
| | Grantha | 31/07 |
| | Telugu | 28/02 |
| | Malyalam | 26/04 |
| | Kannada | 01/00 |

[Table 3.1: Total number of manuscripts used for compilation of first parvan of critical edition of *Mbh.*]

Franklin Edgerton has collated 55 manuscripts for critical editing of the *Sabhāparva*. Among them he utilized 22 manuscripts of Northern recension and 03 from Southern recension. From the descriptive catalogues, published reports, and the lists of manuscripts available in public library in India and abroad, as also from a special search instituted in private library, the existence of more than 125 manuscripts of the *Bhīṣmaparva* (text only) have been so far established, about half of them being written in Devanāgarī, about 20 in Bengali, some 15 each in Grantha and Telugu, over 10 in Malyalam and only one in Śāradā characters. There exist, besides these 125 manuscripts of the text alone, nearly 40 others manuscripts giving the text of the parvan together with a Sanskrit commentary: namely Arjuna Miśra's *Bhāratārthadīpikā*, some 10 Mss.; Nīlakaṇṭha's *Bhāratabhāvadīpa*, about 30 Mss.; and two manuscripts of an anonymous commentator. Manuscripts containing commentaries alone on the *Bhīṣma* are more

¹ Critical Studies in the Mahābhārata (Volume I), pg. 15.

^{*} Among these 30 manuscripts, 07 come from Devanagari transcript of Kashmiri version.

than 30, there being, besides the 2 commentaries mentioned above, about 10 others, including *Devabodha*. This makes a total of about 200 MSS. of the *Bhīṣmaparva*. Among these manuscripts, only 60 Mss. have been actually procured and collated to S.K. Belvalkar, editor of critical edition of *Mbh* (*Bhīṣmaparva*). And finally he selected 34 Mss. to constitute the critical apparatus for *Bhīṣmaparva*.

The list of some manuscripts which is found in several volumes of BORI's critical edition of *Mbh* and from some other sources like Descriptive catalogue of Sanskrit Manuscripts (Vol. III) of Ranbir Sanskrit Research Institute, Jammu², is given below:

| S.No. | Recension | Version | Place* | Accession No. | Remarks, age, Script, Material, Condition |
|-------|--------------------------------|----------------------------------|--|------------------|--|
| 1. | Northern (North-Western group) | Kaśmīrī in Śāradā characters | India Office Lib, London. | 3226 (2137) | 1783AD, 169 folios, Contains first parvan, Indian paper. |
| 2. | " | " | Bombay Govt. Collection, Poona (deposited at BORI) | 159 of 1875-76AD | 16 th or 17 th century, 214 folios, Bhūrjapatra, Contains first three parvans, Incomplete. |
| 3. | " | Kaśmīrī in Devanāgarī characters | " | 229 of 1895-1902 | Samvat 1795, 181 folios, Contains first three parvans, Old Indian paper. |
| 4. | " | " | " | 182 of 1891-96 | VS 1694, 296 folios, Indian paper. |
| 5. | " | " | " | 565 of 1882-83 | Śaka 1616, 237 folios, Complete Ms of Mbh (missing Śāntiparva) Old Indian paper. Complete. |
| 6. | " | " | " | 209 of 1887-91 | 386 folios, Partly Old Indian and partly modern European paper. |
| 7. | " | " | Oriental Institute Lib., Baroda | 632 | VS 1575, 407 folios, Old Indian paper. |
| 8. | " | " | D.A.V.College Lib., Lahore. | 01 | About 400 years old, 28 folios, Contains up to 1.3.152, Incomplete. |
| 9. | Northern (Central group) | Nepālī | One Private Lib, Nepal | ----- | NS 632(1511AD), Contains Ā di (Sauptika-Aiṣika & Viśoka-Strī), On Palm-leaf. |

² There are 54 Mss. of *Mbh* listed in the catalogue of the library.

* The eighth number and some other manuscripts may be unavailable or destroyed after partitioning of India in 1947.

| | | | | | |
|-----|---|------------|--|---------------------|--|
| 10. | ” | Maithili | Nepal Darbar Lib., Nepal. | 1364 | LS 411 (1528AD), On Palm-leaf. |
| 11. | ” | Bengali | Vishvabharati Lib., Shantiniketan. | 01 | Śaka 1662, 207 folios, On Palm-leaf. |
| 12. | ” | ” | ” | 258 | 82 folios, On Palm-leaf. |
| 13. | ” | ” | ” | 782 | Śaka 1681, 199 folios. On paper. |
| 14. | ” | ” | ” | 413 | 164 folios, On Paper. |
| 15. | ” | ” | University Lib., Dacca. | 485 | Śaka 1708, 366 folios, On old Indian yellow paper. |
| 16. | ” | ” | ” | 735 | 19 th century, 346 folios, On old Indian yellow paper. |
| 17. | ” | Devanāgarī | Bombay Govt. Collection, Poona (deposited at BORI) | 30 of A 1879-80 | 416 folios, With Commentary of Arjunamiśra, Contains Ādi, Old Indian glossy paper. |
| 18. | ” | ” | ” | 468 of Viśrāmbāg I | VS 1676, 415 folios, With Commentary of Arjunamiśra, Indian paper. |
| 19. | ” | ” | ” | 234 of 1895-1902 | 683 folios, With Commentary of Nīlakaṇṭha, Thick Indian paper. |
| 20. | ” | ” | ” | 29 of A 1879-80 | Śaka 1677, 230 folios, Contains all parvans, cream-colored paper. |
| 21. | ” | ” | ” | 191 of Viśrāmbāg II | 221 folios, Contains Ādi (from 1.1.205 to end of adhyāya 2), old Indian glossy paper. |
| 22. | ” | ” | ” | 266 of Viśrāmbāg II | 1-121 folios (fol. 122-189 are found under No. 86 MS.), Contains up to end of adhy. 2 of Ādi, old Indian unglazed paper. |
| 23. | ” | ” | Saraswathi Mahal Lib., Tanjore. | 1246 | 448 folios, Paper. |
| 24. | ” | ” | ” | 1199 | 306 folios, Paper. |
| 25. | ” | ” | ” | 1313 | Śaka 1623, 366 folios, With commentary of Ratnagarbha, Paper. |
| 26. | ” | ” | ” | 1339 | 108 folios, With commentary of Ratnagarbha, Paper. |
| 27. | ” | ” | ” | 1152 | VS 1664, 340 folios, Paper. |
| 28. | ” | ” | ” | 1360 | 120 folios, Paper, Incomplete. |
| 29. | ” | ” | ” | 1126 | 235 folios, Paper. |

| | | | | | |
|-----|----------|-----------|--|-------|---|
| 30. | ” | ” | ” | 1269 | 262 folios, Paper. |
| 31. | ” | ” | ” | 1329 | 196 folios, Paper. |
| 32. | ” | ” | ” | 1176 | 279 folios, Paper. |
| 33. | ” | ” | ” | 1293 | 352 folios, Paper. |
| 34. | ” | ” | ” | 1340 | 290 folios, Paper. |
| 35. | ” | ” | ” | 1373 | 21 folios, Paper. |
| 36. | ” | ” | D.A.V.College Lib., Lahore. | 04 | VS 1858, 246 folios, Paper. |
| 37. | Southern | Telugu | Yadugiri Yatiraj Math, Melkote. | ---- | 267 folios, Contains Ādi and Sabhāparva, Palm-leaf. |
| 38. | ” | ” | Saraswathi Mahal Lib., Tanjore. | 11865 | 400 folios, Contains Ādi (Incomplete), Palm-leaf. |
| 39. | ” | ” | ” | 11809 | 164 folios, Contains first five parvans, Palm-leaf. |
| 40. | ” | Grantha | Yadugiri Yatiraj Math, Melkote. | ---- | 110 folios, Palm-leaf. |
| 41. | ” | ” | ” | ---- | 202 folios, Contains first four parvans, Palm-leaf. |
| 42. | ” | ” | ” | ---- | 217 folios, Palm-leaf. |
| 43. | ” | ” | Saraswathi Mahal Lib., Tanjore. | 11823 | 316 folios, Palm-leaf. |
| 44. | ” | ” | ” | 11838 | 477 folios, Palm-leaf. |
| 45. | ” | ” | ” | 11851 | 320 folios, Contains Ādi & Sabhā (up to adhyāya 1-38), Palm-leaf. |
| 46. | ” | ” | ” | 11860 | 324 folios, Palm-leaf. |
| 47. | ” | Malayalam | Private Lib. of Chief of Idapalli, Cochin. | ---- | 79 folios, Palm-leaf, Incomplete (ending with adhy. 53). |
| 48. | ” | ” | State Lib., Cochin. | 05 | 122 folios, Palm-leaf, Incomplete (ending with adhy. 53). |
| 49. | ” | ” | ” | 01 | Kollam 1013 (1838AD), 166 folios, Palm-leaf. |
| 50. | ” | ” | Private Lib. of Kallenkara Pisharam, Cochin. | ---- | 57 folios, Palm-leaf, Incomplete (ending with adhy. 53). |
| 51. | ” | ” | Paliyam MSS Lib., Cochin. | ---- | 245 folios, Palm-leaf. |
| 52. | ” | ” | Private Lib. of Nareri Mana, Malabar. | ---- | 163 folios, Palm-leaf, Incomplete. |
| 53. | ” | ” | Private Lib. of Avaṇapparambu | ---- | 170 folios, Palm-leaf. |

| | | | | | |
|-----|--------------------------------|----------------------------------|--|---------------------|---|
| | | | Mana, Cochin. | | |
| 54. | ” | ” | Poomulli Mana Lib., Malabar. | 297 | Kollam 1017 (1842AD), 183 folios, Palm-leaf. |
| 55. | Northern (North-Western group) | Kaśmīrī in Devanāgarī characters | Bombay Govt. Collection, Poona (deposited at BORI) | 56 of 1882-83 | 99 folios, Contains Sabhā alone, paper, complete. |
| 56. | ” | ” | ” | 29 of 1879-80 | Śaka 1675-80 (1753-58AD), Contains full Mbh (Sabhā on 78 folios), some parvans have Nīlakaṇṭha’s commentary, paper, Complete. |
| 57. | ” | ” | ” | 192 of Viśrāmbāg II | 74 folios, Contains Sabhā alone, paper, complete. |
| 58. | ” | ” | Adyar Library. | XXXV B 131 | VS 1880, 73 folios, Contains Sabhā alone, paper, complete. |
| 59. | Northern (Central group) | Nepālī in Nevārī characters) | Nepal Darbar Lib., Nepal. | I-1635 | About 560 years old, 93 folios (last leaf missing), palf-leaves. |
| 60. | ” | Maithili | ” | I-1559 | LS 335 (1452AD), 192 folios, palm-leaves, complete. |
| 61. | ” | Bengali | University Lib., Dacca. | 971 | Śaka 1680 (1758AD), 145 folios. |
| 62. | ” | ” | ” | 1093 | Śaka 1770 (1848AD), 165 folios (one 85 th missing), palm-leaves. |
| 63. | ” | ” | ” | 592 | Śaka 1662 (1740 AD), 128 folios |
| 64. | ” | ” | Sahitya Parishad Lib., Calcutta. | 76 | Śaka 1690 (1678AD), 102 folios, paper. |
| 65. | ” | ” | Vishvabharati Lib., Shantiniketan | 65 | Śaka 1691 (1769AD), 151 folios, palm-leaf, worm-eaten. |
| 66. | ” | ” | ” | 273 | Undated, 76 folios, palm-leaves. |
| 67. | ” | Devanāgarī | Bombay Govt. Collection, Poona (deposited at BORI) | 167 of 1884-87 | 115 folios, contains Sabhā alone with commentary of Devabodha, paper, complete. |
| 68. | ” | ” | ” | 74 of 1902-07 | VS 1821, 115 folios, contains Sabhā alone with commentary of Devabodha, paper, complete. |
| 69. | ” | ” | ” | 210 of 1887-91 | A composite Mss, in two parts; written on different |

| | | | | | |
|-----|--------------------------------|----------------------------------|--|--------------------|--|
| | | | | | paper by different scribes. Part-I has 52 and II has 34 folios, Contains Sabhā alone, paper, complete. |
| 70. | ” | ” | ” | 469 of Viśrāmbāg I | Śaka 1747 (1825AD), 84 folios, Contains Sabhā alone, complete. |
| 71. | ” | ” | ” | 510 of Viśrāmbāg I | VS 16 (?), 93 folios, Contains Sabhā alone, paper (it is yellow and seems quite old), complete. |
| 72. | ” | ” | ” | 511 of Viśrāmbāg I | 172 folios, Contains Sabhā alone with commentary of Nīlakaṇṭha, paper, complete. |
| 73. | ” | ” | Oriental Institute Lib., Baroda. | 827 | VS 1628, 114 folios (1 st is missing), paper. |
| 74. | Southern | Grantha | State Library, Pudukkottai. | 319 | Date 6 th day of Jyēṣṭha of Dundubhi, 254 folios, contains Ādi, Sabhā, Sautika, Mausala, Mahāprasthānika and Svargārohaṇa parva, palm-leaf. |
| 75. | ” | ” | Saraswathi Mahal Lib., Tanjore. | 11824 | 148 folios, palm-leaf, seems very old but undated. |
| 76. | ” | ” | ” | 11839 | Seems very old but undated, contains Sabhā (from adhy. 39 to last) and Āraṇyakaparvan, palm-leaf. |
| 77. | ” | Malyalam | Poomulli Mana Lib., Malabar. | 298 | 194 folios, palm-leaf. |
| 78. | Northern (North-Western group) | Kaśmīrī in Śāradā characters | Royal Asiatic Society of Bangal, Calcutta. | 3312 (4065) | Laukika Saṁvat 85 (1739AD), 112 folios, Contains Bhīṣma, Āraṇyaka, Virāṭa and Karṇaparvans, Old Kashmirian paper. |
| 79. | ” | Kaśmīrī in Devanāgarī characters | The Bhārata Itihāsa Saṁśodhaka Maṇḍala, Poona. | 207 | Śaka 1690 (1768AD), 227 folios, Contains full Mbh (except Virāṭa, Anuśāsana and Aśvamsdhika), old country paper. |
| 80. | ” | ” | India Office Lib, London. | 3226 | VS 1839 (1783AD), 122 folios, Contains Ādi, Virāṭa, Śānti (Dānadharma portion only), Bhīṣma, Complete. |
| 81. | ” | ” | BORI's Mbh collection, Poona. | 246 | Śaka 1693 (1771AD), 146 folios, contains full Mbh |

| | | | | | |
|-----|--------------------------------|-------------------------------------|---|--------------------------|--|
| | | | | | (missing Aśvamedhika), tough country paper, Complete. |
| 82. | ” | ” | Kamat collection of BORI, Poona. | 23 | Samvat 1742 (1686AD), 288 folios. |
| 83. | ” | Kaśmīrī in Bengali transcript | University Lib., Dacca. | 669 | Śaka 1675 (1753AD), 157 folios, palm-leaf, complete. |
| 84. | Northern (Central group) | Bengali | Vishvabharati Lib., Shantiniketan | 780 | Undated, 248 folios, complete. |
| 85. | ” | ” | ” | 467 | Śaka 1683 (1761AD), I-223 folios, Contains Bhīṣmaparvan, Complete. |
| 86. | ” | ” | ” | 07 | Undated (seems old), 270 folios, worm-eaten, complete. |
| 87. | ” | ” | University Lib., Dacca. | 2789 | Śaka 1671 (1749AD), I-240 folios, Contains Bhīṣmaparvan, paper, Complete. |
| 88. | ” | Devanāgarī | Bombay Govt. Collection, Poona (deposited at BORI) | 482 of Viśrāmbāg I | Undated but seems 300 years old, 285 folios, with commentary of Arjuna Miśra. Country paper. |
| 89. | ” | ” | ” | 483 of Viśrāmbāg I | Undated, 336 folios, with commentary of Nīlakaṇṭha, Country paper. |
| 90. | ” | ” | ” | 481 of Viśrāmbāg I | VS 1729 (1672AD), 187 folios out of which 35 are missing, Indian paper. |
| 91. | ” | ” | ” | 61 of 1882-83 | Undated, 246 folios, rough country paper. |
| 92. | ” | ” | ” | 213 of 1887-91 | Śaka 1606 (1684AD), 240 folios, old country paper. |
| 93. | ” | ” | Oriental Institute Lib., Baroda. | 10801 | VS 1753 (1697AD), contains Bhīṣma and Āśramavāsina, among total 33 folios comes for Bhagavadgītā. |
| 94. | Southern | Telugu | Saraswathi Mahal Lib., Tanjore. | 11810 | Contains 6-11 parvan, palm- leaf. |
| 95. | ” | ” | D.A.V.College Lib., Lahore. | 3945 | Undated, Contains Bhīṣma, Droṇa and Karṇa, Palm-leaf. |
| 96. | ” | Grantha | ” | 3764 | Contains first seven parvan, palm-leaf. |
| 97. | ” | ” | BORI's Mbh collection, Poona. | 267 B | 260 folios, Contains Udyoga and Bhīṣma. |

| | | | | | |
|------|----------|------------|---|-------------------|--|
| 98. | ” | ” | State Library, Pudukottai. | 260 | 270 folios, contains Udyoga, Bhīṣma & Droṇa, palm-leaf, completed. |
| 99. | ” | ” | Saraswathi Mahal Lib., Tanjore. | 11842 | 282 folios, Contains Bhīṣma only, palm-leaf. |
| 100. | ” | Malyalam | Chennamaigalam, Cochin. | 928 | Kollam 1028 (1853AD), 299 folios, Contains Bhīṣma & Udyoga. |
| 101. | ” | ” | Ponnokottu Mana Lib., Travancore, Alwaye. | ---- | 200 folios, palm-leaf, complete. |
| 102. | ” | ” | Nareri Mana Lib, Malabar. | 504 | 172 folios, Contains Bhīṣma only, palm-leaf, complete. |
| 103. | ” | ” | Bombay Govt. Collection, Poona (deposited at BORI) | 62 of 1924-26. | 254 folios, contains bhīṣma only, palm-leaf. |
| 104. | Northern | Devanāgarī | Ranbir Sanskrit Research Institute, Jammu. | 3712 | Undated but not very old, 05 leaves, contains Ādiparva with commentary of Nīlakaṇṭha. |
| 105. | ” | ” | ” | 3837 | Undated but fairly old, 381 leaves, contains Ādiparva with commentary of Nīlakaṇṭha. |
| 106. | ” | ” | ” | 3732 | Undated, 499 leaves, contains Āśramavāsikaparva with commentary of Nīlakaṇṭha. |
| 107. | ” | ” | ” | 3844 | Undated, 44 leaves, contains Aśvamedhikaparva with commentary of Nīlakaṇṭha. |
| 108. | ” | ” | ” | 3716 | 243 leaves, contains Udyogaparva with commentary of Nīlakaṇṭha. |
| 109. | ” | ” | ” | 3947 | 88 leaves, contains Karnaṇaparva with commentary of Nīlakaṇṭha. |
| 110. | ” | ” | ” | 3950 | 8 leaves, contains Droṇaparva. |
| 111. | ” | ” | ” | 3845 | 158 leaves, contains Bhīṣmaparva. |
| 112. | ” | ” | ” | 3838 | 133 leaves, contains Mahāprasthānika with commentary of Nīlakaṇṭha. |
| 113. | ” | ” | ” | 3730 | 12 leaves, contains |

| | | | | | |
|------|---|---|---|--------|---|
| | | | | | Mausalaparva. |
| 114. | ” | ” | ” | 3714 | 524 leaves, contains Vanaparva. |
| 115. | ” | ” | ” | 3715 | 114 leaves, contains Virāṭaparva. |
| 116. | ” | ” | ” | 3722 | 26 leaves, contains Sautikaparva. |
| 117. | ” | ” | ” | 3723 | 23 leaves, contains Strīparvan. |
| 118. | ” | ” | ” | 3979 | 400 leaves, contains Dānadharma of Śāntiparvan. |
| 119. | ” | ” | ” | 518 ka | 212 leaves, contains Rājadharmā of Śāntiparvan. |
| 120. | ” | ” | ” | 3725 | 64 leaves contains Āpaddharma of Śāntiparvan with commentary of Nīlakaṇṭha. |
| 121. | ” | ” | ” | 3731 | 9 leaves, contains Svargārohaṇa. |
| 121. | ” | ” | ” | 3961 | 135 leaves, contains Virāṭaparvan, slightly moth-eaten. |

[Table 3.2: The list of *Mbh.* manuscripts found in the volumes of critical edition of BORI and some other libraries]

3.2. Mahābhārata according to BORI

3.2.1. BORI and the Mahābhārata

The BORI was founded in 1915. The first verse of the *Mbh* was written by Sir R. G. Bhandarkar on 1st April 1919 while inaugurating the work of the Institute on the critical edition of *Mbh*. On 4th August 1925 V. S. Sukthankar took charge of this work as its General Editor and reorganized it on a sound basis. For the subsequent 17 years he worked on it bringing honor to himself and to the Institute. On 5th January 1943, after completion of first volume of critical edition of *Mbh*, Sukthankar gave his statement before the delegates for the Silver Jubilee celebrations in the Tata Hall of the Institute that “The part of the Epic critically dealt with so far is, I imagine, in bulk about four times as great as the Greek Epics, *Iliad* and *Odyssey* put together and one and half times as our *Rāmāyaṇa*. Amid the deepest strands that are woven in the thread of our civilization there is more than one that is drawn originally from *Bhāratavarṣa* and from Sanskrit literature and well in the centre of this vast mass of literature, there stands this deathless traditional book of divine inspiration, unapproachable and far removed from possibilities of human constitution.”

