

GENERAL PRINCIPLES OF CREATING ELECTRONIC DICTIONARIES

Mokhiyakhon Uzokova

mohiyaxonuzkova@gmail.com

Bachelor student, Tashkent State university of Uzbek language and literature

ABSTRACT

The issue of electronic dictionaries has been relevant since the advent of computers and the Internet. Although electronic dictionaries have different forms: web, mobile, desktop, the creation of electronic dictionaries on any topic is based on general principles. The article focuses on these general principles and their role in Uzbek computer lexicography.

Keywords: electronic dictionary, software development, computer lexicography, web application, database, mobile application, desktop application.

INTRODUCTION

The early use of computers in lexicography Computers was first employed in lexicography in the 1960s¹. Random House Dictionary of the English Language (1967) and New Merriam-Webster Pocket Dictionary which was based on the paper-based Webster's 7th New Collegiate Dictionary were initial initiatives in computerized dictionaries. At the same time handheld dictionaries, such as LK-3000 (Nixdorf (now Siemens) in 1979), Craig M100 (1978), and Japanese Speak & Spell (1976) were the first electronic dictionaries with interfaces designed for human users. A need for a portable and well-documented dictionary led to the creation of CD format dictionaries (Oxford English Dictionary, 1988, 1999, 2002). When the Internet became viral, dictionary-makers launched online dictionaries on that (CERN, 1993). But since internet services were usually provided free of charge, and little was done to guard against copyright infringement, they stopped offer valid dictionaries. The convenience of the web-based dictionaries, however, overcame all obstacles, and now dominates among software platforms devoted to electronic dictionaries.

METHODOLOGY

The study of computational lexicography began at the same time when computers became a tool for dictionaries. Work J. of Byrd and others (Tools and

¹ Electronic Lexicography: Traditional and Modern Approaches M. V. Makarych, Yu. B. Popova¹, M. O. Shved².

methods for computational lexicology) describes tools and methods used by the Lexical Systems project at IBM Research. Individual tools and their use in that project have been described elsewhere (Chodorow, et al. (1985), Byrd and Chodorow (1985), Byrd, et al.(1986b), Chodorow and Ravin (1987), Neff and Byrd (1987))². ‘The Oxford History of English Lexicography’ book an author Hilary Nesi (Coventry University) wrote a detailed review about the history of electronic dictionaries (Dictionaries in electronic form, 2009). M. V. Makarych, Yu. B. Popova, M. O. Shved also made a great contribution to the development of such dictionaries. ‘Electronic Lexicography: Traditional and Modern Approaches’ (M. V. Makarych and others, 2020) shows a novel way of creating a database on the example of multilingual English-Belarusian-Russian dictionary. This paper is intended to contribute to those kinds of literature by generalizing concepts that are important to create an optimal electronic dictionary.

RESULTS AND DISCUSSION

A general concept of a dictionary is providing a word with a valid definition and translation if it is multilingual. However, in contemporary dictionaries, only a definition may not satisfy a user’s demand. Factors that need to be considered in creating any type of dictionary are:

1. Creating a database of a dictionary.

The categorization of the database truly depends on an author’s preference. What can be, however, common are:

- 1.1. Word (transcript) + pic (if applicable) + part of a speech.
- 1.2. Meaning.
- 1.3. Etymology (optional).
- 1.4. Usage.
- 1.5. Related words.

Most common online dictionaries dictionary.com, grammota.ru and dictionary.cambridge.org follow this structure. To maintain the attention of a user, dictionaries may offer trending words (‘word of the day’, ‘word of the year’) along with games/quizzes based on the word that has been chosen and/or a list of word categories³. Adding spelling of the word or a whole sentence (if applicable) can also double facilitation.

