

Corpus of American Commencement Speeches (CACS) in the Context of Cognitive Linguistics: Text Selection Criteria and Compilation Stages

V. O. Honhalo

O. O. Potebnia Institute of Linguistics, National Academy of Sciences of Ukraine, Kyiv, Ukraine
Corresponding author. E-mail: viktoriahonhalo@gmail.com

Paper received 15.05.22; Accepted for publication 29.05.22.

<https://doi.org/10.31174/SEND-Ph2022-266X78-04>

Abstract. The article examines the process and criteria for selecting authentic linguistic texts with the help of online YouTube video sharing platform, processing and compilation of texts into an electronic specialized corpus – Corpus of American Commencement Speeches 2000-2021 (CACS). The research is conducted in the field of Cognitive Linguistics and Conceptology, namely in the study of American picture of the world, the phenomenon of commencement speeches, as well as the verbalization of American concept of SUCCESS. The CACS was created with a purpose of profound qualitative research of the structure of the concept of SUCCESS in the modern American culture on the basis of the transcribed American commencement speeches of 2000-2021.

Keywords: *cognitive linguistics, corpus linguistics, conceptology, qualitative methods, concept of SUCCESS, American picture of the world, commencement speech, graduation speech, YouTube, AntConc.*

Introduction. Cognitive Linguistics is the study of language in its cognitive function, where cognitive refers to the crucial role of intermediate informational structures in our encounters with the world. Cognitive Linguistics assumes that our interaction with the world is mediated through informational structures in the mind and focuses on natural language as a means for organizing, processing and conveying the information. Language is seen as a repository of world knowledge, a structured collection of meaningful categories that help to manage and store experiences [9].

Cognitive Linguistics has been thoroughly studied by N. D. Arutyunova, V. Dem'jankov, I. O. Holubovska, E. S. Kubryakova, V. I. Karasik, S. Kh. Lyapin, V. A. Maslova, Z. D. Popova, T. V. Radzievska, J. A. Sternin, N. V. Volosukhina, O. P. Vorobyova, E. A. Uzentsova, S. A. Zhabotinska, S. M. Shcherbyna, G. Lakoff, R. Langacker, R. Jackendoff, M. Johnson, Ch. Fillmore, L. Talmy, A. Goldberg, John R. Taylor, and many others.

In "Introducing Cognitive Linguistics" D. Geeraerts and H. Cuyckens argue that within the primary importance of Cognitive Linguistics are the structural characteristics of natural language *categorization* (such as prototypicality, systematic polysemy, cognitive models, mental imagery, and metaphor); *the functional principles of linguistic organization*; *the conceptual interface* between syntax and semantics; *the experiential and pragmatic background of language-in-use*; the relationship between language and thought, including questions about relativism and conceptual universals. The formal structures of language are studied as reflections of general conceptual organization, categorization principles, processing mechanisms, and experiential and environmental influences [9].

Since Cognitive Linguistics sees language as embedded in the overall cognitive capacities of man, topics of special interest for Cognitive Linguistics include *categorization*. In our research we are aiming at studying the concept (category) of SUCCESS based on American commencement speeches which history traces back to the 17th century, according to Ch. A. Wagner [17].

The U.S. influence on the world political arena is considerable. In order to understand the goals associated with

success in American society, we believe that it is important to study the commencement speeches that are delivered at graduation ceremonies in American educational institutions. These kinds of speeches are a distinctive phenomenon of American culture, that reflect its history, traditions, contemporary American worldview, and vision of the future. The values that are passed on to the younger generation have a huge impact on the future and well-being of the USA and other countries all around the globe. In this regard, the study of American concept of SUCCESS is relevant and to help us increase the accuracy of the cognitive study, it was decided to create a corpus of American commencement speeches, which would give the study not only qualitative, but also decent quantitative results.

The **purpose** of this article is to showcase the process of creating the **Corpus of American Commencement Speeches (CACS 2000-2021)**, which will serve as a basis for a more accurate and independent description and modeling of the prototypical structure of the concept of SUCCESS, as well as its main structure.

Materials and methods. In terms of a Cognitive Linguistic methodology, one of the basic general cognitive abilities reflected in the structure and use of language is prototype categorization: human categories are normally characterized by having one typical member of a category (a prototype), to which other members are related in a motivated way. A. Barcelona and J. Valenzuela claim that our large, complex conceptual structures are invoked on language use and comprehension, and that conventional meanings arise on the basis of experience and general knowledge. This continuum between language and experience explains the fact that the study of conceptual structures or cognitive models as reflected in language has been an important area of research in Cognitive Linguistics [6].

