# LEXICOGRAPHY: FROM EARLIEST TIMES TO THE PRESENT

## SAMPATH BABU TOKALA

**Abstract:** According to Trench (1858), a dictionary is an inventory of the words of a language along with details of meaning and other relevant information. Communities of the world's major and minor languages have recognized the need for the dictionary. Motivation,  for compiling a dictionary comes from various needs of the language, viz. language teaching, developing writers tools, language technology, natural language processing etc.  Initially, collecting words, defining them, and arranging them in certain order would not seem to require any profound theoretical basis. However, when the activity demands for the building of a sophisticated and exhaustive work, involving theoretical and applied linguistic aspects then a number of issues begin to crowd in as in the following:

**Introduction:** With regards to the relationship between words and phrases, the traditional dictionaries have tried to present distinction between words, idioms, 'fixed phrases'. Recent advances, both in construction grammar (e.g. Goldberg, 1995, 2006) and in corpus linguistics (e.g. Sinclair 1991, 1998; Hanks, 2004, 2013), suggest that meaning resides not only in lexical items but also in phraseology.  Other question, often lexicographers juggled with is question of recording the millions of attested nominal phrases, such as *fire escape* and *forest fire*, each of which has at least one unique meaning, which is very often not derivable from the sum of its parts. Different languages dealt with such compounds in different ways. Further, questions regarding the limitation of its inventory to recorded usage. Questions such as possible words and meanings in addition to recorded words and meanings.

It has been claimed that the earliest dictionaries in the world were compiled in China (Li Ming 2006, Yong and Peng 2008). Erya, dating from the 2nd or 3rd century BC, is usually classified as a work of encyclopedic lexicography: it contains explanations of the meanings of words, phrases, and other passages in classic Chinese texts. According to Karen Chung (personal communication), the Erya falls somewhere between a thesaurus and a topically organized lexicon. During  (ca. 2 AD), Xu Shen compiled the Shouwenjiezi, etymology of Chinese words. This is the beginning of all subsequent Chinese lexicography It contains around10,000 entries in Chinese characters, with information about their origins, meanings, and pronunciation. It is organized in 540 sections according to the 'radicals' of each word.

India and Persia Sanskrit dictionaries and thesauruses were compiled over two thousand years ago, and these were the start of a long tradition of native lexicography in Indian languages. Three terms are particularly relevant to the Indian lexicographic tradition: nighantu, kosha, and nirukta. Nighantu simply means 'lexicon'. The earliest known nighantu gives explanations of obscure words found in Vedic texts (sacred literature). In the second or third century BC, a scholar called Yaska, about whom nothing else is known, wrote an etymological commentary (nirukta 'explanation') on words found in a lexicon (nighantu). A kosha is literally a storehouse or treasury. Unusually for lexicography, the earliest kosha was written in verse. It contains entries for nouns and indeclinable forms, but not verbs, and was intended for use by poets. The best-known such work is the Amarakosha by Amarasinha, a Buddhist scholar and poet who probably lived in the sixth century AD.

Amarakośa is the ancient and the most popular lexicon compiled by Amarasiṁha, the renowned scholar, a well-known poet, and grammarian lived around 5th century. It discusses the usefulness and derivations of a word. Scholars ascribed the immortality of nāmaliṅgānuśāsana is due to its faultless work. Essentially,  it is a dictionary of synonymous words but devotes a section called nānārthavarga dealing with homonymous words too. The latter words  are arranged according to the final syllable. It contains a total of 11, 580 words. These words are the words that are used in the synonymous sets of words which define the semantic content of the words. In the beginning of his work, he brings up some special rules, a sub set of meta rules to interpret the gender information of a word. Amarakośa is divided into three kāṇḍas. Each kāṇḍa is further sub-divided into "vargas". The first kāṇḍa has words pertaining to gods, heaven and the five basic elements (pañcamahābhūta) and abstract concepts such as direction (dik), time (kāla), the speech (vāk), etc. This chapter has ten "vargas". The second kāṇḍa deals with the words denoting real physical objects such as earth, human beings, animals, plants etc. This chapter also has ten "vargas". The third kāṇḍa has five "vargas" where words dealing with the description of the grammatical miscellany, polysemous words and other miscellaneous words are dealt with as in the following:

**Chapters (khaṇḍas):**

I.          **First Chapter (prathamakāṇḍa), Vargas from each kāṇḍa are named thus:**

1. Svargavargaḥ (heaven): Heaven, Gods, Demons, their arms, ornaments, symbols or vehicles, and other attributes, Fire, Air, Velocity, Eternity, etc.