Side by side with his text-critical work on the *Mbh*, Sukthankar kept on studying its content and meaning and for that purpose studied many books on Philosophy and Religion. The valuable and scholarly contents of the volumes of BORI's *Mbh* speak for themselves and will continue to speak with greater resonance as years pass by.

The work to prepare and publish a critical edition of *Mbh* which started in 1919 by BORI, was completed in 1966 as an enormous literary project containing 19 volumes on 15,000+ demi quarto pages, and this historic event was formally announced by the then President Dr. S. Radhakrishnan at a special function held at *Mahābhārata* Institute on September 22, 1966. Subsequently, the Institute also prepared and published a critical edition of the *Harivaṃśa* (2 volumes containing 1,711 pages).³

3.2.2. Structure of Mahābhārata

According to the nineteen voluminous parts of BORI's critical edition, the *Mbh*, the greatest epic of the world, contains 18 parvans i.e., *Ādi*, *Sabhā*, *Āraṇyaka*, *Virāṭa*, *Udyoga*, *Bhīṣma*, *Droṇa*, *Karna*, *Śalya*, *Sauptika*, *Strī*, *Śānti*, *Anuśāsana*, *Aśvamedhika*, *Āśramavāsika*, *Mausala*, *Mahāprasthānika* and *Svargārohaṇaparva*. It contains an appendix (*khilaparva*) called *Harivaṃśa* which has 16,375 verses.⁴ All these 18 parvans and one additional *Harivaṃśa* contains 100 sub-parvans. Excluding the *Khilaparva*, there are 1995 adhyāyas and 73,817 (This number is determined after calculating ślokas of the all parvans of *Mbh*. The basis for this calculation is the downloaded soft copy of 18 parvans which is taken from the webpage of Prof. John Smith.⁵ However, the BORI's⁶ webpage declares the number of ślokas are 89000+) stanzas (ślokas) although it is called '*Śatasāhasrī Samhitā*' traditionally. There are many *Ākhyānas* exist in the *Mbh*, although the accurate number of *Ākhyānas* could not be determined in spite of an extensive effort.

³ <http://www.virtualpune.com/html/channel/edu/institutes/html/bhandark.shtml> (accessed on July 05, 08)

⁴ http://www.experiencefestival.com/a/Mahabharata_-_The_Mahabharata_the_epic_story/id/1740155 (accessed on July 05, 08)

⁵ <http://bombay.indology.info/index.html> (accessed on July 04, 08)

⁶ <http://www.bori.ac.in/mahabharata.htm> (accessed on July 05, 08)

3.2.2.1. ĀDIPARVA

It is the introductory parva of *Mbh* which describes the genealogy of the two families, the birth and nurture of *Dhṛtarāṣṭra* and *Pāṇḍu*, their marriages, the births of the hundred sons of the former and the five of the latter, the enmity and rivalry between the young princes of the two branches, and the winning of *Draupadī* at the swayamvara. According to BORI's critical edition of *Mbh*, Ādiparva has 19 sub-parvans, 09 ākhyānas, 225 adhyāyas and 7197 ślokas.

3.2.2.2. SABHĀPARVA

The second *Sabhāparvan* relates some of the most seminal events of the epic, culminating in the famous game of dice between the *Pāṇḍavas* and the *Kauravas*. The *Pāṇḍavas*, happily settled in *Indraprastha*, enjoy one glorious success after another. *Yuddhiṣṭhira*, after erecting the most magnificent hall on earth, decides to perform the Royal Consecration Sacrifice, which will raise his status to that of the world's greatest sovereign. His brothers travel far and wide and conquer all known kingdoms; the existing world sovereign is vanquished. Yet just when the *Pāṇḍavas* are beginning to seem invincible, *Yuddhiṣṭhira* mysteriously gambles everything away in a fateful game of dice to his cousin *Duryodhana*. The story is but a fraction of the infinite expanse of the *Mbh*, yet a pivotal one; the various changes in fortune set the scene for further conflict between the cousins. The grand narrative, rich in lingering, extravagant description, is a monument to the splendor of ancient India.⁷

According to BORI, the present structure of *Sabhāparva* is divided into 09 sub-parvans, 72 adhyāyas, 2390 ślokas with no ākhyānas. The first sub-parvan of *Sabhā* i.e., *Sabhāvarṇana* contains 11 adhyāyas from the beginning. The second *Mantra* (also called *Rājasūyamantra*) includes the 12th to the 17th adhyāyas. *Jarāsaṁdhavadhaparva*, the third sub-parvan contains 18th to 22nd adhyāyas. The fourth *Digvijayaparva* contains the 23rd to the 29th, fifth *Rājasūyikaparva* (also called *Rājasūya*) includes the 30th to the 32nd, sixth sub-parvan *Arghābhīharaṇa* (also called *Arghyābhīharaṇa*) contains the 33rd to the 36th, seventh *Śiśupālavadha* contains the 37th to the 42nd, *Dyūtaparva*, the eighth sub-parvan contains the 43rd to the 65th adhyāyas and the last ninth *Anudyūtaparva* contains the remaining 66th to the 72nd adhyāyas. The *Sabhā* which contains 2390 stanzas in BORI's edition generally is written in Anuṣṭubh chandas (in 2238

⁷ <http://www.claysanskritlibrary.org> (accessed on July 10, 08)

ślokas) but in some places it is written in Triṣṭubh-Jagatī (in 151 ślokas) and Halamukhī (only one) chandas also.

3.2.2.3. ĀRAṆYAKAPARVA

The *Āraṇyakaparva* (also known as *Vanaparva*, *Aranyaparva*) covers the twelve years of the *Pāṇḍavas*' exile in the forest, a penalty imposed upon them by the *Kāuravas* because they had lost a rigged dicing match. A number of colorful stories is told to relieve the tedium of life in *Āraṇyaka*. This parva consists: 'The Story of Rama', 'The Glorification of the Faithful Wife (*Sāvitrī*'s story)', 'The love story of *Nala-Damayantī*', 'The story of *R̥ṣyaśṛṅga* (the horned boy and *ṛṣi*)', 'The Robbing of the Earrings' and 'About the Drilling Sticks'. From a hero overcoming great odds, to a virtuous wife who rescues her family, and *Indra* tricking *Karṇa*, and *Yuddhiṣṭhira*'s victory in the verbal contest with the tree spirit, these disparate stories speak to common human concerns across cultures and centuries.⁸

The present structure shows that *Āraṇyakaparva* contains 16 sub-parvans in itself viz. *Āraṇyaka* (first 11 adhyāyas), *Kirmīravādha* (12th adhyāya), *Kairāta* (13-42 adhyāyas), *Indralokābhigamana* (43-79 adhyāyas), *Tīrthayātrā* (80-153 adhyāyas), *Jaṭāsūravādha* (154th adhyāya), *Yakṣayuddha* (155-172 adhyāyas), *Ājagara* (173-178 adhyāyas), *Mārkaṇḍeyasamasyā* (179-221 adhyāyas), *Draupadī-Satyabhāmasaṁvāda* (222-224 adhyāyas), *Ghoṣayātrā* (225-243 adhyāyas), *Mṛgasvapnabhaya* (244th adhyāya), *Vr̥hīdrauṇika* (245-247 adhyāyas), *Draupadīharaṇa* (248-283 adhyāyas), *Kuṇḍalāharaṇa* (284-294 adhyāyas) and *Āraṇeyaparva* (295-299 adhyāyas). The total is 299 adhyāyas and 10,338 ślokas.

3.2.2.4. VIRĀṬAPARVA

It is the fourth parva of the *Mbh* and details the *Pāṇḍavas*' 13th year in exile, when they live disguised in King *Virāṭa*'s court. The *Pāṇḍavas* suffer the humiliation of becoming servants; a topic explored both through comedy and pathos. They manage to maintain their disguise until the very end of the year, when their troubles truly begin. *Bhīma* is forced to come to *Draupadī*'s rescue when King *Virāṭa*'s general, *Kīcaka*, sets his sights on her. Later, taking advantage of his demise, *Duryodhana* and the *Trigartas* decide to invade *Virāṭa*'s kingdom, unaware the

⁸ <http://www.claysanskritlibrary.org> (accessed on July 07, 08)

Pāṇḍavas are hidden there. In the ensuing battles the *Pāṇḍavas* play a crucial role and, after saving *Virāṭa*, reveal their true identities to him. The parva ends on a note of celebration, with the *Pāṇḍavas* ready to return from exile and reclaim their kingdom. However, the *Virāṭa*'s battles foreshadow the war to come, proving that it will not be easy.⁹

This parva contains four sub-parvans viz. *Vairāṭa*, *Kīcakavadha*, *Gograhaṇa* and *Vaivāhikaparva*. Further it is separated into 67 adhyāyas and 1824 ślokas.

3.2.2.5. UDYOGAPARVA

The *Udyogaparva* begins with the completion of the thirteen year exile by the *Pāṇḍavas*. *Duryodhana* claims that they did not manage to live unknown for the full thirteenth year, since *Arjuna* was recognized in the battle at the end of the *Virāṭaparva*. This parva sees the *Pāṇḍavas* and *Kurus* gathering arms for the coming war and making preparations to fight. However, at the same time Lord *Kṛṣṇa* goes to *Hastināpura* as a messenger of peace but *Duryodhana*, the *Kuru* king, has no intention of negotiating. Both sides are well aware that war is the only outcome, and so this parva contains a great deal of discussion about *Dharma*, which in the context of the looming war, seems well-placed. Most characters are concerned that war between family cannot fail to be sinful. While there are many ślokas advice about caste duties and other general rules, this parva also contains the *Sanat-sujātīya*, a philosophical passage to rival the *Bhagavadgītā*. While not as famous, it contains a similar message, and appears to be a product of the same time and thinking. *Sanat-sujātā* teaches the Vedantic philosophy of seeking Brahman, the ultimate creative power, by truly understanding that one's soul and Brahman are one, and understanding that the universe as we know it is only illusion.¹⁰

It contains 11 sub-parvans (*Udyoga*, *Saṅjayayāna*, *Prajāgara*, *Sanatsujātā*, *Yānasamdhī*, *Bhagavadāyāna*, *Vivāda*, *Niryāṇa*, *Rathātīrathasamkhyā*, *Ulūkādūtāgamana* and *Ambopākhyāna*) which are further divided into 197 adhyāyas. The total number of ślokas in this parvan is 6063.

3.2.2.6. BHĪṢMAPARVA

The *Bhīṣmaparva* is the first of the six war-books (parvans) of the epic, recording the events of the first ten out the eighteen days of the great *Mbh* war, under the leadership if *Bhīṣma*. The

⁹ <http://www.claysanskritlibrary.org> (accessed on July 09, 08)

¹⁰ <http://www.claysanskritlibrary.org> (accessed on July 11, 08)

happenings of several fighting, first day comes between 16th to 45th adhyāyas, second from 46th to 51st adhyāyas, third from 52nd to 55th, fourth from 56th to 64th, fifth from 65th to 70th, seventh from 76th to 82nd, eighth from 83rd to 94th, ninth from 95th to 103rd and the great battle of tenth day exists from 104th to the last chapter i.e., 117th adhyāyas. The *Bhīṣmaparva* contains five sub-parvans (*Bhīṣmābhiṣecana*, *Jambukhaṇḍanirmāṇa*, *Bhūmi*, *Bhagavadgītā* and *Bhīṣmavadhāparva*), 117 adhyāyas and total 5406 śloka.

The most respected *Śrīmadbhagavadgītā* comes here between 23rd to 40th adhyāyas. In this ‘Bible’ of Indian civilization the charioteer *Kṛṣṇa* empowers his disciple *Arjuna* to resolve his personal dilemma: whether to follow his righteous duty as a warrior and slay his opponent relatives in the just battle, or to abstain from fighting and renounce the warrior code to which he is born.

3.2.2.7. DROṆAPARVA

After *Bhīṣma* is arrowed down by the *Arjuna* at the end of previous parva, *Duryodhana* selects *Droṇa* as leader of his forces. *Droṇa* accepts the honor with *Bhīṣma*’s blessing. The fighting rages on, with heavy losses on both sides. Furious and frustrated, *Duryodhana* accuses *Droṇa* of collaborating with the enemy, but he replies that as long as *Arjuna* is on the field, the *Pāṇḍavas* will remain unbeatable. When *Arjuna* is finally diverted from the main action of the battle, *Yudhiṣṭhira* entrust *Arjuna*’s son *Abhimanyu* with the task of making a breach in the *Kaurava* formation. *Abhimanyu* rampages through *Droṇa*’s army, but at last is cornered by several *Kaurava* warriors and finally killed by *Jayadratha*. Returning at night to the *Pāṇḍava* camp, *Yudhiṣṭhira* is racked with guilt and appalled by what he has to tell his brother. *Vyāsa* arrives in the *Pāṇḍava* camp and tries to alleviate the king’s sorrow.

The *Droṇaparva* contains eight sub-parvans (*Droṇābhiṣeka*, *Samśaptakavadha*, *Abhimanyuvadha*, *Pratijñā*, *Jayadrathavadha*, *Ghaṭotkacavadha*, *Droṇavadha* and *Narāyaṇāstramokṣaparva*) in 173 adhyāyas. The total number of śloka in this parva is 8192.

3.2.2.8. KARṆAPARVA

The *Karṇaparva* relates the events of the two dramatic days after the defeat of the great warriors and generals *Bhīṣma* and *Droṇa*, in which *Karṇa*, great hero and the eldest *Pāṇḍava*, leads the *Kaurava* army into combat. This parva depicts mighty battles in gory detail, sets the scene for

Karṇa's tragic death, and includes a remarkable verbal duel between *Karṇa* and his reluctant charioteer *Śalya*, the king of the *Madra*, as they hurl abuse at each other before entering the fray. This parva contains only sub-parvan viz. *Karṇaparva*. Further it is divided into 69 *adhyāyas* which contains 3871 *ślokas*.

3.2.2.9. ŚALYAPARVA

The *Śalyaparva*, ninth parva of *Mbh* portrays the last day of the great battle between the *Kauravas* and the *Pāṇḍavas*, recounting in gory detail the final destruction of King *Duryodhana* and his army. This parva focuses on *Śalya*'s short-lived role as general of *Duryodhana*'s army. Martial speeches, heroic duels, and bloody massacres abound on the battlefield, until finally *Śalya* is killed by king *Yudhiṣṭhira*. At *Śalya*'s death, king *Duryodhana* flees and takes refuge in a lake. *Duryodhana*, the heroic but flawed king of the *Kauravas*, meets his end when he is dishonorably defeated in battle by his arch enemy, *Bhīma*. Framing a fascinating account of the sacred sites along the river *Sarasvatī*, the duel poignantly portrays the downfall of a once great hero in the face of a new order governed by *Kṛṣṇa*, in which the warrior code is brushed aside in order to ensure the predestined triumph of the *Pāṇḍavas*.

This parva contains four sub-parvans viz. *Śalya*, *Hradapraveśa*, *Gadāyuddha* and *Sārasvataparva* and there are 3315 *ślokas* in 64 *adhyāyas*.

3.2.2.10. SAUPTIKAPARVA

It describes the slaying by *Aśvatthāmā* of the children of *Pāṇḍavas*, while they were sleeping. In this parvan there are three sub-parvans (*Sauptika*, *Aiṣika* and *Jalapradānikaparva*) in *adhyāyas*. The total number of *ślokas* is 772.

3.2.2.11. STRĪPARVA

The *Strīparva* shows the lament of *Gāndhārī* and other women after the demise of *Duryodhana* and other *Kauravas* in the great war of *Mbh*. This parva contains five sub-parvan (*Strī*, *Śrāddha*, *Abhiṣecanika*, *Cārvākanigraha* and *Gṛhapravibhāgaparva*) in 27 *adyāyas*. The total number of *ślokas* in *Strīparva* is 730.

3.2.2.12. ŚĀNTIPARVA

The *Śāntiparva* presented as the teachings of *Bhīṣma* as he lays dying on the battlefield, after the epic war between the *Pāṇḍavas* and *Kauravas*. After the ending of great battle *Pāṇḍavas* go to meet *Bhīṣma*, the then *Bhīṣma* teaches *Pāṇḍavas* about their *Rājadharmā*, *Āpaddharma* and *Mokṣadharmā*. This parva shows that the freedom enjoyed by these ancient thinkers was not an end in itself. The record of philosophers seeking liberation (*mokṣa*) from a world they believed unsatisfactory can be found in this parva.

This voluminous twelfth parva of *Mbh* contains three parvan (*Rājadharmā*, *Āpaddharma* and *Mokṣadharmaparva*) in 353 adhyāyas and includes a total of 12,902 ślokas.

3.2.2.13. ANUŚĀSANAPARVA

The *Anuśāsanaparva* has accounts of the stories relating to morals and righteousness of *Bhīṣma*. This parva also popular for *Viṣṇusahasranāma*,¹¹ the most famous hymn to *Viṣṇu*, which describes his 1000 names and for the *Śivasahasranāma Stotra*,¹² wherein *Śiva* is eulogized by *Kṛṣṇa*.

Anuśāsanika and *Bhīṣmasvargārohaṇaparva* are the two sub-parvan of this thirteenth parva of *Mbh*, which is further classified in 154 adhyāyas and a sum of 6439 ślokas.

3.2.2.14. AŚVAMEDHIKAPARVA

The *Aśvamedhikaparvan* covers the royal ceremony or *Aśvamedha* conducted by *Yudhisthira*. In this parvan, there are two sub-parvans viz. *Aśvamedhika* and *Anugītā* which are again sub-divided into 96 adhyāyās and a total of 2743 ślokas.

3.2.2.15. ĀŚRAMAVĀSIKAPARVA

The *Āśramavāsikaparva* starts with returning of *Vidura* to *Hastināpura* after some thirty-six years of great battle of *Kurukṣetra*. When *Vidura* returned to his *Hastināpura*, all the inhabitants hurried to see him in great delight. It so appeared that they had regained their consciousness after a long period. They all offered their obeisance and welcomed each other with embraces. Due to anxieties and long separation, they all cried out of affection. King *Yudhiṣṭhira* then arranged to

¹¹ <http://www.indianetzone.com/2/mahabharata.htm> (accessed on June 29, 08)

¹² <http://shivadarshana.blogspot.com/2007/09/shiva-sahasranama.html> (accessed on July 02, 08)

offer sitting accommodations and a reception. After some days *Vidura* convinced *Dhṛtarāṣṭra* to break the strong network of family affection and leave home to set out on the path of liberation. Finally, *Dhṛtarāṣṭra* left *Hastināpura* in the middle of the night unnoticed by anyone. *Gāndhārī*, *Kuntī* and *Vidura* followed him, seeing that he was going to the southern side of Himalaya Mountains. On the bank of *Saptasrota*, where the sacred Ganges were divided into seven branches, they engaged in rituals to get *Mokṣa* and after some time got their eventual death.¹³ This fifteenth canto of the great epic contains three sub-parvan viz. *Āśramavāsa*, *Putradarśana* and *Nārādāgamanaparva* in 47 adhyāyas. The total numbers of the ślokas are 1062.