In the Eastern Linguistic tradition such categories in Cognitive Linguistics are called *concepts*. In modern Cognitive Linguistics there is no unanimously accepted definition of the term *concept*, but in our study *concept* is understood as a complex culturally determined multi-substrate mental formation.

Drawing on its insights from categorization theory Cognitive Linguistics has been able to provide principled explanations to the most rampant and previously unanalyzed forms in the lexicon [6].

Methods of Cognitive Linguistics have taken a quantitative turn in recent years including quantitative analysis of language data, and results are often measured using statistical methods. L. A. Janda singles out the methods that have proven useful and versatile in linguistic analysis: chi-square, Fisher test, binomial test, ANOVA, correlation, regression, and cluster analysis [13].

In our research we are aiming at investigating the concept (category) of SUCCESS in American culture, and interpreting the structure of this category based on the American Commencement speeches of 2000-2021. The study of the American concept of success exploits the potential of Corpus Linguistics as a powerful and innovative quantitative method. Nowadays there is a rapid development of corpus researches all over the world (V. V. Zhukovska, V. A. Shyrokov, V. V. Rykov, E. A. Karpilovska, V. P. Zakharov, N. B. Gvishiani, O. M. Karpova, D. Biber, T. McEnery, A. Wilson, J. Johansson, Ch. Meyer, W. Croft, S. Adolphs, Stuart A. Yeats, etc.).

Previously, the method has been successfully used in the cognitive study by O. V. Skobnikova, Ukrainian researcher who studied linguistic, cultural and semiotic aspects of representation of the concept of FAMILY in American national corpora and family film texts. The means of verbalization of the concept were represented and verified by quantitative calculations using the corpus analysis on the material of national corpora and Corpus of American Family Movies, compiled by O. V. Skobnikova [3].

In the current research, we draw on the successful experience of researching cognitive categories (concepts) with the help of a personally compiled corpus of relevant texts. Having interpreted and supplemented the method, we applied it to the study of American concept of SUCCESS and compiled our own corpus of authentic texts – **Corpus of American Commencement Speeches (2000-2021) (CACS)**.

The basis of Corpus Linguistics is a corpus of texts that are specially selected according to various criteria. According to Edward Finegan, corpus is a representative collection of texts in electronic format that contains information about the situation in which the text was created, i.e. information about the speaker, author, addressee or audience [8].

T. McEnery and E. Willson define a corpus as a collection of language fragments selected according to definite linguistic criteria for use as a model of language [14].

Corpora have certain features which distinguish corpus from a collection of texts: representativeness, authenticity, selection, balance, machine-readability [2] which we tried to maintain in our CACS.

The main advantage of corpus-linguistic methods used to study American commencement speeches is that they make it possible to carry out quantitative and qualitative studies of the verbalization of the concept of success and to specify the frequency parameters of lexical units and vocabulary patterns. Moreover, methods of Corpus Linguistics help to obtain reliable and free of subjective as-

sumptions data, which represent the structure of a linguistic concept.

There exists a number of tools, used to process natural language and different corpora, e.g. corpus managers such as WordSmith Tools, AntConc, Concordance, Monoconc, Textstate, etc [7].

To work with CACS, the corpus manager AntConc 3.5.9. [5] was used, developed by Dr. Laurence Anthony and designed for first order corpus processing. AntConc can be used to obtain subject-related vocabulary minimums, lists of fixed word collocations (including terminological ones), and samples for theme groups of words. AntConc also enables contextual analysis (analysis of suggested concordances) and searching for contexts. The program can also make frequency lists of all words in the corpus and show keywords of documents. The program also generates frequency lists of all words in the corpus and displays key words of containing documents [5].

Results. In the course of the research, the stages of creating a linguistic corpus were singled out and described, as well as the source and criteria for selecting authentic linguistic material, i.e. commencement speeches, were thought through and elaborated.

Basic CACS 2000-2021 characteristics: a **monolingual corpus** containing texts in English only; a **synchronic corpus** which texts come from the same range of time (2000-2021); a **static corpus** which development is complete at the point; a **specialized corpus** that contains texts limited to one subject area and genre, American commencement speeches specifically; a **full-text corpus; annotated**.