2. Vyomavargaḥ (sky): The Sky and its related concepts.

3. Digvargaḥ (direction): Directions, Deities of the directions, elephants at the points, their female elephants, Cloud, thunder, lightning, rainbow, Rain, hail, rainy day, cloudy day, Moon, types of light, frost, Stars, Planets, sunset, dawn, sunlight, etc.

4. Kālavargaḥ (time): Time, day, night, variations of the moon, eclipse, second, hour, months, year, Weather, seasons, Happy, Sorrow, Soul, Mind, etc.

5. Dhīvargaḥ (cognition): Individuality, consciousness, knowledge, sense, organs, tastes, fragrance, colours, etc.

6. Śabdādivargaḥ (sound): Sarasvatī, voice, word, Vedas, Vedāṅgas, stories, legends, sound, types of sounds, speech, musical sounds, song, ornament's sound, etc.

7. Nāṭyavargaḥ (dance and drama): Seven musical tones, Musical Instruments, dance, theatrical characters, sentiments, desire, affection, kindness, Festival etc.

8. Pātālabhogivargaḥ (nether world): Infernal region, hole, darkness, Snakes, kinds of serpent, parts of snake, etc

9. Narakavargaḥ (hell): Hell, various hells, departed souls, pain, etc.

10. Vārivargaḥ (water): Water, Ocean, wave, whirlpool, shore, channel, island, boat, voyage, pilot, deep, fish, fisherman, net, fish basket, hook, etc., types of fishes, Aquatic animals, crab, turtle, etc. Well, pond, types of ponds, River, Names of rivers, water plants, lotus, water lilly, etc. Parts of these plants etc.

## II.      The Second Chapter (dvitīyakāṇḍa) :

1. Bhūmivargaḥ (earth):Earth, land, soil, clay, world, India, regions, types of lands, country, village, kingdom, hill, road;

2. Puravargaḥ (towns or Cities): City, suburb, Market, fort, wall, house, kinds of houses, parts of house, house holdings, building land, etc.

3. Śailavargaḥ (hills and mountains): Mountains, kinds of mountains, parts of mountains, caves, etc.

4. Vanauṣadhivargaḥ (forests and medicines): Forest, garden, tree, parts of tree, flowers, fruits, leaf, shrub, creeper, names of trees, names of shrubs, names of creepers, names of grass, etc

5. Siṁhādivargaḥ (lions and other animals): Animals, lion, tiger, wolf, deer, kinds of deers, etc. Insects, bee, cricket, birds, hawk, skylark, crow, parrot, etc. parts of birds, wing, beak, etc.

6. Manuṣyavargaḥ (mankind): Man, woman, descriptions of woman, blood-relations like son, daughter, husband, wife etc., manhood, different stages of manhood, parts of our body, diseases such as cough, scab, etc., dress, ornaments, cloths, types of cloths, cosmetics, fragrant plants, sandal etc., hair decoration styles, daily usable things etc.

7. Brahmavargaḥ (priest section):Tribes, religious states, sacerdotal, scholars, characters and descriptions of priests, their occupations and observances, types of fires, sacrifice, its requisites, alms, worship, austerity, study, hypocrisy, marriage, human, pursuits and objects etc.

8. Kṣatriyavargaḥ (military section): Military tribe, kings, ministers, officers, servants, enemies, allies, requisites of government, means of defence, and of success, revenue, foresight, insignia of royalty, army, elephants, parts and kind of elephants, horses, types of horses, vehicles, chariots, litters, warriors, arms and weapons, bow, arrow, javelin etc. war, slaughter, funeral, prison, life, etc.

9. Vaiśyavargaḥ (business section):Third tribe, professions, husbandman, field, implements of husbandry, corn, pulse, oil-seeds, granary, kitchen, vessels, prepared food, dairy, cattle, traffic, weights and measures, commodities, etc.

10. Śūdravargaḥ (Others): The Fourth tribe, the mixed classes, artisans, jugglers, dancers, musician, hunters, servants, barbarians, dogs, hogs, theft, nets, fops, loom, plot for burden, wrought leather, tools, art, images, wages, spirituous gaming, etc

## III.      Third Chapter (tṛtīyakāṇḍa):

1. Viśeṣyanighnavargaḥ (adjective): Epithets of persons, Qualities of things, etc.

2. Saṁkīrṇavargaḥ (miscellaneous): Miscellaneous words;

3. Nānārthavargaḥ (polysemous): Homonymous and polysemous words;

4. Avyayavargaḥ (indeclinables): Indeclinables

5. Liṅgādisaṅgrahavargaḥ (gender): The three genders, and variations of gender: They are Masculine, Feminine, Neuter, Masculine and Feminine, Masculine and Neuter, and Feminine and Neuter.