3.2.2.16. MAUSALAPARVA

The *Mausalaparva* shows the infighting between the *Yādavas* with maces. After the battle of *Kurukṣetra*, *Gāndhārī* had cursed the *Yadu* dynasty to be destroyed by fratricide. She did this because she considered Lord *Kṛṣṇa* to be the cause of her sons' death. This parvan contains only one sub-parvan (*Mausalparvan*) which is divided into 09 adhyāyas and contains a sum of 273 ślokas.

3.2.2.17. MAHĀPRASTHĀNIKAPARVA

The *Mahāprasthānikaparva* covers the first part of the path to death of *Yudhisthira*, his brothers and their wife *Draupadī*. This has only one sub-parvan which is further divided into 03 adhyāyas and a sum of 106 ślokas.

3.2.2.18. SVARGĀROHAṆAPARVA

The eighteenth and the last parvan of the *Mbh* covers *Pāṇḍavas'* return to the spiritual world. This contains only one sub-parvan in 05 adhyāyas. The total number of ślokas are 194.

3.2.2.19. HARIVAMŚA

The *Harivamsaparva*, which is an appendix (*Khilaparva*) of *Mbh*, focuses specifically on the life and genealogy of Lord *Kṛṣṇa*. The two sub-parvan (*Harivamśa* and *Bhaviṣyatparva*) of this appendix contains 16,375 verses.

¹³http://nitaaveda.com/All_Scriptures_By_Acharyas/Historical_Works/Mahabharata/Ashramvasika_Parva.htm
(accessed on July 02, 08)

It is not clear when this kind of division into parvans took place. The parvans are of different length and there is no uniformity regarding their size. The shortest is *Mahāprasthānikaparva* containing only 106 stanzas, whereas the longest is *Śāntiparva*. The combination of *Śāntiparva* and *Anuśāsanaparva* yields 21,000 stanzas. An overview of structure of *Mbh* is given below:

| Mahābhārata | | | | |
|---|-------------------|-----------------|----------------|---------------|
| (According to the BORI's critical edition) | | | | |
| Parvan | Sub-parvan | Ākhyāna* | Adhyāya | Ślokaś |
| Ādiparva | 19 | 09 | 225 | 07197 |
| Sabhāparva | 09 | --- | 072 | 02390 |
| Āraṇyakaparva | 16 | --- | 299 | 10338 |
| Virāṭaparva | 04 | --- | 067 | 01824 |
| Udyogaparva | 11 | --- | 197 | 06063 |
| Bhīṣmaparva | 05 | --- | 117 | 05406 |
| Droṇaparva | 08 | --- | 173 | 08192 |
| Karṇaparva | 01 | --- | 069 | 03871 |
| Śalyaparva | 04 | --- | 064 | 03315 |
| Sauptikaparva | 03 | --- | 018 | 00772 |
| Strīparva | 05 | --- | 027 | 00730 |
| Śāntiparva | 03 | --- | 353 | 12902 |
| Anuśāsanaparva | 02 | --- | 154 | 06439 |
| Aśvamedhikaparva | 02 | --- | 096 | 02743 |
| Āśramavāsikaparva | 03 | --- | 047 | 01062 |
| Mausalaparva | 01 | --- | 009 | 00273 |
| Mahāprasthānikaparva | 01 | --- | 003 | 00106 |
| Svargarohanaparva | 01 | --- | 005 | 00194 |
| Total (in 18 parvans) | 98 | | 1995 | 73,817 |

* The accurate number of *Ākhyānas* could not be determined in spite of an extensive effort.

| | | | | |
|--------------------------------|------------|-----|-------------|---------------|
| Harivaṁśa (Khilaparva) | 02 | --- | ---- | 16,375 |
| Total (including Khila) | 100 | --- | 1995 | 90,192 |

[Table 3.3: Structure of *Mbh.* according to BORI's critical edition]

3.3. Ādiparva of Mahābhārata

3.3.1. Introduction to Ādiparva of Mahābhārata

The *Ādiparva*, first canto of the great epic *Mbh*, has been divided into 19 sub-parvans in which the first *Anukramaṇīparva* gives summary of the *Mbh* and also shows *Dhṛtarāṣṭra*'s lament of 'no hope of success'. The second *Parvasaṁgrahaparva* gives overview of chapters and contents of the epic. The third sub-parvan known as *Pauṣyaparva* narrates the glories of *Uttan̥ka*, and the fourth *Paulomaparva* fully describes the expansion of the *Bhṛgu* dynasty. Fifth, the *Āstikaparvan* explains the origin of all snakes and of the great *Garuḍa*; the churning of the milk ocean; the birth of the celestial steed *Uccaiṣṛavā*; tales of the great *Bharata* kings related at Janamejaya's snake sacrifice; and the origin of various kings and sages and of the great sage *Vyāsa*. The parvan entitled Descent of the First Created Beings describes the origin of demigods, Daityas, Dānavas, and Yakṣas. This parvan also tells the origin of Nāga serpents, snakes, Gandharvas, birds, and various other creatures. The sixth *Ādivaṁśāvataranaparva* narrates the story of the Vasus that how these great souls were forced to take birth from the womb of goddess *Gaṅgā* in the house of King *Śāntanu* and how they regained their position in heaven. All the Vasus invested a portion of their potency in a single *Vasu*, and that one took birth as *Bhīṣma*, who later renounced his father's kingdom, taking the difficult vow of celibacy, which he kept with great determination. *Bhīṣma*'s younger step-brother *Citrāṅgada* assumed the *Kuru* throne under *Bhīṣma*'s guidance, but when the young monarch was killed, *Bhīṣma* installed *Citrāṅgada*'s younger brother *Vicitravīrya* as king and took care of him. Next of the previous sub-parvan, the seventh *Sambhavadaparva* tells how *Vyāsa*, to keep his promise to his mother, begot *Dhṛtarāṣṭra*, *Pāṇḍu*, and *Vidura*, who was actually *Dharma*, lord of justice, forced to take birth as a human being by the curse of the sage *Māṇḍavya*. The eighth *Jatugrhādāhparvan* described are the births of *Pāṇḍu*'s sons, their journey to *Vāraṇāvarta*, the wicked plotting of their cousin *Duryodhana* to kill them, and their effort to escape, based on *Vidura*'s wise counsel, through a

secret underground tunnel. The ninth **Hiḍimbavadhaparva** narrates the meeting of the *Pāṇḍavas* with *Hiḍimba* in the frightful forest; the birth of *Ghaṭotkaca* from that forest encounter; the tenth **Bakavadhaparva** describes the *Pāṇḍava*'s activities while living in disguise in the house of a *Brāhmaṇa*; and the slaying of the monstrous *Baka*, which amazed the *Brāhmaṇa* and all the inhabitants of his city. The eleventh **Caitrarathaparva** described the births of lovely *Draupadī* and her fiery twin brother, *Dhr̥ṣṭadyumna*. Hearing about *Draupadī* from a *Brāhmaṇa* and encouraged as well by the words of *Vyāsa*, the *Pāṇḍavas* decided to win *Draupadī*'s hand in marriage. They eagerly set out for the kingdom of *Pāñcāla*, which was ruled by *Draupadī*'s father, to compete at her *svayamvara* ceremony, where she was to choose her husband. On the way, *Arjuna* defeated the *Gandharva* king *Aṅgārāparṇa* on the banks of the Ganges. Having formed a lasting friendship with him, and having heard many stories from him, *Arjuna* then traveled on with all his brothers towards the kingdom of the *Pāñcālas*. Narrated in this section are the excellent stories of *Tāpatya*, *Vaśiṣṭha*, and *Aurva*. The twelfth **Draupadīsvayamvaraparva** narrates the happenings in the city of *Pāñcāla*, where *Arjuna* alone among all the kings of the earth could pierce an exceptionally difficult target with his arrow, thus winning *Draupadī*'s hand. The losing kings, headed by *Karṇa* and *Śalya*, were furious, but *Bhīma* and *Arjuna* defeated them in a great battle. Seeing the unfathomable, superhuman prowess of *Bhīma* and *Arjuna*, Lord *Kṛṣṇa* and His elder brother, *Balarāma*, recognized them as the sons of *Pāṇḍu*, even though all of the *Pāṇḍavas* were carefully disguised as *Brahmaṇas*. The great minded brothers, *Kṛṣṇa* and *Balarāma*, then went to visit the sons of *Pāṇḍu* at their lodgings in the house of a potter. The thirteenth **Vaivāhikaparva** covers the amazing story of the five *Indras*. It also shows the amazement of King *Drupada*, over the fact that his daughter *Draupadī* was to marry all five *Pāṇḍava* brothers, but Lord *Śiva* had blessed her to enjoy an extraordinary marriage. The fourteenth **Vidurāgamanaparva** and fifteenth **Rājyalambhaparva** narrate the happenings where *Dhṛtarāṣṭra* sent *Vidura* to see the *Pāṇḍavas*, and upon his arrival *Vidura* also met with Lord *Kṛṣṇa*. To prevent a quarrel between the *Pāṇḍavas* and *Kurus*, the kingdom was divided, and the *Pāṇḍavas* went to live in the city of *Khāṇḍavaprastha*. The sixteenth sub-parvan **Arjunavanavāsaparva** tells the story of *Sunda* and *Upasunda*. By the order of *Nārada Muni* the five brothers agreed to spend equal time alone with their lovely wife *Draupadī* (each brother swore that if he ever intruded when another brother was with *Draupadī*, the intruding brother would voluntarily accept banishment). Soon thereafter, when the eldest brother, *Yudhiṣṭhira*, was

alone with *Draupadī*, *Arjuna* unavoidably entered their room to get a weapon that he needed to help a saintly *Brāhmaṇa*. After rescuing the stolen property of the *Brāhmaṇa*, *Arjuna*, determined to honor the *Pāṇḍavas*' mutual pact, left the royal palace and went alone to the forest. Next comes the story of *Arjuna*'s union with the princess *Ulupī*, whom he met on the path while dwelling in the forest; after that, *Arjuna*'s pilgrimage to many sacred spots and the birth of *Babhruvāhana* are described. During that time *Arjuna* saved five *Apsarās* who have been cursed by an ascetic *Brāhmaṇa* to take birth as crocodiles. The seventeenth ***Subhadrāharaṇaparva*** covers the happenings where *Arjuna* met with Lord *Kṛṣṇa* at the holy land of *Prabhāśakṣetra* and went with Him to His capital of *Dvārakā*, (a fabulous city built on the surface of the ocean). While there *Arjuna* fell in love with *Kṛṣṇa*'s lovely young sister, *Subhadrā*, and she also fell in love with *Arjuna*. Taking Lord *Kṛṣṇa*'s permission, *Arjuna* eloped with her. The eighteenth ***Haraṇahārikaparvan*** and nineteenth ***Khāṇḍavadāhaparvan*** narrates the happenings where Lord *Kṛṣṇa*, son of *Devakī*, brought a dowry for His new brother-in-law. Upon arriving at the *Pāṇḍava*'s capital (*Khāṇḍavaprastha*, also known as *Indraprastha*), *Śrī Kṛṣṇa* acquired His famous whirling weapon, the *Sudarśana* disc, and *Arjuna* acquired his famous *Gāṇḍīva* bow. The *Khāṇḍava* forest was burned to ashes, and *Subhadrā* gave birth to mighty *Abhimanyu*. *Arjuna* saved the great mystic *Maya* from the fiery forest, while a special serpent escaped. The great sage *Mandapālā* begot a son in the womb of the bird *Śārṅga*.

3.3.2. Structure of Ādiparva (according to BORI's edition)

The liberated sage *Vyāsa* affirms in ***Parvasamgraha*** that *Ādiparva* contains 218 chapters composed of 7,984 verses.¹⁴ But going through the Critical Edition of *Mahābhārata*, we get in the *Ādiparva* have 19 sub-parvans and 09 ākhyanas in 225 adhyāyas. The total number of śloka is 7,197. The tabular form of the structure of this first canto is given below:

¹⁴ ityetadādhīparvoktaṁ prathamam bahuvistaram |
adhyāyānāṁ śate dve tu sanikhyāte paramarṣiṇā |
aṣṭādaśaiva cādhyāyā vyāsenottamatejasā ||
sapta ślokaśahasrāṇi tathā nava śatāni ca |
ślokāśca caturāśītirdṛṣṭo grantho mahātmanā ||Mbh.01.02.02.95-96||

| Ādiparva (according to BORI's edition) | | | |
|--|--|-----------------|-------|
| Upaparva | Ādhyāya | Ākhyāna | Śloka |
| 1. Anukramaṇīparva | 01 (i.e., 1 st adhyāya only) | ---- | 210 |
| 2. Parvasaṁgraha | 01 (i.e., 2 nd adhyāya only) | ---- | 243 |
| 3. Pauṣyaparva | 01 (i.e., 3 rd adhyāya only) | ---- | 195 |
| 4. Paulomaparva | 09 (from 4 th to 12 th) | ---- | 150 |
| 5. Āstikaparva | 40 (from 13 th to 53 rd) | ---- | 1025 |
| 6. Ādivaṁśāvatarapaṇaparva | 08 (from 54 th to 61 st) | ---- | 482 |
| 7. Saṁbhavaparva | 62 (from 62 nd to 123 rd) | 03 [*] | 2169 |
| 8. Jatugṛhadāhaparva | 15 (from 124 th to 138 th) | ---- | 373 |
| 9. Hiḍimbavadha | 04 (from 139 th to 142 nd) | ---- | 111 |
| 10. Bakavadha | 10 (from 143 rd to 152 nd) | ---- | 264 |
| 11. Caitraratha | 21 (from 153 rd to 173 rd) | 03 [#] | 554 |
| 12. Draupadīsvayaṁvara | 16 (from 174 th to 189 th) | 01 [*] | 381 |
| 13. Vaivāhika | 02 (from 190 th and 191 st) | ---- | 37 |
| 14. Vidurāgamanaparva | 07 (from 192 nd to 198 th) | ---- | 174 |
| 15. Rājyalambha | 01 (consists only 199 th adhyāyas) | ---- | 50 |
| 16. Arjunavanavāsa | 11 (from 200 th to 210 th) | 01 [⊗] | 295 |
| 17. Subhadrāharaṇa | 02 (211 th and 212 th) | ---- | 57 |
| 18. Haraṇahārika | 01 (consists only 213 rd adhyāya) | ---- | 82 |
| 19. Khāṇḍavadāha | 12 (from 214 th to 225 th) | 01 [∇] | 344 |

[Table 3.4: Structure of Ādiparva of Mbh. according to BORI's critical edition]

* These three are- *Śakuntalopākhyāna* (starts at 62nd and end with 69th adhyāya and includes a sum of 305 ślokas), *Yayātyupākhyāna* (starts at 70th and end with 80th adhyāya and composed of 365 ślokas) and *Uttarayāyātam* (starts at 81st and end with 88th adhyāya and a sum of 151 ślokas).

The *Caitrarathaparva* also contains three ākhyānas viz. *Tāpatyopākhyāna* (it begins from 160th adhyāya and end with 163rd consisting of 102 ślokas), *Vāśiṣṭhopākhyāna* (includes a sum of 173 ślokas in 164th to 168th and also in 173rd adhyāya) and *Aurvopākhyāna* (in 169th to 172nd adhyāyas and a sum of 86 ślokas).

* This subparvan has only ākhyāna namely, *Pañcendropākhyāna* which composed of 49 ślokas in its last chapter.

⊗ The *Sundopasunda* is only upākhyāna which includes 119 ślokas from 201st to 204th adhyāya.

∇ The only ākhyāna *Śārngakopākhyāna* which includes 147 ślokas of last six chapters of the Ādiparva (i.e., from 220 to 225).

3.3.2. Computer adaptation of Ādiparvan

Ādiparva, the first canto of *Mbh.* has a multi-layered division. It is primarily divided into sub-parvans and then further divided into adhyāyas which are composed of several stanzas. The Ākhyānas are also an important part of the structure but they cannot be regarded as a sub-division because the maximum portion of the parva is not included in any of them. Thus each stanza has two types of hierarchy; (a) compulsory i.e., Parva> Upaparva> Adhyāya, (b) optional i.e. Ākhyāna which comes under upaparva and contains some adhyāyas.

Mbh → Parva → Upa-parva → [Ākhyāna] → Adhyāya → Śloka

This structure of the *Ādiparva* is adapted to prepare an RDBMS based system for online indexing of the text. Each layer of hierarchy is stored in separate tables. Each section of every level is given a unique identity and adjoined to the tables of other levels. The basic table is “**shloka**” which contains the columns *ShlokaSamhita*, *Shlokapada*, *Shloka_Id*, *Adhyaya_Id* and *Shloka_Nm*. The *Shloka_Id* is the unique id of each stanza. *Shloka_Nm* is the serial number of the stanzas in the corresponding adhyāyas. The column *Adhyaya_Id* denotes the unique id of adhyāya which is connected with ‘*Id*’ column of “**adhyaya**” table. The structure for database storage is as follows:

| Shloka_Id | Adhyaya_Id | Shloka_Nm | ShlokaSamhita | ShlokaPada |
|-----------|------------|-----------|---|---|
| 01 | 01 | 00 | नारायणं नमस्कृत्य नरं चैव नरोत्तमम् । देवीं सरस्वतीं चैव ततो जयमुदीरयेत् ॥ | नारायणम् नमः कृत्य नरम् च एव नर उत्तमम् । देवीम् सरस्वतीम् च एव ततः जयम् उदीरयेत् ॥ |
| 02 | 01 | 01 | लोमहर्षणपुत्र उग्रश्रवाः सूतःपौराणिको । नैमिषारण्ये शौनकस्य कुलपतेर्द्वादशवार्षिके सत्रे ॥ | लोमहर्षणपुत्र उग्रश्रवाः सूतः पौराणिकः । नैमिष अरण्ये शौनकस्य कुलपतेः द्वादशवार्षिके सत्रे ॥ |

| | | | | |
|------|-----|-----|--|---|
| 03 | 01 | 02 | समासीनानभ्यः छद्ब्रह्मर्षीन् संशितव्रतान् । विनयावनतो भूत्वा कदाचित् सूतनन्दनः ॥ | समासीनान् अभ्यः छद् ब्रह्मर्षीन् संशितव्रतान् । विनय अवनतः भूत्वा कदाचित् सूतनन्दनः ॥ |
| --- | --- | --- | --- | --- |
| --- | --- | --- | --- | --- |
| 2603 | 69 | 43 | तामेवमुक्त्वा राजर्षिर्दुःषन्तो महिषीं प्रियाम् । वासोभिरन्नपानैश्च पूजयामास भारत ॥ | ताम् एवम् उक्त्वा राजर्षिः दुःषन्तः महिषीम् प्रियाम् । वासोभिः अन्नपानैः च पूजयामास भारत ॥ |

[Table 3.5: Structure of database storage, it is first table among five called ‘shloka’]

The table “**adhyaya**” has a description of the adhyāyas in *Id*, *UP_Id*, *Akh_Id*, and *Adh_Nm* columns. The column *Id* represents the uniqueness of adhyāyas. As mentioned, it relates the description of adhyāya to each stanza. The *UP_Id* denotes the number of sub-parvans of Ādiparva which is bound to *Id* column of the table “**upaparva**”. The column *Akh_Id* is a unique identification number of ākhyāna and is linked up to *Id* column of the table “**akhyana**”. The column *Adh_Nm* gives the number of adhyāyas in Sanskrit words.

| Id | UP_Id | Akh_Id | Adh_Nm |
|-----|-------|--------|-----------------|
| 01 | 01 | 00 | प्रथमोऽध्यायः |
| 02 | 02 | 00 | द्वितीयोऽध्यायः |
| 03 | 03 | 00 | तृतीयोऽध्यायः |
| --- | --- | --- | --- |
| --- | --- | --- | --- |

| | | | |
|----|----|----|---------------------|
| 62 | 07 | 01 | द्विषष्टितमोऽध्यायः |
| 63 | 07 | 01 | त्रिषष्टितमोऽध्यायः |