Additionally, during the creation process, SketchEngine [15] was used to determine the quantitative characteristics of the CACS corpus: **317,186 tokens** (267,987 words) and **6,918 concordances**. The value of concordances is in the fact that they are not an individual list of separate words, but the context of the researched language units' usage. The corpus manager allows finding all occurrences of a unit's usage (e.g. *success*) in KWIC (keyword in context) format in the corpus [6].

Conditionally referential to CACS is the structure of the lexical-semantic field (LSF) SUCCESS, which was previously modeled taking into account the quantitative indicators of LSF units' usage on the basis of ten English lexicographic dictionaries [1].

The process of compiling a linguistic corpus went through **seven stages:**

1. Genre and source of linguistic texts – commencement speeches posted on <https://www.youtube.com/> [18]:

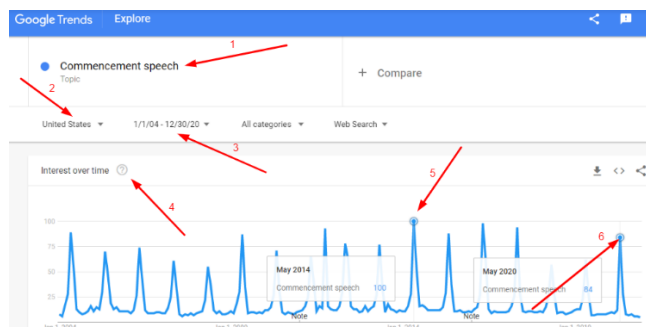
YouTube is an American online video sharing and social media platform. It has been chosen as a source of authentic material for the CACS as it is the most popular video hosting worldwide with over 200 million users in the US [19]. We believe that YouTube gives a unique opportunity to conduct a bias-free and independent research on commencement speeches by forming a list of popular commencement speeches of all times.

2. Search query – commencement speech.

Having researched the notion of commencement address, we have taken out a few options of what people call them: *commencement speech*, *commencement address*, *graduation speech*, *valedictorian address (valediction)* or

even *baccalaureate sermon*. While working with YouTube, it is important to have only one search query (keyword/term), and to select the most popular option we used the Google Trends service (2020) – a website by Google that analyzes the popularity of top search queries in Google Search across various regions and languages [10].

Having checked out all five possible options mentioned above, it was clear that the ‘*commencement speech*’ term has been the best among all, due to the frequency and popularity of its usage among the Internet users in the United States since 2004 (earliest possible) to 2020. See the following detailed graph on the ‘commencement speech’ term:

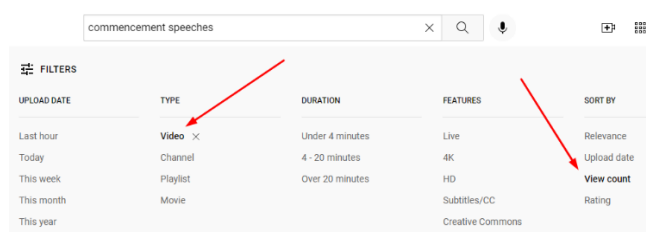


Pic.1. Google Trends results on “Commencement speech” [10] (1.Search query – Commencement speech; 2. Region – United States; 3. Time range – 1/01/2004 – 12/30/2020; 4. Interest over time – numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term (5) – May 2004. The last point seen on the chart is May 2020 (6) with a score of 84, which is more than half popular (50)).

As you can see on the screenshot (Pic.1), the graph has a jump-like character. This may be attributed to the fact that interest in the commencement speeches increases in May, exactly when graduation ceremonies are held at all U.S. institutions. The term peaked its popularity in May 2014. After 2017, you can see a significant decline in interest in commencement speeches. But in 2018-2020, the situation changes, in this period there is a significant growth in the popularity of commencement speeches among Internet users, YouTube included.

1. Setting YouTube search filters:

According to Google Trends, commencement speech is the best keyword to search for the recorded videos of official commencement speech deliveries on YouTube. Based on the search results, 100 speeches were selected to compile the CACS 2000-2021. It was also decided to set a few search results filters for more relevancies: – TYPE: Video; – SORT BY: View count (Pic.2).



Pic.2. YouTube search results screenshot with set filters [18].