The lexicon, Amarakośa is of great importance and is seriously studied, used and one of the most extensively translated text into various Indian and foreign languages. Scores of commentaries are written and published. It has influenced lexicographic traditions, both in the east and the west.

During the early middle age there has been a lot of cultural exchange between India and neighbouring Persia. There is some evidence that Persian dictionaries existed before the modern times, particularly mention must be made of Loghat-e-Fors (Lexicon of Persian), compiled by Abu Mansur Ali ibn Ahmad AsadiTusi (c. 1072). The main aim was not only to record and explain words found in Persian literature and that might be unfamiliar to his contemporaries. The entries are illustrated with

citations from literature and are arranged according to the alphabetical order, unlike the Amarakośa. However, another Persian lexicon compiled by Faxr-e-QavasQaznavi in India in 1291, the Farhang-e-PanjBaxši 'culture in five sections'.   The Persian lexicographical tradition flourished in India, where many Persian writers lived and practiced. In the last millennium there are about 130 Persian dictionaries were compiled.

The history of Greek lexicography accounts both the efforts of Greek lexicographers in classical times and of the lexicography of ancient Greek since the Renaissance. From the fifth century BC onwards, Greek scribes used glosses to read and understand manuscripts of earlier scholars to explain obsolete and unusual words.   Later these glosses were compiled into separate glossaries.   In the second century AD, a     dictionary was compiled to distinguish 'correct' words and terms from the colloquialisms. Eklogē (Ἐκλογὴ 'selection') by Phrynichos of Bithynia.  Ancient Latin lexicographers are comparatively less known; works of classical Latin lexicography have been hardly available.   An ambitious monolingual dictionary called De Verborum Significatu 'on the meaning of words' was compiled by the philologist and educationist Marcus VerriusFlaccus (c. 55 BC – 20 AD). It was a huge work and was concerned with etymology and cultural history as well as word meaning. Entries were supported by citations from literature. In the 2nd century AD Sextus Pompeius Festus edited a revised version of this work, partly survived.  In the eighth century, the historian Paulus Diaconus created an abridged version of Flaccus's dictionary, and this has survived. It is said to be a cultural work rather than a linguistic compendium.

Arabic and Hebrew dictionaries were compiled between the 7th and the 13th centuries AD, with a variety of functions, including the prescriptive use of language, the facilitation of the use of the words of the Qur'an.  As noted in Roth (1994), both Arabic and Hebrew lexicography flourished in medieval Spain, before they were all swept away by the Christian 'suzerainty'.   Latin and Greek enjoyed a universally accepted conventional alphabetical order, but the history of the conventional order of letters in the Arabic writing system is ridden with problems. The fact that, in written Arabic, normally only consonants are represented, while readers are left to supply the vowels for themselves.   Arab lexicographers experimented with various ways of ordering of words. In the Lexicon, Al-Jawhariyy's As-sihah 'the Strong' (ca. 10 A.D.) and Ibn Manzur's monumental Lisan Al-'Arab 'Language of the Arabs' (ca. 13 A.D.) words are ordered according to the last consonant and arrange the consonants in an order that is determined to some extent by the mode of articulation (as in Indian languages). Modern Arabic dictionaries invented a different alphabetical order. The most important work of medieval Arabic lexicography is the Kitab Al-'Ayn (literally, Book of the 'ayn), compiled by Al-Khalil ibn Ahmad in the 8th century. An 'ayn is a written symbol representing the Arabic voiced pharyngeal fricative consonant /ʕ/ to indicate a comprehensive dictionary of the Arabic language. The earliest known works of Hebrew lexicography were compiled in the Middle East in the 10th century AD (see Drory 2000; Cohen and Choueka 2006). Hebrew had already become rare or extinct during the Roman Empire, so these were, in effect, dictionaries of a dead language. For two millennia Hebrew survived mainly (or only) as a liturgical and literary language, being preserved and cherished as a symbol of the ethnic and religious identity of Jews during the diaspora. From at least the 2nd century AD it was no longer a medium of everyday communication. For that purpose, it was supplanted first by Aramaic and subsequently by Arabic. Sefer ha-Egron (902) is a lexicon of approximately 1000 Hebrew words for poetic purposes, compiled in Egypt by Sa'adiah ben Josef. The words are presented in two arrangements: first, they are listed alphabetically, while the second list is of words according to their final consonant, in order to facilitate rhyming. Some years later Sa'adiah issued a version of his work with glosses in Arabic, to facilitate understanding of the meaning of Hebrew words. Kitab Jami al-Alfaz (c. 945) is a Hebrew-Arabic biblical dictionary compiled in Fez, Morocco, by David ben-Abraham El-Fasi. According to Cohen and Choueka (2006), "probably the finest achievement of medieval Hebrew lexicography is 'the Book of Roots' by Yonah ibn-Janah." Ibn-Janah, otherwise known as Abu al-WalīdMarwān ibn-Janāh. His lexicon, Kitab al-'usul, is the second part of a work known as Kitab al-Anqih 'the Book of Exact Investigation'.  The order of the words of the lexicon is based on the three-letter root system that is now recognized as universal for Semitic languages. The description and details of glosses are in Arabic. In the thirteenth century, a revised grammar and lexicon based on the work of IbnJanah was compiled in Narbonne by the Biblical scholar and philosopher Rabbi David Qimhi (ca. 12 A.D.).