[Table 3.6: the second ‘adhyaya’ table works as bridge table between previous and next two tables]

The table ‘**akhyana**’ contains two columns viz. *Id* and *Akhyana*. The column *Id* is adjoined to *Akh_Id* of ‘**adhyaya**’ table and it is used to relate the name of ākhyāna with adhyāya. The *Akhyana* column gives the name of ākhyāna.

| Id | Akhyana |
|----|-----------------|
| 01 | शकुन्तलोपाख्यान |
| 02 | ययात्युपाख्यान |
| 03 | उत्तरयायातम् |
| 04 | तापत्योपाख्यान |
| 05 | वाशिष्ठोपाख्यान |

[Table 3.7: the table ‘akhyana’ helps to show the ākhyana]

The table “**upaparva**” contains *Id*, *Parva_Id* and *Upaparva* columns. The *Id* column is adjoined to *UP_Id* column of table “**adhyaya**”. The *Parva_Id* column gives the unique number of each parva of *Mbh.* and it is adjoined with the column *Id* of table ‘**parva**’. The *Upaparva* column of this table shows the name of sub-parvans in Unicode Devanagari.

| Id | Parva_Id | Upaparva |
|----|----------|-----------------|
| 01 | 01 | अनुक्रमणिकापर्व |

| | | |
|----|----|-----------------|
| 02 | 01 | पर्वसङ्ग्रहपर्व |
| 03 | 01 | पौण्ड्यपर्व |
| 04 | 01 | पौलोमपर्व |

[Table 3.8: the fourth table ‘upaparva’ gives the sub-parvan reference and also calls to fifth for completing the reference]

The table “**parva**” has two columns viz. *Id* and *Parva* where the *Id* column shows the unique number of each parvan of *Mbh.* and it is connected with *Parva_Id* column of table ‘Up-4’. The *Parva* column of this table gives the name of each parvan of *Mbh.* in Unicode Devanāgarī.

| Id | Parva |
|----|------------|
| 01 | आदिपर्व |
| 02 | सभापर्व |
| 03 | वनपर्व |
| 04 | विराटपर्व |
| 05 | उद्योगपर्व |

[Table 3.9: the table ‘Parva’ contains the unique id and name of all parvans of *Mbh.*]

Chapter-4

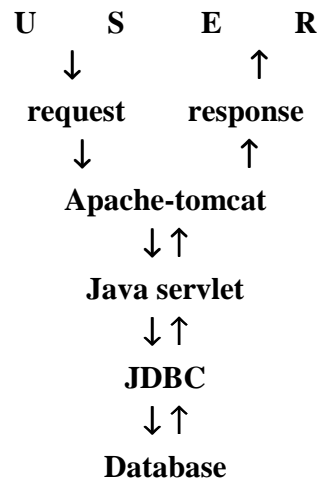
Online Indexing of Mahābhārata (Ādiparva)

4.1. Introduction

This chapter describes the partial implementation of Online Indexing of *Mahābhārata* (Ādiparvan) as part of the present M. Phil. R&D. The computational model uses Java in the web format for the indexing of words occur in *Mahābhārata* (Ādiparvan) through the identification and connection with original ślokas stored in the database. The system accepts three kinds of searches and gives analyzed output in the same format. The first input mechanism is ‘**Direct Search**’ where the user can enter any key-word in Devanāgarī UTF-8 and get all the references and details from *Mahābhārata* (Ādiparvan). The second is ‘**Alphabetical Search**’ facility where one can just click any letter of Devanāgarī alphabet to get the index of the words starting with that alphabet, and the third input mechanism is ‘**Search by Classes**’ where the user can click on “Parva” → “Upa-parvas” → “Adhyāyas” → “Ākhyānas” to get the index. Clicking on an indexed word will display the details with the śloka in which it occurs. It also gives facility to search that word in some other online lexical resources.

4.2. Architecture of the system

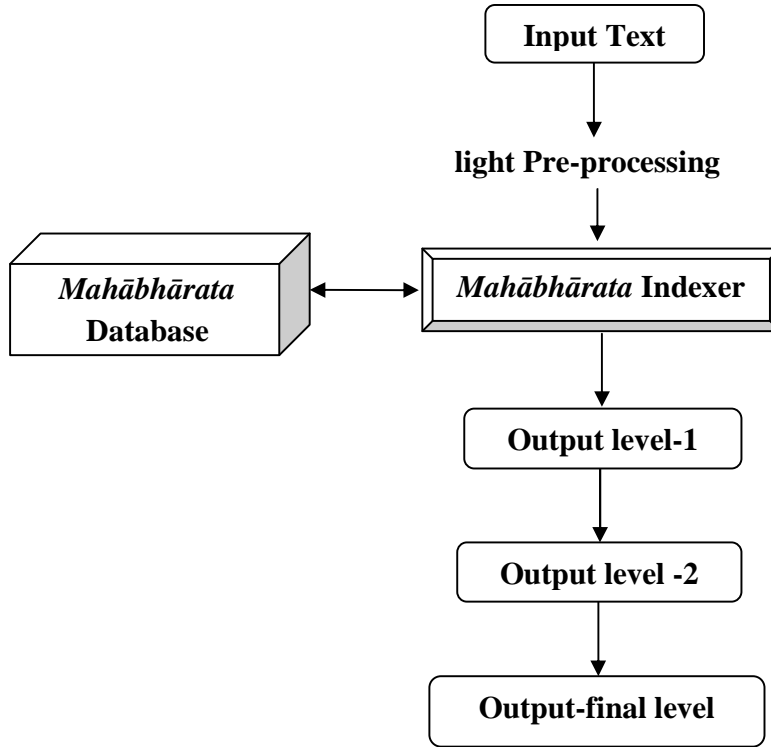
The indexing system of *Mahābhārata* (Ādiparvan) is developed in front-end of JSP with Java servlet, Back-end of RDBMS and JDBC connectivity. The webserver for Java/JSP is Apache Tomcat 4.0 and the RDBMS used is MS-SQL Server 2005 Database in Unicode scheme. The following model describes the interaction between multi-tier architecture of the indexing system of *Mahābhārata*:



[Fig 4.1: Multi-tier architecture of the Mahabharata Indexer]

4.3. Process of the Indexing system

There are three ways to give input to the system e.g. Direct Search, Alphabet search and Search by the structure of the text in Devanāgarī UTF-8 format



[Fig 4.2: Process of Indexing system]

Step I: Preprocessing

Preprocessing a text mainly consists of normalizing it.

Step II: Mahābhārata Indexer and Database

At this step, the indexer makes an indexed list of exact and partially matching words. Getting the query as an input, the indexer, after a slight preprocessing, sends it to the database. If the word has its occurrence in the database, the system gives the output.

Step III: Output level-1

At this level, the indexer gives all the occurrences of the searched query with its numerical reference in a hyperlinked mode.

Step IV: Output level -2

Clicking any hyperlinked word, system shows its original place in the śloka and also gives its full reference in the text. It also asks for further information from other online lexical resources.

Step V: Output - final level

Here, the indexer gives a list of online lexical resources and also gives the facility to do morphological analysis of the query with the help of POS Tagger¹ and Subanta Analyzer.²

4.4. The Front-end of the Mahābhārata Indexer

4.4.1. Java Server Pages

Java Server Pages (JSP) technology provides a simplified, fast way to create dynamic web content. JSP technology enables rapid development of web-based applications that are server and platform-independent.³ It is one of the most sophisticated tools available for high performance and secures web applications.

The front-end of the system is developed in utf-8 enabled Java Server Pages (JSP) and HTML. The front-end of the software enables the user to interact with the indexing system with the help of Apache Tomcat web-server. The JSP technology helps to create web based applications by combining Java code with HTML. The web server runs the Java code and displays the results as HTML. For this system, there are two JSP pages, one is the main search page and the other is the cross-referential search page which searches the searched query in different online lexical and linguistic resources. The below written code snippet instructs the page to set the language and content encoding of input and output UTF-8:

```
<%@ page
    language=" java "
    pageEncoding="utf-8 "
    contentType="text/html; charset=utf-8 "
    import=" java.util.* "
%>
```

¹ <http://sanskrit.jnu.ac.in/post/post.jsp>

² <http://sanskrit.jnu.ac.in/subanta/rsubanta.jsp>

³ <http://java.sun.com/products/jsp/>

The following code initializes the different values of main class, strings and integers:

```
<%
    Mahabharata mb = new Mahabharata();
    akhyanasFound = 1;
    request.setCharacterEncoding("UTF-8");
    String searchtype = "";
    int parva = 0;
    int upaparva = 0;
    int adhyaya = 0;
    int akhyana = 0;
    int shloka = 0;
    String ddfocus="parva";
    String token="";
    String searchstr="";
```

The following code obtains the different values of the search queries:

```
try{
    ddfocus = request.getParameter("ddfocus");
}
catch(Exception e){
    ddfocus="parva";
}
try{
    searchtype = request.getParameter("searchtype");
}
catch(Exception e){
    searchtype="partial";
}
try{
    searchstr = request.getParameter("searchstr");
}
catch(Exception e){
    searchstr="";
}
try{
    parva = Integer.parseInt(request.getParameter("parva"));
```

```
}  
catch(Exception e){  
    parva=0;  
}  
  
try{  
    upaparva = Integer.parseInt(request.getParameter("upaparva"));  
}  
catch(Exception e){  
    upaparva=0;  
}
```

The following code defines the values of different strings:

```
if (searchtype==null)  
    searchtype="partial";  
if (searchstr==null)  
    searchstr="";  
if (ddfocuss==null)  
    ddfocuss="parva";  
if (token==null)  
    token="";
```

The following code draws the text box and search button for the exact search query:

```
DIRECT SEARCH (unicode Sanskrit)  
<% if (searchstr.length()>0) { %>  
    <br><input type=text name=searchstr value="<%= searchstr %>" >  
<% } else { %>  
    <br><input type=text name=searchstr value="<%= token %>" >  
<% } %>  
<input type=submit value="search Mahabharata Database">
```

The following code searches the result of exact search query entered in the text box as partial search:

```
<% if (searchtype.equals("partial") && searchstr !=null &&  
searchstr.length()>0){ %>  
    <%= mb.searchIndex(searchstr) %>
```

The code which provides the list of alphabets for partial alphabetical search:

```
ALPHABET SEARCH<br>
<a href=ibasic.jsp?searchtype=partial&searchstr=अ>अ</a>
<a href=ibasic.jsp?searchtype=partial&searchstr=आ>आ</a>
<a href=ibasic.jsp?searchtype=partial&searchstr=इ>इ</a>
<a href=ibasic.jsp?searchtype=partial&searchstr=ई>ई</a>
.
.
.
```

This code takes the input by clicking on a letter, and displays search result through the `alphabetSearch()` function of Mahabharata class:

```
<% if (searchtype.equals("partial") && searchstr !=null &&
searchstr.length()>0){ %>
    <%= mb.alphabetSearch(searchstr) %>
```

The following code gets the search input by selecting the class of the text (e.g. Parva, Upa-parva, Adhyāya and Ākhyāna):

```
SEARCH BY CLASS
<select name=parva>
<option value=1 <% if (parva==1){ %> selected <% } %> >आदिपर्व</option>
<option value=2 <% if (parva==2){ %> selected <% } %> >सभापर्व</option>
<option value=3 <% if (parva==3){ %> selected <% } %> >वनपर्व(आरण्यक)</option>
<option value=4 <% if (parva==4){ %> selected <% } %> >विराटपर्व</option>
... ..
```

The following code gives the result of class search as the list of words in the selected terminal class:

```
<% if (searchtype.equals("partial")){ %>
<% if ( ddfocus.equals("adhyaya") || ddfocus.equals("akhyana") ){ %>
```

If Adhyaya is selected and Akhyana is not found give the result from that Adhyaya.

```
<% if ( adhyaya>0 && akhyanasFound == 0 ) { %>
    <%=mb.getIndexForAdhyaya(adhyaya) %>
```

If Akhyana is selected and it is found, give the result according to that Akhyana.

```
<%} else if (akhyana>0 && akhyanasFound ==1) { %>
    <%=mb.getIndexForAkhyana(akhyana) %>
```

The following code displays the result generated when a particular search token is submitted.

```
Results
<table>
<td><b>Index Search for <font>'<%=token %>'</font></b>
<td><b>Reference:</b><td><%= mb.getShlokaRefActual()
%>(<%=mb.getShlokaRefDescriptive() %>)
    <td><b>Shloka:</b><td><%= mb.getShlokaText() %>
    <td><b>Parva:</b><td><%= mb.getParva() %>
    <td><b>Upaparva:</b><td><%= mb.getUpaparva() %>
    <td><b>Adhyaya:</b><td><%= mb.getAdhyaya() %>
    <td><b>Akhyana:</b><td><%= mb.getAkhyana() %>
```

This code gives options to search the queried word in different online resources.

```
<td><b>Search other sources:</b> <td>
<a href=searchNet.jsp?word=<%= token %>>search this word on other online
resources</a>
```

The following code of cross-referential JSP page searches the searched string in other online resources.

```
<a
href="http://sanskrit.jnu.ac.in/amara/viewdata.jsp?searchtype=direct&searchst
r=<%= word %>">search Amarakosha(JNU)</a>
```



```
<a href="http://www.spokensanskrit.de/index.php?script=HK&tinput=
<%=word %>&country_ID=&trans=Translate&direction=AU">search "Spoken Sanskrit"
Dictionary (by Klaus Glashoff, Germany)</a>
```

4.4.2. Java Servlet Technology

Java Servlet technology provides web developers with a simple, consistent mechanism for extending the functionality of a web server and for accessing existing business systems. A servlet can almost be thought of as an applet that runs on the server side--without a face. Java servlets make many web applications possible. Servlets have access to the entire family of Java APIs, including the JDBC API to access enterprise databases. Servlets can also access a library of HTTP-specific calls and receive all the benefits of the mature Java language, including portability, performance, reusability, and crash protection.⁴

For the indexing system of *Mahābhārata*, the following code snippet of Java Servlet Technology has been used. This code snippet imports the java packages to be used in this class:

```
import java.lang.*;
import java.util.*;
import java.io.*;
import java.sql.*;
```

It is the main class:

```
public class Mahabharata {
}
```

It is the class constructor which creates a copy of the class to use temporarily:

```
public Mahabharata(){
    .. .. .
}
```

The following method loads the configuration of data files:

```
public void loadConf(){
    .. .. .
}
```

⁴ <http://java.sun.com/products/servlet/> (accessed on 11/07/08)

The following code gets the list of *Upa-parva* of a selected *Parva*:

```
public Hashtable getUpaparvaByParva(int parva,int upaparva){
    .. .. .
    return upaparvas;
}
```

The following code gets the list of *Adhyaya* of a selected *Upaparva*

```
public Hashtable getAdhyayaByUpaparva(int upaparva, int adhyaya){
    .. .. .
    return adhyayas;
}
```

The following code gets the list of *Akhyana* (if available) of a selected *Adhyaya*:

```
public Hashtable getAkhyanaByAdhyaya(int adhyaya, int akhyana){
    .. .. .
    return akhyanas;
}
```

The following code gets the list of words in selected *Akhyana* or *Adhyaya*:

```
public String getIndexForAkhyana(int akhyana){
    .. .. .
    return "<b>search found "+tknCount+" results for above
akhyana</b><br><b>"+ r + "</b>";
}
```

The following code is the main function which searches the queried word:

```
public String searchIndex(String word){
    .. .. .
    return "<b>Exact Search found "+tknCount+" results for
'"+word+" '</b><br><b>"+ r + "</b>";
}
```

The following code searches the alphabetical partial string search:

```
public String alphabetSearch(String alph){
    .. .. .
    return "<b>Alphabet search found "+tknCount+" results for
'"+alph+" '</b><br><b>"+ r + "</b>";
}
```

The following function gives the list of words of a selected *Adhyaya*

```
public String getIndexForAdhyaya(int adhyaya){
    .. .. .
    return "<b>search found "+tknCount+" results for above
adhyaya</b><br><b>"+ r + "</b>";
}
```

The following method organizes the search result in the incremental order of reference

```
public void getShlokaById(int shlokaid, String tkn){

    setShlokaid_incremental(shlokaid);
    setBaseWord(tkn);
    ResultSet rs = null;
}
```

The following function gives description of the base word in the either case whether *Akhyana* occurs or not:

```
public String getShlokaRefDescriptive(){
    if (getAkhyanaid()==0)
        return
getParva()+">"+getUpaparva()+">"+getAdhyaya()+">"+getShlokaid();
    else
        return
getParva()+">"+getUpaparva()+">"+getAdhyaya()+">"+getAkhyana()+">"+getShlokai
d();
}
```

The following function gives reference of the Shloka:

```
public String getShlokaRef(){
    return
getParvaId()+"."+getUpaparvaId()+"."+getAdhyayaId()+"."+getAkhyanaid()+"."+ge
tShlokaid_incremental();
}
```

The following function gives actual reference of the Shloka:

```
public String getShlokaRefActual(){
    return
    getParvaId()+"."+getUpaparvaId()+"."+getAdhyayaId()+"."+getAkhyanaId()+"."+ge
    tShlokaId();
}
```

The following function gives the Shloka in which the searched base word occurs. If *Shloka pada* available then it gives *Shloka Pada* otherwise *Shloka Samhita*.

```
public String getShlokaText(){
    if (getShlokapada().length()>0)
        return shlokapada;
    else
        return shlokasamhita;
}
```

4.4.3. Apache Tomcat 4.0 Web Server

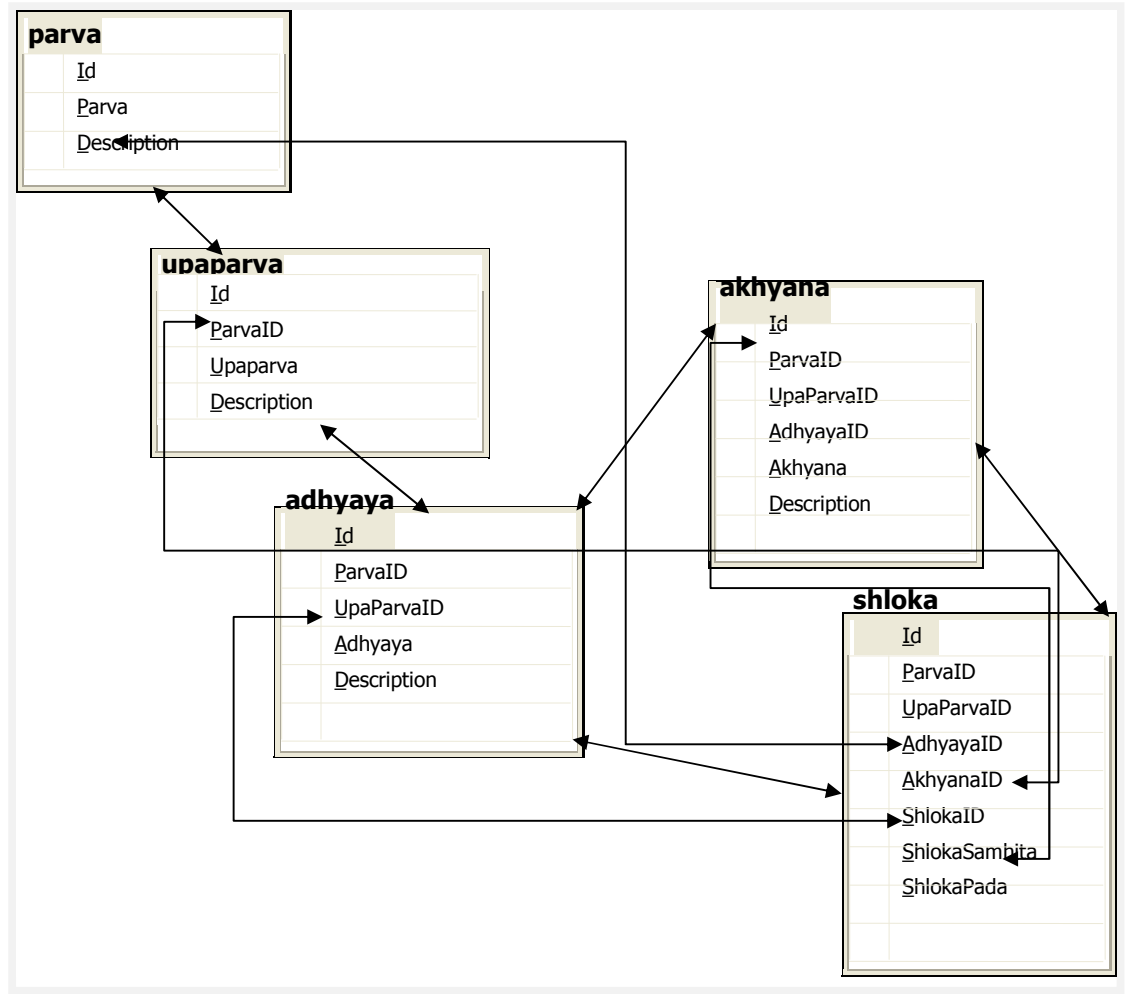
The Apache Tomcat is the servlet container that is used for the Java Servlet and Java Server Pages technologies. The Java Servlet and Java Server Pages specifications are developed by Sun under the Java Community Process. Apache Tomcat is developed in an open and participatory environment and released under the Apache Software License. Apache Tomcat is intended to be a collaboration of the best-of-breed developers from around the world.⁵ The *Mahābhārata* indexer runs on this Apache Tomcat 4.0 platform.