The reason for selecting the **View count** filter is that it shows the level of popularity of the video despite all possible heated debates or negative feedbacks on any of the speeches. In the meantime, the **Rating** filter outputs only unpopular commencement speeches (less than 10000 views) with almost 100% of likes.

2. Selection criteria for commencement speeches for CACS.

Thorough attention has been paid to work out the selection criteria for commencement speeches found in the search results on YouTube.

a. Video format: the commencement speech delivery is recorded on a video and is posted on YouTube (speech transcript is not enough).

b. Integrity: all recorded videos found on the search query must be a recording of the whole speech, not a cut out fragment.

c. Credibility: all recorded videos must include the following official information: date, location of the ceremony and name of the speaker. All videos must contain a real formal speech at the graduation ceremony of a particular school. The selected video cannot be a parody, reality show fragment, news broadcast etc.

d. Time range: all graduation ceremonies must be held within 2000-2021 time range, because perception of *success* evolves over time and it was decided to limit the list of commencement speeches to the 21st century deliveries.

The list of researched speeches was manually supplemented with commencement speeches of 2020 and 2021 (8 and 6 speeches respectively), since they have not yet reached the first hundred search results for the search query, but are unique and unprecedented in its nature because of first online formats, world pandemic of Covid-19 and United States racial unrest related issues. This time range is also of great importance because it reflects the American worldwide view changes, the shift in American values and priorities in defining and achieving *success*.

e. Popularity: all videos must have a large number of views and are filtered by View count.

f. Language: all speeches must be in English.

g. Location: institutions (university, college, school) where the speech was delivered must be located in the USA.

h. One-time listing: regardless of the number of one speech output on the search query, the speech is included only once to the corpus and all the other following videos of the same speech are skipped.

3. Data input. After the list of 100 commencement speeches was created, there was a need for their transcripts, as the commencement speech texts are presented in plain text format (*plain TEXT,.txt*) in CACS. **Plain text** is a simple sequence of letters, spaces and punctuation marks, so this format is read by most corpus managers. One of the difficulties of the stage was the transcription of more than a half of commencement speeches, which proved to be quite time-consuming (transcription of a 30-minute speech can take up to 5 hours) and challenging [4]. **The total duration** of all selected videos is *14hrs 34 minutes*.

4. Text markup. At this stage all transcripts underwent additional revision and editing. **Managed metadata:**

- short info about the speaker (name and profession)
- title of commencement speech (if any)
- speech specificity (e.g. in a form of a poem/song; political, year of presidential elections; related to Covid-19 pandemics, United States racial unrest, the UN sustainable development goals etc.)
- time, year and location where the graduation ceremony was held
- speech format – *onset/online*
- duration of video
- speech links on YouTube
- information on existing/needed/created transcripts

The usual step of compiling a corpus is **PoS (part-of-speech) tagging** is a special label assigned to each token in a text corpus to indicate the part of speech and often also other grammatical categories. PoS tags are used in corpus searches and in text analysis tools and algorithms [16]. However, the use of AntConc corpus manager does not require any PoS tagging and works well with the ‘simple_word_indexer’ used to process the raw files [4].

5. Corpus manager. The stage of conducting our research on the American concept of SUCCESS using the already compiled CACS. When the text corpus was ready, we proceeded to working with these texts in the **AntConc** corpus manager, which offers a KWIC, Plot Tool, Cluster Tool, N-Grams, Collocates, frequency list generator, a powerful concordance generator, keyword and word lists [4].

It is very important to note, that the identity of the speakers was neither limited, nor specified, by any selection criteria. Having carefully examined history of American commencement speeches and the practice of transferring [11; 17] the right of speech delivery to the invited university guests, it was obvious that the speaker has always been an integral part of the tradition.

The tradition of commencement speech delivery has changed drastically over time along with those who delivered those speeches. From the very beginning of the tradition, speeches were delivered by students during their graduation day exercises; later the tradition was passed on to the university rectors and guests – politicians and state officials. After that, the right to deliver commencement speeches was also given to businessmen, entrepreneurs, actors and singers, athletes, patrons, scientists and other prominent figures. Recently, the tradition of delivering commencement speeches has become popular among valedictorians (*valedictorian address*) – the best (first in the ranking) high school graduates. Thus, the list of 100 commencement speeches includes speeches delivered by world known prominent figures, university students and high school graduates.