² TOOLS AND METHODS FOR COMPUTATIONAL LEXICOLOGY Roy J. Byrd Nicoletta Calzolari* Martin S. Chodorow** Judith L. Klavans Mary S. Neff Omneya A. Rizk*** IBM T. J. Watson Research Center Yorktown Heights, New York 10598

³ Dictionaries in electronic form, Hilary Nesi. in: Cowie, A. P. (ed) The Oxford History of English Lexicography Oxford: Oxford University Press 458-478 (2008)

Here is below an example of a word “Spectacle” taken from dictionary.com online English dictionary:

1.1. Word (transcript): *Spectacle* [spek-tuh-kuhl], noun.

1.2. Meaning:

1. anything presented to the sight or view, especially something of a striking or impressive kind.
2. a public show or display, especially on a large scale.
3. spectacles. eyeglasses, especially with pieces passing over or around the ears for holding them in place.
4. Often **spectacles**.
 - a) something resembling spectacles in shape or function.
 - b) any of various devices suggesting spectacles, as one attached to a semaphore to display lights or different colors by colored glass.

5. *Obsolete*. a spyglass.

1.3. Etymology (optional): 1300–50; Middle English < Latin *spectāculum* a sight, spectacle, derivative of *spectāre*, frequentative of *specere* to look, regard. See cle².

1.4. Usage:

1. This is not high production value YouTube, or YouTube driven by *spectacle* or personality.
2. More meta-comedy than action *spectacle*, it’s the rare superhero story that could potentially appeal to viewers, like me, whose eyes glaze over when battle scenes run longer than a few minutes.
3. To be sure, the football *spectacle* changed to accommodate the realities of the war.
4. Wilkie speculated in an email that Takano was “laying the grounds for a *spectacle*.”

1.5. Related words (synonyms, antonyms and phrases; nearby words; words from the word):

OTHER WORDS FROM SPECTACLE

spec-ta-cle-less – adjective; *spec-ta-cle-like* – adjective; *su-per-spec-ta-cle* – noun.

WORDS RELATED TO SPECTACLE

pageant, tableau, parade, spectacular, movie, scene, display, drama, sight, performance, comedy, demonstration, event, extravaganza, phenomenon, play, marvel, show, production, exposition

IDIOMS ABOUT SPECTACLE

make a spectacle of oneself, to call attention to one's unseemly behavior; behave foolishly or badly in public: *They tell me I made a spectacle of myself at the party last night.*

OTHER WORDS FOR SPECTACLE

*marvel, wonder, sight, show.*⁴

When it comes to a database, a source of data collection prioritizes its significance. In a traditional method of collection, data can be compiled through a large number of paper dictionaries. This, on the one hand, might be a time-consuming process, as a collector first needs to find related words and enter them into an electronic spreadsheet. On the other hand, since paper books cannot be updated as frequently as e-sources, combined data may become outdated in a short period of time. The wealth of information in published dictionaries can be tapped by semi-automatic and fully-automatic methods, to help build the computerized lexicons⁵. Online sequential processing of data using related periodicals can be considered another innovative way of data collection, which significantly reduces man-involved procedures and gathers up-to-date information. Scientists and creators of an electronic multilingual translation dictionary of Belarusian used the second method in the data processing⁶.

2. Text normalization.

Once machine-readable dictionary tapes are acquired from the publishers, they usually have to be cleaned up and / or normalized⁷. Row data simply needs to be technically corrected and converted into one format. Reviewing and editing might be done before or after storing the data in a computerized format, as in both ways an automatic process is involved.

3. Developing a software for an electronic dictionary.

When a database is ready to process, the next step would be thinking of an appropriate software environment where all this will be called. Since the object of the research is depicting key factors of creating digital dictionaries, these software platforms should be considered: web, mobile, and desktop. Each of them has both advantages and disadvantages when it serves as a basic environment for a database.

⁴ <https://www.dictionary.com/browse/spectacle>

⁵ TOOLS AND METHODS FOR COMPUTATIONAL LEXICOLOGY Roy J. Byrd Nicoletta Calzolari* Martin S. Chodorow** Judith L. Klavans Mary S. Neff Omneya A. Rizk*** IBM T. J. Watson Research Center Yorktown Heights, New York 10598

⁶ M. V. Makarych, Yu. B. Popova, M. O. Shved. Electronic Lexicography: Traditional and Modern Approaches. -Science and Technique. V. 19, No 5 (2020).