4.5. The Back-end

The back-end of the indexing system consists of RDBMS, which contains co-relative data tables. This Tomcat server based program connects to MS-SQL Server 2005 RDBMS through JDBC connectivity. The lexical resources are stored as Devanāgarī utf-8.

There are five tables namely; '*shloka*', '*adhyaya*', '*akhyana*', '*upaparva*' and '*parva*'. The descriptions of the tables have been given in the previous chapters. A design of the indexing system of *Mahābhārata* database is given below:

⁵ <http://www.apache.org/>



4.6. Database connectivity

The database connectivity has been done through the JDBC driver software. JDBC Application Programming Interface (API) is the industry standard for database independent connectivity for Java and a wide range of database-SQL databases. JDBC technology allows to use the Java programming language to develop ‘Write once, run anywhere’ capabilities for applications that require access to large scale data. JDBC works as bridge between Java program and Database. SQL server 2005 and JDBC support input and output in Unicode, so this system accepts Unicode Devanāgarī text as well as prints result in the same format.⁶

⁶ <http://java.sun.com/javase/technologies/database/> (accessed on 08/07/08)

4.7. How does the indexing system work?

This *Mahābhārata Indexer* has mainly online version which the user can access on <http://sanskrit.jnu.ac.in/mb/ibasic.jsp>. A CD version of the system has been enclosed herewith (in the back side of the dissertation). It is a testing version and does not have complete data and functionality. It is better to use the online version of the system.

4.7.1. How to use online version of the system

The system takes input and gives output in Devanāgarī UTF-8 encoding. For this, a Unicode input mechanism is required like Baraha,⁷ INSCRIPT key-board.⁸ Thus, one can type one's desired word for exact search. There are other facilities to search the queries like- one can choose a character from the Devanāgarī alphabet with the desired word begins. The search result will display a list of hyperlinked words with their references and the list of indexed words could be exact or partial string. The further information can be obtained by clicking the specific word. The drop-down boxes, according to the structure of the text, are the third searching facility where the user can first choose a "Parva" from the box. Entering the "Parva" the "Upa-parva" comes. In this chain, the third drop-down of "Adhyaya" appears. If the "Adhyaya" has any *ākhyāna* then the fourth drop-down of "Akhyana" will appear and the indexer will make the index according to the "Akhyana", if not then system will make an index of the words according to the specific *adhyāya*.

At second step, the user has to click a word among the list of indexed words. Clicking the required word the page will move to another page where he can find the searched query with its origin (i.e., in a *śloka*) and full reference on the basis of the *Mahābhārata (Ādiparva)*. On that same page, there is an option to search additional (i.e., linguistic and cultural) information from another online lexical resources. Clicking that *icon*, the user will get the links of several sites where he can find further information.

4.7.2. How to use CD version of the system

For running the *Mahābhārata Indexer*, the CD version requires installing Java Development Kit (JDK), Apache Tomcat web server 4.0 and the MS-SQL Server 2005 RDBMS based Database. It

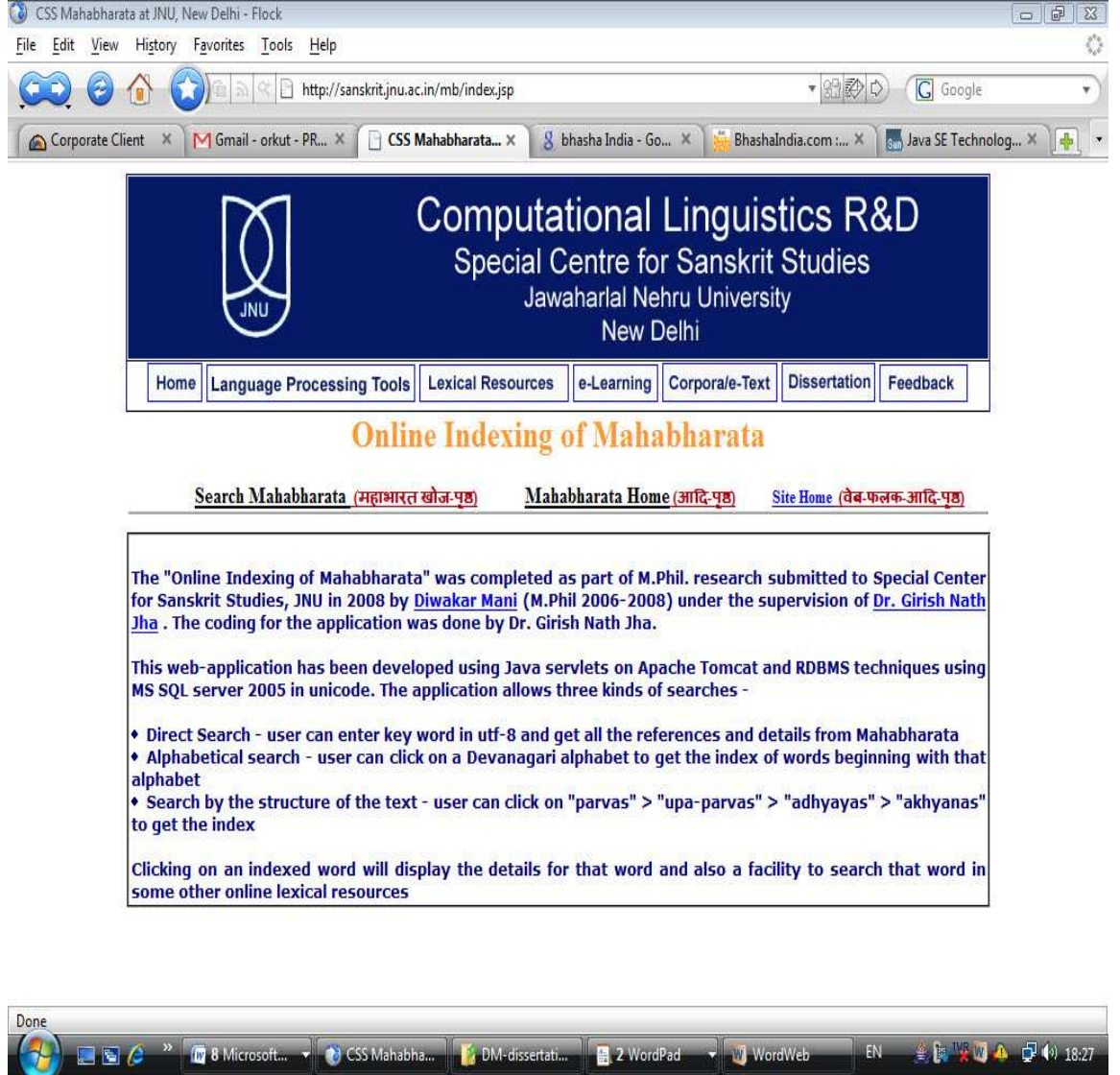
⁷ The Baraha can be freely downloaded at- www.baraha.com

⁸ It can be freely downloaded from- <http://www.bhashaindia.com/Downloadsv2/ListCategories.aspx>

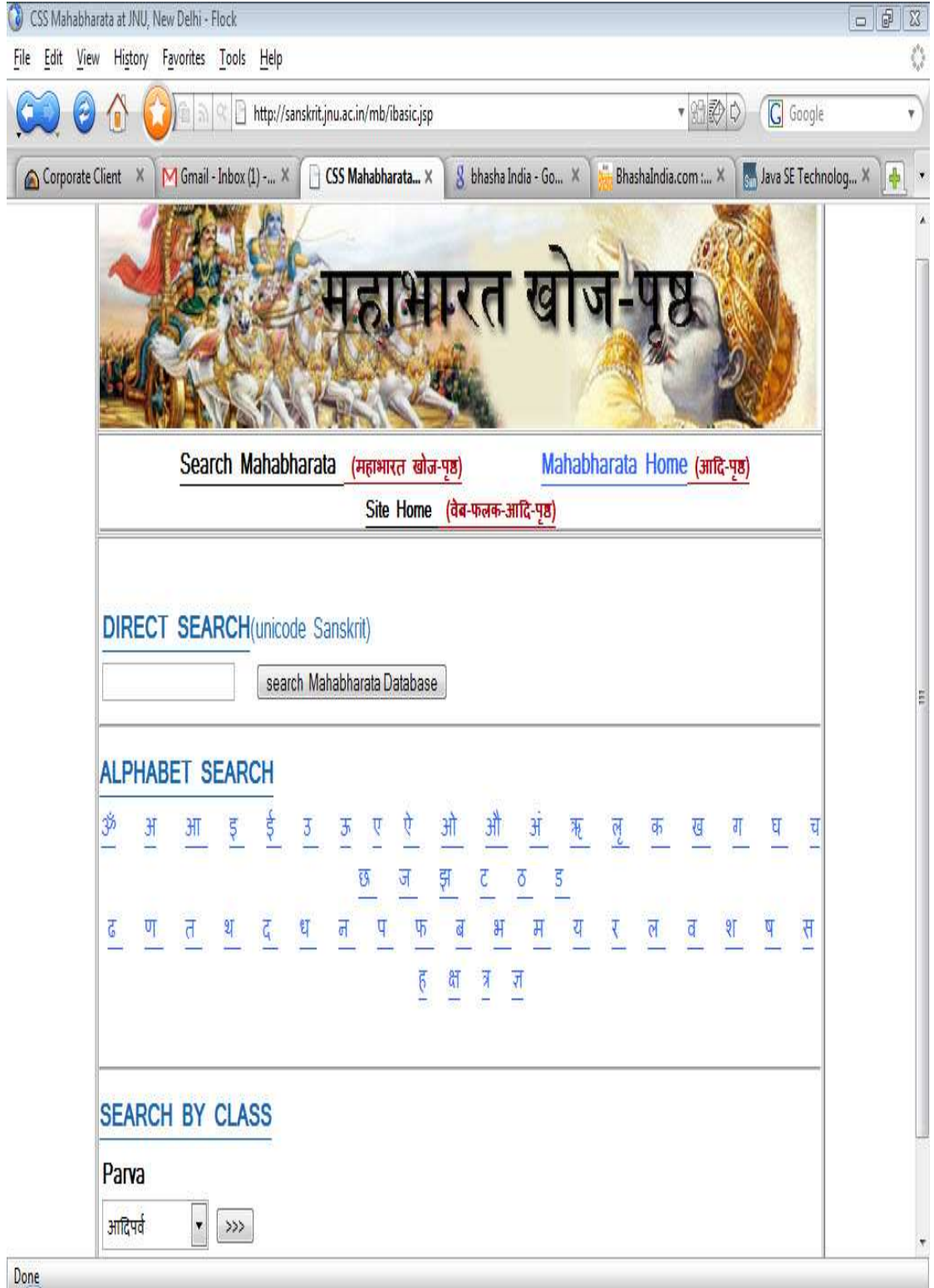
also needs a UTF-8 Devanāgarī input mechanism such as Baraha7.0. The installer files of these three are also given in the CD. The installation process for running the system is given here in respective order:

- ▲ First install the JDK of any updated version then set the path of 'java' by editing the path variable and adding the path of 'bin' folder of java (i.e., C:\jdk1.5.0_14\bin).
- ▲ Now install the Apache Tomcat web server. To run Tomcat Web Server two environment variables are to be defined-
 - JAVA_HOME= C:\jdk1.5.0_14
 - CATALINA_HOME= C:\Program Files\Apache Tomcat 4.0
- ▲ For installing the database the procedure should be:
 - First install MS-SQL server 2005 and MS-JDBC driver to connect the database with Apache Tomcat web server.
 - Run *Mahabharata* database script to create the database.
 - Import data from excel file. Set permissions to database objects.
- ▲ After installation of the database, copy the 'mb' folder from the CD to C:\Program Files\Apache Tomcat 4.0\webapps folder.
- ▲ Now start Tomcat Web Server and type this address in the web browser-
<http://localhost:8080/mb>


4.8. The Snapshots of the Mahābhārata Indexer



[Figure 4.4: Homepage of the Mahabharata Indexer]



[Figure 4.5: Search page of the Mahabharata indexer]



महाभारत खोज-पुस्त

[Search Mahabharata](#) (महाभारत खोज-पुस्त)
 [Mahabharata Home](#) (आदि-पुस्त)

[Site Home](#) (वेब-फलक-आदि-पुस्त)

DIRECT SEARCH(unicode Sanskrit)

Exact Search found 6 results for 'गरुड'

| | |
|------|---------------|
| गरुड | (1.5.24.0.5) |
| गरुड | (1.5.25.0.4) |
| गरुड | (1.5.27.0.35) |
| गरुड | (1.5.28.0.15) |
| गरुड | (1.5.30.0.1) |
| गरुड | (1.5.30.0.8) |

ALPHABET SEARCH

ॐ अ आ इ ई उ ऊ ए ऐ ओ औ अं ऋ लृ क ख ग घ च
छ ज झ ट ठ ड
ढ ण त थ द ध न प फ ब भ म य र ल व श ष स
ह क्ष त्र ज

Alphabet search found 31 results for 'गरुड'

| | |
|----------------|-----------------|
| गरुडस्य | (1.2.2.0.73) |
| गरुडः | (1.5.14.0.22) |
| गरुडः | (1.5.20.0.4) |
| गरुडः | (1.5.20.0.8) |
| गरुडम् | (1.5.20.0.9) |
| गरुडम् | (1.5.20.0.14) |
| गरुडः | (1.5.21.0.5) |
| गरुडः | (1.5.24.0.1) |
| गरुडः | (1.5.24.0.5) |
| गरुडम् | (1.5.25.0.3) |
| गरुडः | (1.5.25.0.4) |
| गरुडम् | (1.5.25.0.5) |
| गरुडैन | (1.5.26.0.1) |
| गरुडस्य | (1.5.26.0.13) |
| गरुडः | (1.5.26.0.20) |
| गरुडः | (1.5.26.0.26) |
| गरुडः | (1.5.27.0.1) |
| गरुडम् | (1.5.27.0.34) |
| गरुडः | (1.5.27.0.35) |
| गरुडः | (1.5.28.0.7) |
| गरुडः | (1.5.28.0.15) |
| गरुडः | (1.5.28.0.24) |
| गरुडः | (1.5.29.0.15) |
| गरुडः | (1.5.29.0.18) |
| गरुडः | (1.5.30.0.1) |
| गरुडः | (1.5.30.0.8) |
| गरुडारुणौ | (1.6.59.0.39) |
| गरुडो | (1.6.60.0.38) |
| गरुडारुणौ | (1.6.60.0.67) |
| गरुडपासितप्वजः | (1.7.114.0.62) |
| गरुडा | (1.19.218.0.20) |


SEARCH BY CLASS

Parva

आदिपर्व

>>>

[Figure 4.6: results of the searched query in hyperlinked mode with their numerical references]



महाभारत खोज-पृष्ठ

[Search Mahabharata \(महाभारत खोज-पृष्ठ\)](#)
[Mahabharata Home \(आदि-पृष्ठ\)](#)

[Site Home \(वेब-फलक-आदि-पृष्ठ\)](#)

DIRECT SEARCH(unicode Sanskrit)

ALPHABET SEARCH

ॐ अ आ इ ई उ ऊ ए ऐ ओ औ अं ऋ लृ क ख ग घ च

छ ज झ ट ठ ड

ढ ण त थ द ध न प फ ब भ म य र ल व श ष स

ह क्ष त्र ज्ञ

SEARCH BY CLASS

Parva

Results

Index Search for 'गरुड'

Reference: 1.5.24.0.5(आदिपर्व>आस्तीकपर्व>चतुर्विंशोऽध्यायः>5)

Shloka: गरुड उवाच- यथा अहम् अभिजानीयाम् ब्राह्मणम् लक्षणैः शुभैः । तत् मे कारणतः मातः पृच्छतः वक्तुम् अर्हसि ॥

Parva: आदिपर्व

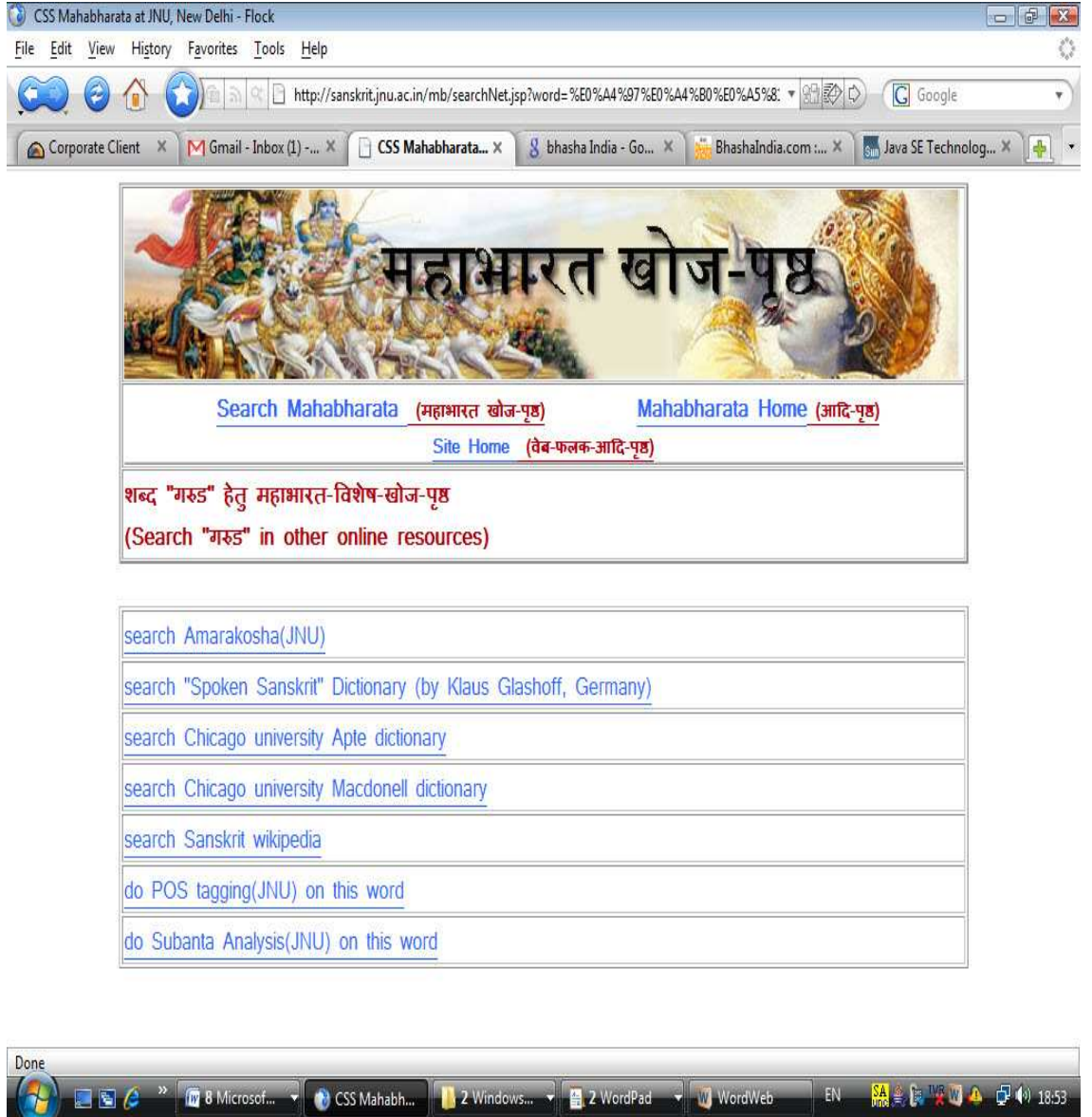
Upaparva: आस्तीकपर्व

Adhyaya: चतुर्विंशोऽध्यायः

Akhyana:

Search other sources:
[search this word on other online resources](#)

[Figure 4.7: referential page for the searched query]



[Figure 4.8: Cross-referential page where the user can click any one link to know further linguistic, grammatical and cultural knowledge]

अमरकोश-खोज-पृष्ठ (Search Page for Amarakosha)

| | | |
|--|---|-----------------------------------|
| Data Entry/Edit (डेटा-निवेश) | Search Amarakosha (डेटा-चिन्तन) | The team (अमरकोश) |
| मंडली | Home (आदि-पृष्ठ) | |

DIRECT SEARCH

(unicode Sanskrit/Hindi/English/Bangla/Punjabi/Oriya/Assamese/Maithili/Kannada languages)

गरुड

ALPHABET SEARCH

अ आ इ ई उ ऊ ए ऐ ओ औ क ख ग घ च छ ज
 झ ट ड त थ द ध न प फ ब भ म य र ल व श ष स ह
 क्ष त्र ज

SEARCH BY CLASS

स्वर्ग

| | | |
|--------------------|---|--------------------------|
| Search Results for | 'गरुड' | Multilingual gloss |
| Base Word: | गरुत्मान् | English: A heavenly bird |
| Reference: | | Hindi: गरुड |
| Semantic Category: | स्वर्ग | Kannada: |
| Category: | Noun | Bangla: गरुड |
| Number: | Singular | Oriya: ଗରୁଡ଼/ବିଷ୍ଣୁକ ବର |
| Gender: | Pum-linga | Punjabi: |
| Synonyms (if any): | पन्नगाशनः, सुपर्णः, विष्णुरथः, नागान्तकः, खगेश्वरः ,वैनतेयः, ताक्षर्यः, गरुडः(8) | Assamese: |
| | | Maithili: |

| | |
|-----------|---|
| Ontology: | स्वर्ग > गरुत्मान् > > गरुडः > ताक्षर्यः > वैनतेयः > खगेश्वरः > नागान्तकः > विष्णुरथः > सुपर्णः > पन्नगाशनः |
|-----------|---|

[Click here to edit this data](#)

[Figure 4.9: clicking on any site the page will move on that specific site to give further additional information (if the additional info is available). For example, clicking the 'Amarakosh site' it shows all the information of 'गरुड']

The screenshot shows a web browser window titled 'Sanskrit Dictionary for Spoken Sanskrit - Flock'. The address bar shows the URL 'http://www.spokensanskrit.de/index.php?script=HK&input=%20%E0%A4%97%E0%A4%'. The search bar contains the text 'गरुड' (Garuda). The 'Translate' button is visible. Below the search bar, there is a table of recent entries into the dictionary.

| Sanskrit word | Transliteration | Grammar | English word | Edit |
|---------------|-----------------|---------|--------------------|------|
| गरुड | garuDa | m. | eagle | Edit |
| म्लान | mLaana | adj. | feeble | Edit |
| तान्त { तम् } | taanta { tam } | ppp. | fainted away | Edit |
| शीर्ण | ziirNa | adj. | crushed | Edit |
| केत्वातु | ketvaatu | | cobalt [Chem.] | Edit |
| वेला | velaa | f. | coast [seashore] | Edit |
| शालिन् | zaalin | adj. | full of | Edit |
| रोपित | ropita | adj. | coated | Edit |
| शंसिन् | zaMsin | adj. | telling | Edit |
| अनुबोध | anubodha | m. | recollection | Edit |
| तान्त { तम् } | taanta { tam } | ppp. | fainted away | Edit |

[Figure 4.10: Here, the searched query gets additional informal from ‘Spoken Sanskrit Dictionary’ site]

Revised and enlarged edition of Prin. V. S. Apte's The practical Sanskrit-English Dictionary - Flock

File Edit View History Favorites Tools Help

http://dsal.uchicago.edu/cgi-bin/philologic/search3advanced?dbname=apte3new&qu... Google

विदेश-आलयम् खोजें डल पत्रिका लिंक चिट्ठे क्रिकेट 1:PM दबिटर [2] समाचार [50]

Corporate Client Gmail - Inbox (2) -... Revised and enla... file:///C:/...ions.html how can the text b... hwg-techniques a...

Digitals
Dictionaries
of South Asia

Vaman Shivaram Apte (1858-1892)
THE PRACTICAL SANSKRIT-ENGLISH DICTIONARY
Revised and Enlarged Edition of Prin. V. S. Apte's

Searching the **entire dictionary** for गरुड. Your search located 2 occurrences.

[Click here for a key-word-in-context display.](#)

1. विनता vinata : (page 1444)

a. with downcast face, dejected. विनता vinataविनता 1 N. of the mother of Aruṇa and Garuḍa, said to be one of the wives of Kaśyapa; see गरुड.-2 A kind of basket.-3 An abscess on the back or abdo- men.-Comp.-नन्दनः,-सुतः,-सूनुः epithets of Garuḍa or Aruṇa. विनतिः

2. सुधा sudha : (page 1693)

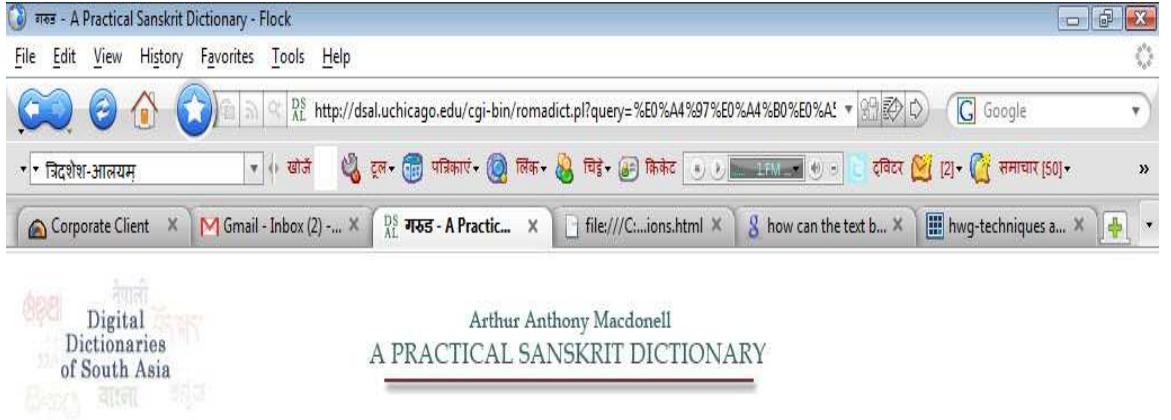
ambrosial, flowing with nectar; नेतुं वाञ्छति यः खलान् पथि सतां सूक्तैः सुधास्यन्दिभिः Bh.2.6.-सवा uvula or soft palate.-हरः,-हृत् an epithet of Garuḍa; see गरुड. सुधित sudhitaसुधित a. Nectar-like; गौप्यः स्फुरत्पुरटकुण्डलकुन्तलत्विङ्गण्ड-

[Back to the Search Page](#) | [Back to the DDSA Page](#)

Done

Revised and enlarge... 3 Microsoft Office... DM-dissertation-26j... EN 15:01

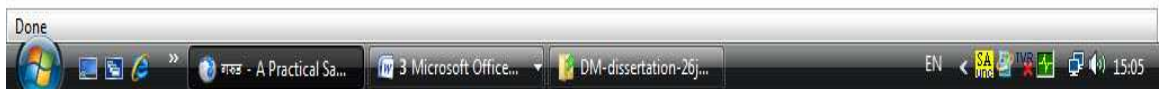
[Figure 4.11: the searched query 'गरुड' gets information from Online Apte's Sanskrit-English Dictionary]



A search found 1 entries with गरुड in the entry word or full text. The results are displayed using roman characters without diacritics and South Asian scripts.

garuda (p. 081) [garuḍa] m. N. of a fabulous bird, son of Vinatā and elder brother of Aruna, *king of the feathered race, vehicle of Vishnu or Krishna; kind of battle-array; -mānikya maya, a. of emerald; -vega, m. N. of a horse.*

[Back to the Search Page](#) / [Back to the DDSA Page](#)



[Figure 4.12: the searched query ‘गरुड’ gets information from Online Macdonell’s Sanskrit-English Dictionary]

~ 115 ~

Conclusion

The present work is an R&D effort at the M.Phil. level for developing an Online Indexing system of *Mahābhārata* (*Ādiparva*). It was a two year program: one year for course work and the next one year for R&D. Within this one year, a research on making database structure for the text was done. Also splitting the saṁhitā form of *ślokas* was done according to Pāṇinian rules. Besides this, the evaluation of tools and techniques- JSP-Java for front-end, servlet objects and Apache Tomcat for web server was studied and an online interface was developed which is live at <http://sanskrit.jnu.ac.in/mb>. The system has certain limitations which can be described as follows:

Limitations

- ▲ As the computation of the Sanskrit texts is opening a new perspective in the Sanskrit studies, the availability of the index of *Mbh* in an interactive format will be of great help to explore the *Mbh* worldwide. The absence of any such work increases the utility and importance of the work greatly. As the text is very large, this work is restricted to *Ādiparva* only, where 7,197 *ślokas* exists in 19 sub-parvans, 09 *ākhyānas* and, in general, a total of 225 adhyāyas.
- ▲ This system has fixed input and output mechanism. One can search his query in Unicode Devanāgarī only and the output will be in the same format.
- ▲ At present, the system is unable to give the translation in any language.
- ▲ In this version, it may fail to search a word which is in sandhi form. While sandhi-split version of text has been stored for only *Ādiparva*, efforts are being made to do this for all other parvans.
- ▲ If a base word is searched it cannot be found in all its forms. For example, if “*Brahman*” (*Prātipadika*) is searched, it will not return results for “*Brahma*”, “*Brhmā*”, “*Brahmaṇi*” etc. A subanta generation module is being developed separately to solve this problem.
- ▲ It has only a string search facility so it cannot search synonymous words. For example, if ‘*Kṛṣṇa*’ is searched, it cannot return ‘*Nārāyaṇa*’, ‘*Vāsudeva*’, ‘*Hari*’, ‘*Muralīdhara*’, ‘*Yogeśvara*’ etc. A separate module of *Amarakośa* has been developed. In near future, this module will be used to handle this issue.

Future research and development

The present *Online Indexing system for Ādiparva of Mahābhārata* has tremendous potentials in the field of Sanskrit Computational Lexicography and M(A)TS. Some of the immediate and future applications of the system are discussed below:

- ▲ **An Online Indexer for Sanskrit documents-** the present work only deals with the *Ādiparvan* of *Mahābhārata* but the same methodology can be applied to build an indexing system for remaining parvans of the text and also for another Sanskrit documents which are based on this epic.
- ▲ **Support for other encoding schemes-** at this point, the system takes the input in UTF-8 format but in the future, it may be upgraded to process the input in other encoding schemes also.
- ▲ **Machine Translation System (MTS)-** This R&D has potential for M(A)TS from and to Sanskrit which is a major goal of this and other R&D currently in progress at the SCSS , JNU.

Appendices

APPENDIX-I**List of the Parvans of Mahābhārata**

| Id | Parva |
|----|-------------------|
| 1 | आदिपर्व |
| 2 | सभापर्व |
| 3 | वनपर्व (आरण्यक) |
| 4 | विराटपर्व |
| 5 | उद्योगपर्व |
| 6 | भीष्मपर्व |
| 7 | द्रोणपर्व |
| 8 | कर्णपर्व |
| 9 | शल्यपर्व |
| 10 | सौप्तिकपर्व |
| 11 | स्त्रीपर्व |
| 12 | शान्तिपर्व |
| 13 | अनुशासनपर्व |
| 14 | अश्वमेधपर्व |
| 15 | आश्रमवासिकपर्व |
| 16 | मौसलपर्व |
| 17 | महाप्रस्थानिकपर्व |
| 18 | स्वर्गारोहणिकपर्व |
| 19 | हरिवंशपर्व (खिल) |

APPENDIX-II**List of the Upa-parvans**

| Id | Parva_Id | Upaparva |
|----|----------|-----------------|
| 1 | 1 | अनुक्रमणीपर्व |
| 2 | 1 | पर्वसंग्रहपर्व |
| 3 | 1 | पौष्यपर्व |
| 4 | 1 | पौलोमपर्व |
| 5 | 1 | आस्तीकपर्व |
| 6 | 1 | आदिवंशावतरणपर्व |
| 7 | 1 | संभवपर्व |

| | | |
|----|---|---------------------------|
| 8 | 1 | जनुगृहदाहपर्व |
| 9 | 1 | हिडिम्बवधपर्व |
| 10 | 1 | बकवधपर्व |
| 11 | 1 | चैत्ररथपर्व |
| 12 | 1 | द्रौपदीस्वयंवरपर्व |
| 13 | 1 | वैवाहिकपर्व |
| 14 | 1 | विदुरागमनपर्व |
| 15 | 1 | राज्यलम्भपर्व |
| 16 | 1 | अर्जुनवनवासपर्व |
| 17 | 1 | सुभद्राहरणपर्व |
| 18 | 1 | हरणहारिकपर्व |
| 19 | 1 | खाण्डवदाहपर्व |
| 20 | 2 | सभावर्णनपर्व |
| 21 | 2 | मन्त्रपर्व |
| 22 | 2 | जरासंधवधपर्व |
| 23 | 2 | दिग्विजयपर्व |
| 24 | 2 | राजसूयिकपर्व |
| 25 | 2 | अर्घाभिहरणपर्व |
| 26 | 2 | शिशुपालवधपर्व |
| 27 | 2 | धूतपर्व |
| 28 | 2 | अनुधूतपर्व |
| 29 | 3 | आरण्यकपर्व |
| 30 | 3 | किर्मीरवधपर्व |
| 31 | 3 | कैरातपर्व |
| 32 | 3 | इन्द्रलोकाभिगमनपर्व |
| 33 | 3 | तीर्थयात्रापर्व |
| 34 | 3 | जटासुरवधपर्व |
| 35 | 3 | यक्षयुद्धपर्व |
| 36 | 3 | आजगरपर्व |
| 37 | 3 | मार्कण्डेयसमस्यापर्व |
| 38 | 3 | द्रौपदी-सत्यभामासंवादपर्व |
| 39 | 3 | घोषयात्रापर्व |
| 40 | 3 | मृगस्वप्नभयपर्व |
| 41 | 3 | व्रीहीद्रौणिकपर्व |

| | | |
|----|---|-----------------------|
| 42 | 3 | द्रौपदीहरणपर्व |
| 43 | 3 | कुन्दलाहरणपर्व |
| 44 | 3 | आरण्यपर्व |
| 45 | 4 | वैराटपर्व |
| 46 | 4 | कीचकवधपर्व |
| 47 | 4 | गोग्रहणपर्व |
| 48 | 4 | वैवाहिकपर्व |
| 49 | 5 | उद्योगपर्व |
| 50 | 5 | सञ्जययानपर्व |
| 51 | 5 | प्रजागरपर्व |
| 52 | 5 | सनत्सुजातापर्व |
| 53 | 5 | यानसंधिपर्व |
| 54 | 5 | भगवद्यानपर्व |
| 55 | 5 | विवादपर्व |
| 56 | 5 | निर्याणपर्व |
| 57 | 5 | रथातिरथसंख्यापर्व |
| 58 | 5 | उलूकदूतागमनपर्व |
| 59 | 5 | अम्बोपाख्यानपर्व |
| 60 | 6 | भीष्माभिषेचनपर्व |
| 61 | 6 | जम्बुखण्डनिर्माणपर्व |
| 62 | 6 | भूमिपर्व |
| 63 | 6 | भगवद्गीतापर्व |
| 64 | 6 | भीष्मवधपर्व |
| 65 | 7 | द्रोणाभिषेकपर्व |
| 66 | 7 | संशप्तकवधपर्व |
| 67 | 7 | अभिमन्युवधपर्व |
| 68 | 7 | प्रतिज्ञापर्व |
| 69 | 7 | जयद्रथवधपर्व |
| 70 | 7 | घटोत्कचवधपर्व |
| 71 | 7 | द्रोणवधपर्व |
| 72 | 7 | नारायणास्त्रमोक्षपर्व |
| 73 | 8 | कर्णपर्व |
| 74 | 9 | शल्यपर्व |
| 75 | 9 | हृदप्रवेशपर्व |

| | | |
|-----|----|----------------------|
| 76 | 9 | गदायुद्धपर्व |
| 77 | 9 | सारस्वतपर्व |
| 78 | 10 | सौप्तिकपर्व |
| 79 | 10 | ऐषीकपर्व |
| 80 | 10 | जलप्रदानिकपर्व |
| 81 | 11 | स्त्रीपर्व |
| 82 | 11 | श्रद्धापर्व |
| 83 | 11 | अभिषेचनिकपर्व |
| 84 | 11 | चार्वीकनिग्रहपर्व |
| 85 | 11 | गृहप्रविभागपर्व |
| 86 | 12 | राजधर्मपर्व |
| 87 | 12 | आपद्धर्मपर्व |
| 88 | 12 | मोक्षधर्मपर्व |
| 89 | 13 | अनुशासनिकपर्व |
| 90 | 13 | भीष्मस्वर्गारोहणपर्व |
| 91 | 14 | अश्वमेधिकपर्व |
| 92 | 14 | अनुगीतापर्व |
| 93 | 15 | आश्रमवासपर्व |
| 94 | 15 | पुत्रदर्शनपर्व |
| 95 | 15 | नारदागमनपर्व |
| 96 | 16 | मौसलपर्व |
| 97 | 17 | महाप्रस्थानिकपर्व |
| 98 | 18 | स्वर्गारोहणपर्व |
| 99 | 19 | हरिवंशपर्व |
| 100 | 19 | भविष्यत्पर्व |

APPENDIX-III

List of Ākhyanas of the Ādiparva

| Id | Akhyana |
|----|-----------------|
| 1 | शकुन्तलोपाख्यान |
| 2 | ययात्युपाख्यान |
| 3 | उत्तरयायातम् |
| 4 | तापत्योपाख्यान |

| | |
|---|----------------------|
| 5 | वाशिष्ठोपाख्यान |
| 6 | और्वोपाख्यान |
| 7 | पञ्चेन्द्रोपाख्यान |
| 8 | सुन्दोपसुन्दोपाख्यान |
| 9 | शार्ङ्गकोपाख्यान |