The list of 100 commencement speeches (from top one to 100th) manually selected from the YouTube search results and compiled into the CACS 2000-2021: Steve Jobs (Stanford University, 2005), Denzel Washington (Dillard University, 2017), Rick Rigsby (California State University Maritime Academy, 2018), Jim Carrey (Maharishi University, 2014), Kyle Martin (The King’s Academy, 2019), William H. McRaven (University of Texas, 2014). Chase Dahl (Weber High School, 2015), Donovan

Livingston (Harvard University, 2016), Sacha Baron Cohen (Harvard University, 2004), Andy Samberg (Harvard University, 2012), J. K. Rowling (Harvard University, 2008), Barack Obama (Arizona State University, 2009), Matthew McCounaughy (University of Houston, 2015), David Foster (Kenyon College, 2005), Conan O’Brien (Dartmouth College, 2011), Mark Zuckerberg (Harvard University, 2017), Will Ferrell (University of South California, 2017), Sara Abushaar (Harvard University, 2014), Oprah Winfrey (Harvard University, 2013), Natalie Portman (Harvard University, 2015), Ed Helm (Cornell University, 2014), David McCullough (Wellesley High School, 2012), Bill Gates (Harvard University, 2007), Yasmin Younis (Boston University, 2018), Rhett&Link (Harnett Central High, 2013), Charlie Day (Merrimack College, 2014), Sundar Pinchai (online, 2020), Chedwick Boseman (Howard University, 2018), Lou Holtz (Franciscan University, 2015), Chimamanda Ngozi (Harvard University, 2018), Randy Pausch (Carnegie Mellon University, 2008), Oprah Winfrey (Stanford University, 2008), Barack Obama (online, 2020), Michelle Obama (online, 2020), Mindy Kaling (Harvard University, 2014), Beyonce (online, 2020), Stephen Colbert (Northwestern University, 2011), Ellen DeGeneres (Tulane University, 2009), Lisa Kamal (University of Wisconsin-Madison, 2019), Lisa Kudrow (Vassar College, 2010), Barack Obama (Howard University, 2016), Denzel Washington (University of Pennsylvania, 2011), Jaclyn Martson (Rolling Hills Community Church, 2016), Elon Musk (California Institute of Technology, 2012), Chimamanda Ngozi (Wellesley College, 2015), Tim Allen (Hillsade College, 2021), Victor Wooten (University of Vermont, 2016), Fared Zakaria (Harvard University, 2012), Meryl Streep (Columbia University, 2010), Arjun Byju (Pine View High School, 2013), Bill and Melinda Gates (Stanford University, 2014), Dr. Michael J. Burry (University of California, 2012), Fred Rogers (Dartmouth College, 2002), Jeff Bezos (Princeton University, 2010), John Krasinski (Brown University, 2001), Wesley Chan (University of California, 2012), Steven Spielberg (Harvard University, 2016), Sean Combs (Howard University, 2014), Dolla Parton (University of Tennessee, 2009), Matt Damon (Massachusetts Institute of Technology, 2016), John G. Roberts (Cardigan Mountain School, 2016), Amy Poehler (Harvard University, 2011), Sheryl Sandberg (Harvard University, 2014), Tom Hanks (Yale University, 2011), Donald Trump (US Military Academy, 2020), Neil Gaiman (University of the Arts, 2012), Sheryl Sandberg (UC Berkeley, 2016), Michelle Obama (City College of New York, 2016), Kumail Nanjini (Grinnell College, 2017), Salman Khan (Khan Academy, 2012), Barack Obama (Rutgers University, 2016), Larry Page (University of Michigan, 2009), Tyler Perry (Tuskegee University, 2016), Barack Obama (Morehouse College, 2013), Lady Gaga (online, 2020), Abby Wambach (Barnrd College, 2018), Shonda Rhimes (Dartmouth College, 2014), Joss Whedon (Wesleyan University, 2013), George Saunders (Syracuse University, 2013), Michael Lewis (Princeton University, 2012), Aaron Sorkin (Syracuse University, 2012), John Green (Butler University, 2013), Larry Ellison (University of South California, 2016), Sheryl Sandberg (Barnard College, 2011), Daveed Diggs (Brown University, 2017), Stephen