⁷ TOOLS AND METHODS FOR COMPUTATIONAL LEXICOLOGY Roy J. Byrd Nicoletta Calzolari* Martin S. Chodorow** Judith L. Klavans Mary S. Neff Omneya A. Rizk*** IBM T. J. Watson Research Center Yorktown Heights, New York 10598

Web-app dictionaries

Advantages	Disadvantages
A web application is a client-server application (the client is a browser, and the server is a web server).	The user needs access to the Internet or to the organization local network if the server is located there
Data in this app is stored mainly on the server and is exchanged over the network	It is dependent on the network, so it can't be used offline.
Users do not depend on the operating system, so web applications are cross-platform and a number of people can use it at the same time.	Do not give a chance people to use simultaneously.
They do not require installation on a computer, tablet or smartphone (they do not occupy a device memory unlike mobile applications)	Might take a huge amount of space on a storage.
Mobile / Desktop app dictionaries	
The application works offline.	For changing the design, making any improvements, developing additional features the user needs to download it himself from the store.
All the data that the mobile application works with is stored on the user's device.	Longer and more expensive development process than for client-server development with similar functionality

<p>The user’s confidence in the security of the product, because app developers make thorough checking of the proposed product for viruses.</p>	<p>The mobile application is almost always in addition to the web application, i. e. the company begins with the development of a web application and only then moves to its mobile counterpart.</p>
---	--

3. Evaluating suitability of a dictionary for an auditory, as well as, various systems (IOS, Android, etc.)

When the product is ready to use, it should be tested on various software systems installed on devices with different screen radiuses, such as tablets, laptops, and mobile phones. It increases the reliability of the product, helping both authors and users to review the current condition.

4. Vocabulary control.

Authorship must be controlled in both creators’ and users’ addition and interaction. It is better if users are also added to the list of vocabulary editors, who may add a new word or edit a definition while searching.

CONCLUSION

Taking into account the abovementioned key factors might help to create high-quality electronic dictionaries at the level of world standards, such as dictionaries created in Uzbek, but have not been electrified yet: “Explanatory dictionary of Uzbek language”, “Etymological dictionary of Uzbek language”, “Morpheme dictionary of Uzbek language”, “Explanatory language of Navo’i works dictionary” encyclopedic dictionaries, translation dictionaries; Explanatory dictionaries on various fields, “Dictionary of synonyms of the Uzbek language”, “Dictionary of antonyms of the Uzbek language”, “Dictionary of compound words”, etc. This, in turn, will take Uzbek computer lexicography to another level, encouraging new solutions to the problems associated with natural language processing.

REFERENCES

1. Hilary Nesi. Dictionaries in electronic form// The Oxford History of English Lexicography Oxford: Oxford University Press, 2008.

2. Roy J. Byrd Nicoletta Calzolari, Martin S. Chodorow, Judith L., Klavans Mary S., Neff Omneya A. Rizk. TOOLS AND METHODS FOR COMPUTATIONAL LEXICOLOGY/. - Computational Linguistics, Volume 13, Numbers 3-4, July-December 1987.
3. M. V. Makarych, Yu. B. Popova, M. O. Shved. Electronic Lexicography: Traditional and Modern Approaches. -Science and Technique. V. 19, No 5 (2020).
4. Yu.B. Popova, M.V. Makarych, M.A. Shved. MOBILE APPLICATION FOR THE ENGLISH-BELARUSIAN-RUSSIAN TECHNICAL DICTIONARY.
5. Online dictionary of English words [Electronic resource]. Mode of access: <https://www.dictionary.com/>
6. Cambridge Online dictionary of English words [Electronic resource]. Mode of access: <https://dictionary.cambridge.org/>
7. Online dictionary of English words [Electronic resource]. Mode of access: <https://grammota.ru/>