APPENDIX-IV

List of Adhyāyas of the Ādiparva

| Id | UP_Id | Akh_Id | Adh_Nm |
|----|-------|--------|-------------------|
| 1 | 1 | 0 | प्रथमोऽध्यायः |
| 2 | 2 | 0 | द्वितीयोऽध्यायः |
| 3 | 3 | 0 | तृतीयोऽध्यायः |
| 4 | 4 | 0 | चतुर्थोऽध्यायः |
| 5 | 4 | 0 | पञ्चमोऽध्यायः |
| 6 | 4 | 0 | षष्ठोऽध्यायः |
| 7 | 4 | 0 | सप्तमोऽध्यायः |
| 8 | 4 | 0 | अष्टमोऽध्यायः |
| 9 | 4 | 0 | नवमोऽध्यायः |
| 10 | 4 | 0 | दशमोऽध्यायः |
| 11 | 4 | 0 | एकादशोऽध्यायः |
| 12 | 4 | 0 | द्वादशोऽध्यायः |
| 13 | 5 | 0 | त्रयोदशोऽध्यायः |
| 14 | 5 | 0 | चतुर्दशोऽध्यायः |
| 15 | 5 | 0 | पञ्चदशोऽध्यायः |
| 16 | 5 | 0 | षोडशोऽध्यायः |
| 17 | 5 | 0 | सप्तदशोऽध्यायः |
| 18 | 5 | 0 | अष्टादशोऽध्यायः |
| 19 | 5 | 0 | एकोनविंशोऽध्यायः |
| 20 | 5 | 0 | विंशोऽध्यायः |
| 21 | 5 | 0 | एकविंशोऽध्यायः |
| 22 | 5 | 0 | द्वाविंशोऽध्यायः |
| 23 | 5 | 0 | त्रयोविंशोऽध्यायः |
| 24 | 5 | 0 | चतुर्विंशोऽध्यायः |

| | | | |
|----|---|---|------------------------|
| 25 | 5 | 0 | पञ्चविंशोऽध्यायः |
| 26 | 5 | 0 | षड्विंशोऽध्यायः |
| 27 | 5 | 0 | सप्तविंशोऽध्यायः |
| 28 | 5 | 0 | अष्टविंशोऽध्यायः |
| 29 | 5 | 0 | एकोनविंशोऽध्यायः |
| 30 | 5 | 0 | त्रिंशोऽध्यायः |
| 31 | 5 | 0 | एकत्रिंशोऽध्यायः |
| 32 | 5 | 0 | द्वात्रिंशोऽध्यायः |
| 33 | 5 | 0 | त्रयस्त्रिंशोऽध्यायः |
| 34 | 5 | 0 | चतुस्त्रिंशोऽध्यायः |
| 35 | 5 | 0 | पञ्चत्रिंशोऽध्यायः |
| 36 | 5 | 0 | षट्त्रिंशोऽध्यायः |
| 37 | 5 | 0 | सप्तत्रिंशोऽध्यायः |
| 38 | 5 | 0 | अष्टत्रिंशोऽध्यायः |
| 39 | 5 | 0 | एकोनचत्वारिंशोऽध्यायः |
| 40 | 5 | 0 | चत्वारिंशोऽध्यायः |
| 41 | 5 | 0 | एकचत्वारिंशोऽध्यायः |
| 42 | 5 | 0 | द्विचत्वारिंशोऽध्यायः |
| 43 | 5 | 0 | त्रिचत्वारिंशोऽध्यायः |
| 44 | 5 | 0 | चतुश्चत्वारिंशोऽध्यायः |
| 45 | 5 | 0 | पञ्चचत्वारिंशोऽध्यायः |
| 46 | 5 | 0 | षट्चत्वारिंशोऽध्यायः |
| 47 | 5 | 0 | सप्तचत्वारिंशोऽध्यायः |
| 48 | 5 | 0 | अष्टचत्वारिंशोऽध्यायः |
| 49 | 5 | 0 | एकोनपञ्चाशत्तमोऽध्यायः |
| 50 | 5 | 0 | पञ्चाशत्तमोऽध्यायः |
| 51 | 5 | 0 | एकपञ्चाशत्तमोऽध्यायः |
| 52 | 5 | 0 | द्विपञ्चाशत्तमोऽध्यायः |
| 53 | 5 | 0 | त्रिपञ्चाशत्तमोऽध्यायः |
| 54 | 6 | 0 | चतुःपञ्चाशत्तमोऽध्यायः |
| 55 | 6 | 0 | पञ्चपञ्चाशत्तमोऽध्यायः |
| 56 | 6 | 0 | षट्पञ्चाशत्तमोऽध्यायः |
| 57 | 6 | 0 | सप्तपञ्चाशत्तमोऽध्यायः |
| 58 | 6 | 0 | अष्टपञ्चाशत्तमोऽध्यायः |

| | | | |
|----|---|---|----------------------|
| 59 | 6 | 0 | एकोनषष्ठितमोऽध्यायः |
| 60 | 6 | 0 | षष्ठितमोऽध्यायः |
| 61 | 6 | 0 | एकषष्ठितमोऽध्यायः |
| 62 | 7 | 1 | द्विषष्ठितमोऽध्यायः |
| 63 | 7 | 1 | त्रिषष्ठितमोऽध्यायः |
| 64 | 7 | 1 | चतुःषष्ठितमोऽध्यायः |
| 65 | 7 | 1 | पञ्चषष्ठितमोऽध्यायः |
| 66 | 7 | 1 | षट्षष्ठितमोऽध्यायः |
| 67 | 7 | 1 | सप्तषष्ठितमोऽध्यायः |
| 68 | 7 | 1 | अष्टषष्ठितमोऽध्यायः |
| 69 | 7 | 1 | एकोनसप्ततितमोऽध्यायः |
| 70 | 7 | 2 | सप्ततितमोऽध्यायः |
| 71 | 7 | 2 | एकसप्ततितमोऽध्यायः |
| 72 | 7 | 2 | द्विसप्ततितमोऽध्यायः |
| 73 | 7 | 2 | त्रिसप्ततितमोऽध्यायः |
| 74 | 7 | 2 | चतुःसप्ततितमोऽध्यायः |

APPENDIX-V

List of Ślokasamhitā of the Ādiparva

| Shloka_Id | Adhyaya_Id | Shloka_Nm | ShlokaSamhita |
|-----------|------------|-----------|--|
| 1 | 1 | 0 | नारायणं नमस्कृत्य नरं चैव नरोत्तमम् । देवीं सरस्वतीं चैव ततो जयमुदीरयेत् ॥ |
| 2 | 1 | 1 | लोमहर्षणपुत्र उग्रश्रवाः सूतःपौराणिको । नैमिषारण्ये शौनकस्य कुलपतेर्द्वादशवार्षिके सत्रे । |
| 3 | 1 | 2 | समासीनानभ्यगच्छद्ब्रह्मर्षीन् संशितव्रतान् । विनयावनतो भूत्वा कदाचित् सूतनन्दनः ॥ |
| 4 | 1 | 3 | तमाश्रममनुप्राप्तं नैमिषारण्यवासिनः । चित्राः श्रोतुं कथास्तत्र परिववृस्तपस्विनः ॥ |
| 5 | 1 | 4 | अभिवाद्य मुनींस्तांस्तु सर्वानेव कृताञ्जलि । अपृच्छत्स तपोवृद्धिं सद्भिश्चैवाभिनन्दितः ॥ |
| 6 | 1 | 5 | अथ तेषूपविष्टेषु सर्वेष्वेव तपस्विषु । निर्दिष्टमासनं भेजे विनयाल्लोमहर्षणिः ॥ |
| 7 | 1 | 6 | सुखासीनं ततस्तंतु विश्रान्तमुपलक्ष्य च । अथापृच्छदृषिस्तत्र |

| | | | |
|----|---|----|---|
| | | | कश्चित्प्रस्तावयन्कथाः ॥ |
| 8 | 1 | 7 | कुत आगम्यते सौते क्व चायं विहृतस्त्वया । कालः कमलपत्राक्ष शंसैतत्पृच्छतो मम ॥ |
| 9 | 1 | 8 | सूत उवाच - जनमेजयस्य राजर्षेः सर्पसत्रे महात्मनः । समीपे पार्थिवेन्द्रस्य सम्यक्पारिक्षितस्य च ॥ |
| 10 | 1 | 9 | कृष्णद्वैपायनप्रोक्ताः सुपुण्या विविधाः कथाः । कथिताश्चापि विधिवद्या वैशंपायनेन वै ॥ |
| 11 | 1 | 10 | श्रुत्वाहं ता विचित्रार्था महाभारतसंश्रिताः । बहूनि संपरिक्रम्य तीर्थान्यायतनानि च ॥ |
| 12 | 1 | 11 | समन्तपञ्चकं नाम पुण्यं द्विजनिषेवितम् । गतवानस्मि तं देशं युद्धं यत्राभवत्पुरा । पाण्डवानां कुरूणां च सर्वेषां च महीक्षिताम् ॥ |
| 13 | 1 | 12 | दिदक्षुरागतस्तस्मात्समीपं भवतामिह । आयुष्मन्तः सर्व एव ब्रह्मभूता हि मे मताः ॥ |
| 14 | 1 | 13 | अस्मिन्यज्ञे महाभागाः सूर्यपावकवर्चस । कृताभिषेकाः शुचयः कृतजप्या हुताग्नयः । भवन्त आसते स्वस्था ब्रवीमि किमहं द्विजाः ॥ |
| 15 | 1 | 14 | पुराणसंश्रिताः पुण्याः कथा वा धर्मसंश्रिताः । इतिवृत्तं नरेन्द्राणामृषीणां च महात्मनाम् ॥ |
| 16 | 1 | 15 | ऋषय ऊचुः - द्वैपायनेन यत्प्रोक्तं पुराणं परमर्षिणा । सुरैर्ब्रह्मर्षिभिश्चैव श्रुत्वा यदभिपूजितम् ॥ |
| 17 | 1 | 16 | तस्याख्यानवरिष्ठस्य विचित्रपदपर्वणः । सूक्ष्मार्थन्याययुक्तस्य वेदार्थैर्भूषितस्य च ॥ |
| 18 | 1 | 17 | भारतस्येतिहासस्य पुण्यां ग्रन्थार्थसंयुताम् । संस्कारोपगतां ब्राह्मीं नानाशास्त्रोपबृंहिताम् ॥ |
| 19 | 1 | 18 | जनमेजयस्य यां राज्ञो वैशंपायन उक्तवान् । यथावत्स ऋषिस्तुष्ट्या सत्रे द्वैपायनाज्ञया ॥ |
| 20 | 1 | 19 | वेदैश्चतुर्भिः समितां व्यासस्याद्भुतकर्मणः । संहितां श्रोतुमिच्छामो धर्म्या पापभयापहाम् ॥ |
| 21 | 1 | 20 | सूत उवाच - आद्यं पुरुषमीशानं पुरुहूतं पुरुष्टुतम् । ऋतमेकाक्षरं ब्रह्म व्यक्ताव्यक्तं सनातनम् ॥ |
| 22 | 1 | 21 | असच्च सच्चैव च यद्विश्वं सदसतः परम् । परावराणां स्रष्टारं पुराणं परमव्ययम् ॥ |
| 23 | 1 | 22 | मङ्गल्यं मङ्गलं विष्णुं वरेण्यमनघं शुचिम् । नमस्कृत्य हृषीकेशं चराचरगुरुं हरिम् ॥ |
| 24 | 1 | 23 | महर्षेः पूजितस्येह सर्वलोके महात्मनः । प्रवक्ष्यामि मतं कृत्स्नं व्यासस्यामिततेजसः ॥ |
| 25 | 1 | 24 | आचख्युः कवयः केचित्संप्रत्याचक्षते परे । आख्यास्यन्ति तथैवान्ये |

| | | | |
|----|---|----|---|
| | | | इतिहासमिमं भुवि ॥ |
| 26 | 1 | 25 | इदं तु त्रिषु लोकेषु महज्ज्ञानं प्रतिष्ठितम् । विस्तरैश्च समासैश्च धार्यते यदिद्वजातिभिः ॥ |
| 27 | 1 | 26 | अलंकृतं शुभैः शब्दैः समयैर्दिव्यमानुषैः । छन्दोवृत्तैश्च विविधैरन्वितं विदुषां प्रियम् ॥ |
| 28 | 1 | 27 | निष्प्रभेऽस्मिन्निरालोके सर्वतस्तमसावृते । बृहदण्डमभूदेकं प्रजानां बीजमक्षयम् ॥ |
| 29 | 1 | 28 | युगस्यादौ निमित्तं तन्महद्वित्यं प्रचक्षते । यस्मिंस्तच्छ्रूयते सत्यं ज्योतिर्ब्रह्म सनातनम् ॥ |
| 30 | 1 | 29 | अद्भुतं चाप्यचिन्त्यं च सर्वत्र समतां गतम् । अव्यक्तं कारणं सूक्ष्मं यत्तत्सदसदात्मकम् ॥ |
| 31 | 1 | 30 | यस्मात्पितामहो जज्ञे प्रभुरेकः प्रजापतिः । ब्रह्मा सुरगुरुः स्थाणुर्मनुः कः परमेष्ठ्यथ ॥ |
| 32 | 1 | 31 | प्राचेतसस्तथा दक्षो दक्षपुत्राश्च सप्त ये । ततः प्रजानां पतयः प्राभवन्नेकविंशतिः ॥ |
| 33 | 1 | 32 | पुरुषश्चाप्रमेयात्मा यं सर्वमृषयो विदुः । विश्वेदेवास्तथादित्या वसवोऽथाश्विनावपि ॥ |
| 34 | 1 | 33 | यक्षाः साध्याः पिशाचाश्च गुह्यकाः पितरस्तथा । ततः प्रसूता विद्वांसः शिष्टा ब्रह्मर्षयोऽमलाः ॥ |
| 35 | 1 | 34 | राजर्षयश्च बहवः सर्वेः समुदिता गुणैः । आपो द्यौः पृथिवी वायुरन्तरिक्षं दिशस्तथा ॥ |
| 36 | 1 | 35 | संवत्सरर्तवो मासाः पक्षाहोरात्रयः क्रमात् । यच्चान्यदपि तत्सर्वं संभूतं लोकसाक्षिकम् ॥ |
| 37 | 1 | 36 | यदिदं दृश्यते किञ्चिद्भूतं स्थावरजङ्गमम् । पुनः संक्षिप्यते सर्वं जगत्प्राप्ते युगक्षये ॥ |
| 38 | 1 | 37 | यथार्तावृतुलिङ्गानि नानारूपाणि पर्यये । दृश्यन्ते तानि तान्येव तथा भावा युगादिषु ॥ |
| 39 | 1 | 38 | एवमेतदनाद्यन्तं भूतसंहारकारकम् । अनादिनिधनं लोके चक्रं संपरिवर्तते ॥ |
| 40 | 1 | 39 | त्रयस्त्रिंशत्सहस्राणि त्रयस्त्रिंशच्छतानि च । त्रयस्त्रिंशच्च देवानां सृष्टिः संक्षेपलक्षणा ॥ |
| 41 | 1 | 40 | दिवस्पुत्रो बृहद्भानुश्चक्षुरात्मा विभावसुः । सविता च ऋचीकोऽर्को भानुराशावहो रविः ॥ |
| 42 | 1 | 41 | पुत्रा विवस्वतः सर्वे मह्यस्तेषां तथावरः । देवभ्रातृ तनयस्तस्य तस्मात्सुभ्राडिति स्मृतः ॥ |
| 43 | 1 | 42 | सुभ्राजस्तु त्रयः पुत्राः प्रजावन्तो बहुश्रुताः । दशज्योतिः शतज्योतिः सहस्रज्योतिरात्मवान् ॥ |

| | | | |
|----|---|----|---|
| 44 | 1 | 43 | दश पुत्रसहस्राणि दशज्योतेर्महात्मनः । ततो दशगुणाश्चान्ये शतज्योतेरिहात्मजाः ॥ |
| 45 | 1 | 44 | भूयस्ततो दशगुणाः सहस्रज्योतिषः सुताः । तेभ्योऽयं कुरुवंशश्च यदूनां भरतस्य च ॥ |
| 46 | 1 | 45 | ययातीक्ष्वाकुवंशश्च राजर्षीणां च सर्वशः । संभूता बहवो वंशा भूतसर्गाः सविस्तराः ॥ |
| 47 | 1 | 46 | भूतस्थानानि सर्वाणि रहस्यं विविधं च यत् । वेदयोगं सविज्ञानं धर्मोऽर्थ काम एव च ॥ |
| 48 | 1 | 47 | धर्मकामार्थशास्त्राणि शास्त्राणि विविधानि च । लोकयात्राविधानं च संभूतं दृष्टवानृषिः ॥ |
| 49 | 1 | 48 | इतिहासाः सवैयाख्या विविधाः श्रुतयोऽपि च । इह सर्वमनुक्रान्तमुक्तं ग्रन्थस्य लक्षणम् ॥ |
| 50 | 1 | 49 | विस्तीर्यैतन्महज्ज्ञानमृषिः संक्षेपमब्रवीत् । इष्टं हि विदुषां लोके समासव्यासधारणम् ॥ |
| 51 | 1 | 50 | मन्वादि भारतं केचिदास्तीकादि तथापरे । तथोपरिचराद्यन्ये विप्राः सम्यगधीयते ॥ |
| 52 | 1 | 51 | विविधं संहिताज्ञानं दीपयन्ति मनीषिणः । व्याख्यातुं कुशलाः केचिद्ग्रन्थं धारयितुं परे ॥ |

APPENDIX-VI

List of Ślokapada (Split form) of the Ādiparva

| Shloka_Id | Adhyaya_Id | Shloka_Nm | ShlokaPada |
|-----------|------------|-----------|---|
| 1 | 1 | 0 | नारायणम् नमः कृत्य नरम् च एव नर उत्तमम् । देवीम् सरस्वतीम् च एव ततः जयम् उदीरयेत् ॥ |
| 2 | 1 | 1 | लोमहर्षण पुत्र उग्रश्रवाः सूतः पौराणिकः । नैमिष अरण्ये शौनकस्य कुलपतेः द्वादश वार्षिके सत्रे ॥ |
| 3 | 1 | 2 | समासीनान् अभ्यगच्छद् ब्रह्मर्षीन् संशितव्रतान् । विनयं अवनतः भूत्वा कदाचित् सूतनन्दनः ॥ |
| 4 | 1 | 3 | तम् आश्रमम् अनुप्राप्तम् नैमिष अरण्यवासिनः । चित्राः श्रोतुम् कथाः तत्र परिवव्रुः तपस्विनः ॥ |
| 5 | 1 | 4 | अभिवाद्य मुनीन् तान् तु सर्वान् एव कृत अञ्जलिः । अपृच्छत् सः |

| | | | |
|----|---|----|---|
| | | | तपोवृद्धिम् सद्भिः च एव अभिनन्दितः ॥ |
| 6 | 1 | 5 | अथ तेषु उपविष्टेषु सर्वेषु एव तपस्विषु । निर्दिष्टम् आसनम् भजे विनयात् लोमहर्षणिः ॥ |
| 7 | 1 | 6 | सुख आसीनम् ततः तम् तु विश्रान्तम् उपलक्ष्य च । अथ अपृच्छत् ऋषिः तत्र कः चित् प्रस्तावयन् कथाः ॥ |
| 8 | 1 | 7 | कुत आगम्यते सौते क्व च अयम् विहृतः त्वया । कालः कमलपत्र अक्ष शंस एतत् पृच्छतः मम ॥ |
| 9 | 1 | 8 | सूत उवाच - जनमेजयस्य राज ऋषेः सर्पसत्रे महात्मनः । समीपे पार्थिव इन्द्रस्य सम्यक् पारिक्षितस्य च ॥ |
| 10 | 1 | 9 | कृष्णद्वैपायन प्र उक्ताः सुपुण्या विविधाः कथाः । कथिताः च अपि विधिवद्या वैशंपायनेन वै ॥ |
| 11 | 1 | 10 | श्रुत्वा अहम् ता विचित्र अर्था महाभारत संश्रिताः । बहूनि संपरिक्रम्य तीर्थानि आयतनानि च ॥ |
| 12 | 1 | 11 | समन्तपञ्चकम् नाम पुण्यम् द्विजनिषेवितम् । गतवान् अस्मि तम् देशम् युद्धम् यत्र अभवत् पुरा । पाण्डवानाम् कुरूणाम् च सर्वेषाम् च महीक्षिताम् ॥ |
| 13 | 1 | 12 | दिदक्षुः आगतः तस्मात् समीपम् भवताम् इह । आयुष्मन्तः सर्व एव ब्रह्मभूता हि मे मताः ॥ |
| 14 | 1 | 13 | अस्मिन् यज्ञे महाभागाः सूर्य पावक वर्चसः । कृत अभिषेकाः शुचयः कृत जप्या हुत अग्नयः । भवन्तः आसते स्वस्था ब्रवीमि किम् अहम् द्विजाः ॥ |
| 15 | 1 | 14 | पुराण संश्रिताः पुण्याः कथा वा धर्म संश्रिताः । इतिवृत्तम् नरेन्द्राणाम् ऋषीणाम् च महात्मनाम् ॥ |
| 16 | 1 | 15 | ऋषय ऊचुः - द्वैपायनेन यत् प्र उक्तम् पुराम् परम ऋषिणा । सुरैः ब्रह्म ऋषिभिः च एव श्रुत्वा यत् अभिपूजितम् ॥ |
| 17 | 1 | 16 | तस्य आख्यान वरिष्ठस्य विचित्र पद पर्वणः । सूक्ष्म अर्थ न्याय |