In the middle ages,     Europe has developed its lexicography in the form of interlinear vernacular glossing of words of Latin manuscripts. These glosses then compiled in alphabetical order.  The earliest known example is the 8th-century Glosses de Reichenau consisted of over 5,000 words of the Latin Vulgate with glosses in Gallo-Roman a precursor of Old French. Similarly, in England    a glossary, Promptorium     Parvulorum ('Young People's Storeroom'), compiled in about 1440 by Galfridus Anglicus with 10,000 entries. Then, in 1499, the

Promptorium was set in type and printed. Thus, the invention of printing is of the great importance in the history of lexicography.

**References:**

1. Bharath, Ravipati. 2008. An Ontological Classification of Telugu Nouns. (Unpublished thesis). University of Hyderabad. Hyderabad.
2. Cohen, Chaim, and Ya'akov Choueka. 2006. 'Hebrew Lexicography' in K. Brown (ed.), *Encyclopedia of Language and Linguistics,* second edition; vol.5, pp.168-272. Elsevier.
3. Drory, Rina. 2000. *Models and contacts: Arabic literature and its impact on medieval Jewish culture*. Leiden: E. J. Brill.
4. Dr. N. S. Sarap, Teachers' Perception Of Students' Attitude And Motivation Towards B.Sc. Agriculture English Course; English Studies International Research Journal : Volume 3 Issue 2 (2015) Issn 2347-3479, Pg 13-16
5. Fellbaum, Christiane (ed.). 1998. *WordNet: an Electronic Lexical Database.* Cambridge, MA: MIT Press.
6. Ferri, Rolando (ed.). 2011. *The Latin of Roman Lexicography*. Pisa: Fabrizio Serra Editori.
7. Goldberg, Adele (1995) *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.
8. Dr. Gargi Talapatra, S. Radhakrishnan's Translation Of The Principal Upanishads: A Study In Spiritual Diaspora; English Studies International Research Journal : Volume 3 Issue 2 (2015) Issn 2347-3479, Pg 17-19
9. Goldberg, Adele (2006) *Constructions at Work*. Oxford: Oxford University Press.
10. Hanks, Patrick. 2010. 'Lexicography, Printing Technology, and the Spread of Renaissance Culture.' In A. Dykstra et al. (eds), *Euralex Proceedings*. Leeuwarden: Frisian Institute.
11. Hanks, Patrick. 2013. *Lexical Analysis: Norms and Exploitations.* MIT Press.
12. Eliza Binte Elahi, Emancipation Of Women And Feminist Ideas In Begum Rokeya Sakhawat Hossain's Utopias; English Studies International Research Journal : Volume 3 Issue 2 (2015) Issn 2347-3479, Pg 20-23
13. Li Ming. 2006. 'Chinese Lexicography' in K. Brown (ed.), *Encyclopedia of Language and Linguistics,* Second Edition. Elsevier.
14. Sinclair, J. M. 1998. 'The Lexical Item' in E. Weigand (ed.), *Contrastive Lexical Semantics*. Benjamins.
15. Dr. D. R. Pratima Roy, Evolution Of Sri Sri From Translator To Transcreator In To Poesy: A Rhapsody; English Studies International Research Journal : Volume 3 Issue 2 (2015) Issn 2347-3479, Pg 24-26
16. Sinclair, John. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.
17. Trench, Richard Chenevix. 1858. 'On some deficiencies in our English dictionaries'. In: *Proceedings of the Philological Society*.
18. Yong Heming and Peng Jing. 2008. *Chinese Lexicography: a History from 1046 BC to AD 1911.* Oxford: Oxford University Press.
19. Shashikala Walmiki, Portrait Of Iconoclastic Female Character In Jhumpa Lahiri's "The Lowland" ; English Studies International Research Journal : Volume 3 Issue 2 (2015) Issn 2347-3479, Pg 27-31

Sampath Babu Tokala
Doctoral Scholar, Centre for ALTS
School of Humanities, University of Hyderabad, Hyderabad – 500046.