| | | | |
|----|---|----|---|
| | | | युक्तस्य वेद अर्थः भूषितस्य च ॥ |
| 18 | 1 | 17 | भारतस्य इतिहासस्य पुण्याम् ग्रन्थ अर्थ संयुताम् । संस्कार उपगताम् ब्राह्मी नाना शास्त्र उपबृंहिताम् ॥ |
| 19 | 1 | 18 | जनमेजयस्य याम् राज्ञः वैशंपायन उक्तवान् । यथावत् सः ऋषिः तुष्ट्या सत्रे द्वैपायन आजया ॥ |
| 20 | 1 | 19 | वेदैः चतुर्भिः समिताम् व्यासस्य अद्भुत कर्मणः । संहिताम् श्रोतुम् इच्छामः धर्म्याम् पाप भय अपहाम् ॥ |
| 21 | 1 | 20 | सूत उवाच - आद्यम् पुरुषम् ईशानम् पुरुहूतम् पुरुष्टुतम् । ऋतम् एक अक्षरम् ब्रह्म व्यक्त अव्यक्तम् सनातनम् ॥ |
| 22 | 1 | 21 | असत् च सत् च एव च यत् विश्वम् सत् असतः परम् । पर अवराणाम् स्रष्टाम् पुराणम् परम् अव्ययम् ॥ |
| 23 | 1 | 22 | मङ्गल्यम् मङ्गलम् विष्णुम् वरेण्यम् अनघम् शुचिम् । नमः कृत्य हृषीकेशम् चर अचर गुरुम् हरिम् ॥ |
| 24 | 1 | 23 | महा ऋषेः पूजितस्य इह सर्वलोके महात्मनः । प्रवक्ष्यामि मतम् कृत्स्नम् व्यासस्य अमित तेजसः ॥ |
| 25 | 1 | 24 | आचख्युः कवयः केचित् संप्रति आचक्षते परे । आख्यास्यन्ति तथा एव अन्ये इतिहासम् इमम् भुवि ॥ |
| 26 | 1 | 25 | इदम् तु त्रिषु लोकेषु महत् ज्ञानम् प्रतिष्ठितम् । विस्तरैः च समासैः च धार्यते यत् द्विजातिभिः ॥ |
| 27 | 1 | 26 | अलंकृतम् शुभैः शब्दैः समयैः दिव्य मानुषैः । छन्दः वृत्तैः च विविधैः अन्वितम् विदुषाम् प्रियम् ॥ |
| 28 | 1 | 27 | निष्प्रभे अस्मिन् निरालोके सर्वतः तमः आवृत्ते । बृहदण्डम् अभूत् एकम् प्रजानाम् बीजम् अक्षयम् ॥ |
| 29 | 1 | 28 | युगस्य आदौ निमित्तम् तत् महत् दिव्यम् प्रचक्षते । यस्मिन् तत् श्रूयते सत्य ज्योतिः ब्रह्म सनातनम् ॥ |

| | | | |
|----|---|----|---|
| 30 | 1 | 29 | अद्भुतम् च अपि अचिन्त्यम् च सर्वत्र समताम् गतम् । अव्यक्तम् कारणम् सूक्ष्मम् यत् तत् सद् असदात्मकम् ॥ |
| 31 | 1 | 30 | यस्मात् पितामहः जज्ञे प्रभुः एकः प्रजापतिः । ब्रह्मा सुरगुरुः स्थाणुः मनुः कः परमेष्ठि अथ ॥ |
| 32 | 1 | 31 | प्राचेतसः तथा दक्षः दक्षपुत्राः च सप्त ये । ततः प्रजानाम् पतयः प्राभवन् एक विंशतिः ॥ |
| 33 | 1 | 32 | पुरुषः च अप्रमेय आत्मा यम् सर्वम् ऋषयः विदुः । विश्वेदेवाः तथा आदित्या वसवः अथ अश्विनौ अपि ॥ |
| 34 | 1 | 33 | यक्षाः साध्याः पिशाचाः च गुह्यकाः पितरः तथा । ततः प्रसूता विद्वांसः शिष्टा ब्रह्मर्षयः अमलाः ॥ |
| 35 | 1 | 34 | राजर्षयः च बहवः सर्वेः समुदिता गुणैः । आपः द्यौः पृथिवी वायुः अन्तरिक्षम् दिशः तथा ॥ |
| 36 | 1 | 35 | संवत्सर ऋतवः मासाः पक्ष अहोरात्रयः क्रमात् । यत् च अन्यत् अपि तत् सर्वम् संभूतम् लोकसाक्षिकम् ॥ |
| 37 | 1 | 36 | यत् इदम् दृश्यते किञ्चिद् भूतम् स्थावर जङ्गमम् । पुनः संक्षिप्यते सर्वम् जगत् प्राप्ते युग क्षये ॥ |
| 38 | 1 | 37 | यथा ऋतौ ऋतु लिङ्गानि नाना रूपाणि पर्यये । दृश्यन्ते तानि तानि एव तथा भावा युगादिषु ॥ |
| 39 | 1 | 38 | एवमेतद् अनादि अन्तम् भूत संहार कारकम् । अनादि निधनम् लोके चक्रम् संपरिवर्तते ॥ |
| 40 | 1 | 39 | त्रयः त्रिंशत् सहस्राणि त्रयः त्रिंशत् शतानि च । त्रयः त्रिंशत् च देवानाम् सृष्टिः संक्षेप लक्षणा ॥ |
| 41 | 1 | 40 | दिवः पुत्रः बृहद् भानुः चक्षुः आत्मा विभावसुः । सविता च ऋचीकः अर्कः भानुः आशावहो रविः ॥ |
| 42 | 1 | 41 | पुत्रा विवस्वतः सर्वे मह्यस्तेषाम् तथावरः । देवभाट् तनयः तस्य तस्मात् सुभाट् इति स्मृतः ॥ |

| | | | |
|----|---|----|---|
| 43 | 1 | 42 | सुभ्राजः तु त्रयः पुत्राः प्रजायन्तः बहुश्रुताः । दशज्योतिः शतज्योतिः सहस्रज्योतिः आत्मवान् ॥ |
| 44 | 1 | 43 | दश पुत्रसहस्राणि दशज्योतेः महात्मनः । ततः दशगुणाः च अन्ये शतज्योतेः इह आत्मजाः ॥ |
| 45 | 1 | 44 | भूयः ततः दशगुणाः सहस्रज्योतिषः सुताः । तेभ्यः अयम् कुरुवंशः च यदूनाम् भरतस्य च ॥ |
| 46 | 1 | 45 | ययाति ईक्ष्वाकुवंशः च राजर्षीणाम् च सर्वशः । संभूता बहवः वंशा भूतसर्गाः सविस्तराः ॥ |
| 47 | 1 | 46 | भूतस्थानानि सर्वाणि रहस्यम् विविधम् च यत् । वेदयोगम् सविज्ञानम् धर्मः अर्थं काम एव च ॥ |
| 48 | 1 | 47 | धर्मं कामं अर्थशास्त्राणि शास्त्राणि विविधानि च । लोक यात्रा विधानम् च संभूतम् दृष्टवान् ऋषिः ॥ |
| 49 | 1 | 48 | इतिहासाः सवैयाख्या विविधाः श्रुतयः अपि च । इह सर्वम् अनुक्रान्तम् उक्तम् ग्रन्थस्य लक्षणम् ॥ |
| 50 | 1 | 49 | विस्तीर्यै तत् महत् ज्ञानम् ऋषिः संक्षेपम् अब्रवीत् । इष्टम् हि विदुषाम् लोके समास व्यास धारणम् ॥ |
| 51 | 1 | 50 | मनु आदि भारतम् केचिद् आस्तीक आदि तथा अपरे । तथा उपरिचरात् अन्ये विप्राः सम्यक् अधीयते ॥ |
| 52 | 1 | 51 | विविधम् संहिता ज्ञानम् दीपयन्ति मनीषिणः । व्याख्यातुम् कुशलाः केचित् ग्रन्थम् धारयितुम् परे ॥ |

Bibliography

Books

- Allen, James, 2004, **Natural Language Understanding**, Pearson Education Pvt. Ltd. (Singapore), New Delhi.
- Bakharia, Aneesha. 2001, **Java Server Pages**, Prentice Hall of India Pvt. Ltd, New Delhi.
- Barbara J. Grosz et al. (ed.), 1990, **Readings in Natural Language Processing**, Morgan Kaufmann: Los Altos, CA.
- Belvalkar, S.K. (Ed.), 1947, **The Mahābhārata (Bhīṣmaparvan)**, Bhandarkar Oriental Research Institute, Poona.
- Bharati, Akshar, Chaitanya, Vineet & Sangal, Rajeev, **A computational framework for Indian languages**, Technical Report TRCS-90-100, Dept. of CSE, IIT Kanpur, July 1990b. (Course Notes for Intensive Course on NLP for Linguists, Vol.1)
- Bharati, Akshar, Sangal, Rajeev & Chaitanya, Vineet, 2004, **Natural language Processing: A Paninian Perspective**, Prentice Hall of India Pvt. Ltd, New Delhi.
- Bharati, Akshar, Sharma, Dipti M., Chaitanya, Vineet, Kulkarni, Amba P., Sangal, Rajeev & Rao, Durgesh D., **LERIL: Collaborative Effort for Creating Lexical Resources**, In Proc. of Workshop on Language Resources in Asian Languages, together with 6th NLP Pacific Rim Symposium, Tokyo.
- Bhattacharya, J.N. & Sarkar, Nilanjana (eds.), 2004, **Encyclopedic Dictionary of Sanskrit (Vol-III)**, Global Vision Publishing House, Delhi.
- Boguraev, Bran & Ted Briscoe (eds.), 1989, **Computational Lexicography for Natural Language Processing**, Longman, London.
- Borko, Harold, and Bernier, Charles L., 1978. **Indexing Concepts and Methods**, Academic Press INC., New York.
- Briggs, Rick ,1985, **Sanskrit & Artificial Intelligence-NASA ,Knowledge Representation in Sanskrit and Artificial Intelligence** (Roacs, NASA Ames Research Center, Moffet Field, California)
- Brockington , John , 1998, **A Comprehensive Guide to the Mahabharata (and the Ramayana) and Scholarship on them**, Brill, Lei den.
- Bucknell, Roderic S., 1994, **Sanskrit Manual**, Motilal Banarsidas, Delhi.
- Chattopadhyay, Santanu, 2005, **Compiler Design**, Prentice Hall of India, New Delhi.

- Coward, David F. & Charles E. Grimes, 1995, **Making Dictionaries: A guide to lexicography and the Multi-Dictionary Formatter**, Summer Institute of Linguistics, Waxhaw, North Carolina.
- Dādhimatha, Śivadattaśāstrī, *Nirukta* with Durgācārya commentary, Venkateshwara Press, Mumbai.
- Date, C.J., 1987, **Introduction to Database Systems**, Addison-Wesley, Reading, MA.
- Dowson, John (1820-1881), 1979, **A classical dictionary of Hindu mythology and religion, geography, history, and literature**, Routledge, London.
- Dvivedi, Balmukund, 1979, **Sanskrit Dictionary: Origin and Development**, Smriti Publishers, Allahabad.
- Edgerton, Franklin (Ed.), 1944, **The Mahābhārata (Sabhāparvan)**, Bhandarkar Oriental Research Institute, Poona.
- Gairola, Vachaspati, 1997, *Samśkr̥ta Sāhitya kā Itihāsa*, Chaukhambha Vidya Bhavan, Varanasi.
- Garnham, Allan, 1988, **Artificial Intelligence: An Introduction**, Routledge and Kegan Paul, London, NY.
- Grishman, R., 1986, **Computational Linguistics: An Introduction**, Cambridge University Press, New York.
- Grosz Barbara J., Jones Karen Sparck & Webber Bonnie Lynn, 1986, **Readings in Natural Language Processing**, Morgan Kaufmann Publishers, Inc, California.
- Handke, Jürgen, 1995, **The Structure of the Lexicon: Human versus Machine**, Mouton de Gruyter, Berlin.
- Hopkins, E. Washburn, 1898, an article on **The Bharata and the Great Bharata** (from The American Journal of Philology, Vol. 19, No. 1, pp. 1-24), The Johns Hopkins University Press, Baltimore, Maryland, USA.
- Jamie Joworski, **Java 2 Platform unleashed**, BPB Publication, Techmedia, Ansari Road, Daryaganj, New Delhi.
- Jha, Girish N, 2005, **Information technology applications for Sanskrit lexicography: case of Amarakosha**, procs of the AsiaLex conference at National University of Singapore, Singapore, June 1-3, 2005.

- Jha, Girish N, 2005, **Language Technology in India: A survey**, Computer Society of India, November 2005.
- Jurafsky, Daniel & Martin, James H., 2005, **Speech and Language Processing**, Pearson Education (Singapore) Pvt. Ltd., New Delhi.
- Kanetkar, Yashavant P., 1996, **Unix Shell Programming**, BPB Publications, New Delhi.
- McArthur, Tom & Kernerman, Ilan (Ed.), 1998, **Lexicography in Asia** (selected papers from the Dictionaries in Asia Conference, Hong Kong University of Science and Technology, 1997), Password Publishers Ltd, Tel Aviv, Israel.
- Mitkov, Raslan (Ed.), 2003, **The Oxford Handbook of Computational Linguistics**, Oxford University Press, Oxford.
- Narang, Satyapal, 1998, **Sam̐skṛta Koṣa-Śāstra ke Vividha Āyāma**, Rashtriya Sanskrit Sansthan, New Delhi.
- Naur, Peter, Randell, Brian & Buxton, J. N. (Ed.), 1976, **Software engineering: concepts and techniques**, Mason/Charter Publisher Inc, New York.
- Navrat P. & Ueno H. (eds.), 1998, **Knowledge Based Software Engineering**, IOS Press, Amsterdam, Netherlands.
- Patkar, M.M., 1984, **Descriptive Catalogue of Manuscripts (Vol-III)**, Shri Ranbir Sanskrit Research Institute, Jammu.
- Patkar, Madhukar M., 1981. **History of Sanskrit Lexicography**, Munshiram Manoharlal Publishers Pvt. Ltd., New Delhi.
- Rajan M.A.S. & Srinivasan S.H., 1993, **Sanskrit and Computer-based Linguistics**, Academy of Sanskrit Research Melkote, Karnataka.
- Ringland, G.A. & Duce, D.A., 1988, **Approaches to Knowledge Representation: An Introduction**, Research Studies Press Ltd, Letchworth, Hertfordshire, England.
- Russell, Joseph P., 2002, **Java Programming**, Prentice Hall of India Pvt. Ltd, New Delhi.
- Sangal, Rajeev & Bendre, S.M. (eds.), 2003, **Recent Advances in Natural Language Processing** (Proceedings of the International Conference ICON-2002), Vikas Publishing House Pvt. Ltd., New Delhi.

- Sharma, Indu, Sharma, Ranavir & Sharma, Shrikrishna (Ed.), 2002, ***Mahābhāratapadānukramakoṣe (Ādiparvakhaṇḍa)***, Sanskrit Ewam Prācya Vidyāsamsthāna, Kurukshetra University, Kurukshetra.
- Sharma, Ishvar Chandra & Bimali, O.N. (Eds.), 2006, **Mahābhārata (Sanskrit Text with English Translation of M.N. Dutt, Volume-I)**, Parimal Publications, Delhi.
- Sörensen S. (Ed.), 1963, **An Index to the Names in the Mahābhārata**, Motilal Banarasidas, Delhi.
- Sukthankar, V.S. (Ed.), 1933, **The Mahābhārata (Ādiparvan)**, Bhandarkar Oriental Research Institute, Poona.
- Sukthankar, V.S. (Ed.), 1942, **The Mahābhārata (Āraṇyakaparvan-I)**, Bhandarkar Oriental Research Institute, Poona.
- Sukthankar, V.S. (Late), 1944, **Critical Studies In The Mahābhārata**, V.S. Sukthankar Memorial Edition Committee, Poona.
- Winternitz, Maurice, 1987, **A History of Indian Literature**, Vol-I, Motilal Banarasidas, Delhi.
- Yogi, Satyabhushan & Shashikumar, 1985, ***Nighaṇṭu tathā Nirukta***, Motilal Banarasidas, Delhi.

Thesis & Dissertation

- Chandra, Subash. 2006, **Machine Recognition and Morphological Analysis of Subanta-padas**, M.Phil dissertation submitted to SCSS, JNU.
- Chandrashekar R., 2007, **Part-of-Speech Tagging for Sanskrit**, Ph.D. thesis submitted to SCSS, JNU.
- Chaudhary, Narayan K, 2006, **Developing a Computational Framework for the Verb Morphology of Great Andamanese**, M.Phil dissertation submitted to CL/SLL&CS, JNU.
- Kumar, Sachin, 2007, **Sandhi Splitter and Analyzer for Sanskrit (with special reference to aC sandhi)**, M.Phil dissertation submitted to SCSS, JNU.

e-Resources

- Apache Tomcat Servlet, <http://tomcat.apache.org/> (accessed: 07/05/2008)
- Back Words Indexing, <http://www.backwordsindexing.com/Comp.html> (accessed: 18/05/2008).
- Baraha, Software, <http://www.baraha.com/BarahaIME.htm> (accessed: 12/05/2008)
- Boguraev, Bran and Briscoe, Ted, Large Lexicons for NLP, <http://acl.ldc.upenn.edu/J/J87/J87-3002.pdf> (accessed: 05/06/2008).
- Contents and structure of the Mahābhārata, <http://www.spiritus-temporis.com/mahabharata/the-mahabharata,-the-epic-story.html> (accessed: 14/05/2008).
- Developments in MTS 93, <http://www.hutchinsweb.me.uk/MTS-1993.pdf> (accessed: 09/06/2008).
- Electronic Database, <http://www.jstor.org/stable/pdfplus/1694720.pdf> (accessed: 15/06/2008).
- Index (Database), http://en.wikipedia.org/wiki/Index_%28database%29 (accessed: 10/06/2008).
- Java Server Pages, <http://java.sun.com/products/jsp/> (accessed: 02/06/2008).
- JAVA Servlet, <http://java.sun.com/products/servlet/> (accessed: 07/06/2008).
- LingBench IDE, http://www.natlantech.com/lingbench_ide.html (accessed: 11/06/2008).
- Multimedia indexing and searching environment, <http://www.dcs.shef.ac.uk/nlp/mumis/> (accessed: 11/06/2008).
- OntoLex 2004: Ontologies and Lexical Resources in Distributed Environments,ch-1, <http://www.loa-cnr.it/ontolex2004.html> (accessed: 14/06/2008).
- Practical NLP lexicons, <http://www.ilc.cnr.it/EAGLES96/synlex/node17.html> (accessed: 17/06/2008).
- Review of Lexical Resources, <http://delivery.acm.org/10.1145/980000/976170/p113-sampson.pdf?key1=976170&key2=6832772121&coll=GUIDE&dl=ACM&CFID=31102481&CFTOKEN=78941241> (accessed: 17/06/2008).
- Structure of the Mahābhārata, <http://www.experiencefestival.com/mahabharata> (accessed: 14/05/2008).
- Structure of the Mahābhārata, <http://www.indopedia.org/Mahabharata.html> (accessed: 13/05/2008).

- Mahabharata Resources, <http://mahabharata-resources.org/> (accessed on: 15/07/2008).
- Mahabharata Samshodhana Pratishthana,
http://mahabharata.org/index.php?current_tab=index (accessed on: 18/06/2008).
- Mahabharata, <http://www.smso.net/Mahabharata> (accessed on: 19/06/2008).
- Goettingen's Mahabharata, http://www.sub.uni-goettingen.de/ebene_1/fiindolo/gretil/1_sanskrit/2_epic/mbh/sas/mahabharata.htm (accessed on: 16/06/2008).