Curry (online, 2020), Elon Musk (University of South California, 2014), Michael Bloomberg (Harvard University, 2014), Tim Cook (Massachusetts Institute of Technology, 2017), Oprah Winfrey (University of South California, 2018), Tim Cook (Stanford University, 2019), Michelle Obama (Tuskegee University, 2015), Brené Brown (University of Texas (online), 2020), Jane Lynch (Smith College, 2012), Michelle Obama (Eastern Kentucky University, 2013), Gwynne Shotwell (Northwestern University (online), 2021), John Legend (University of Pennsylvania, 2021), George RR Martin (Northwestern University, 2021), Robert A. Iger (University of Texas, 2021), Bina Venkataraman (Hillsade College, 2021) [18]*.

*The search results on YouTube are dynamic, as the number of views of recorded commencement speeches is constantly changing. Therefore, it is crucial to underline that *the actual list of 100 speeches considered in the research and added to CACS 2000-2021 is compiled as of September 1, 2021. Future search results for the same search query may vary.*

Conclusions. To conclude, during the cognitive studies in the field of cognitive linguistics and conceptology the method of corpus linguistics was exploited – the creation of an electronic corpus of linguistic texts – Corpus of American Commencement Speeches 2000-2021 (CACS). The corpus is characterized as monolingual, synchronous, specialized, annotated corpus and consists of 100 full-text commencement speeches within 2000-2021 time range.

All 100 videos (14 hrs 34 mins) were transcribed and saved as plain text. CACS consists of **317,186 tokens** (267,987 words) and **6,918 concordances**. Such services as YouTube, Google Trends, AntConc, SketchEngine were used to create the CACS in seven stages. Considerable attention and efforts were paid to the development of commencement speech selection criteria and their manual compilation into a corpus.

The creation of the corpus as part of the study of American concept of success will help to conduct qualitative and quantitative research on the concept. The corpus is also expected to successfully conduct experimental tasks, such as comparing the lexical semantic-field and the concept of SUCCESS by conducting a profound qualitative research and comparison.

CACS will also serve as a powerful tool in the study of American culture, its representation of success both of an individual and the whole country. Commencement speeches are an appreciated American phenomenon that has flourished over the years. It is now used as a powerful message, a transfer of values, experience and life guidelines to the younger generation, not only in the U.S., but all around the world. Additionally, the corpus can be used to study other key concept of American picture of the world, such as *government, freedom, family, career* and many others. Moreover, commencement speeches have been a reflection of political and social issues, which makes it possible to track the evolution of concepts of American reality and its influence on the whole world.

REFERENCES

1. Гонгало В.О. Структурна організація лексико-семантичного поля SUCCESS у сучасній англійській мові. *Закарпатські філологічні студії*. 2021. Вип. 19. Т. 1. С. 112–118.
2. Жуковська В.В. Вступ до корпусної лінгвістики: навчальний посібник. Житомир: Вид-во ЖДУ ім. І. Франка, 2013. 140 с.
3. Скобнікова О. В. Створення власного корпусу американських кіносценаріїв. *Науковий вісник ДДПУ імені І. Франка. Серія «Філологічні науки»*. 2018. Вип. 9. С. 204–207.
4. Anthony L. AntConc Releases/ Guide. Tokyo: Waseda University, 2022. URL: <https://www.laurenceanthony.net/software/antconc/>
5. Anthony L. AntConc. Tokyo: Waseda University, 2022. URL: <https://www.laurenceanthony.net/software/antconc/>
6. Barcelona A., Valenzuela J. “An Overview of Cognitive Linguistics”. 25p. URL: https://www.researchgate.net/publication/242182668_An_Overview_of_Cognitive_Linguistics
7. Corpus Manager: A tool for multilingual corpus analysis/ G. K. Mikros et al., 2007. URL: https://www.researchgate.net/publication/236583638_Corpus_Manager_A_tool_for_multilingual_corpus_analysis
8. Finegan E. “Language: its structure and use”. N.Y.: Harcourt Brace College Publishers, 2004, 607p.
9. Geeraerts D., Cuyckens H. (2012). “Introducing Cognitive Linguistics.” 21 p. URL: https://www.researchgate.net/publication/286721291_Introducing_Cognitive_Linguistics
10. Google Trends. URL: <https://trends.google.com/trends>
11. Gurock J. S. “The Colonial and Early National Period 1654-1849: American Jewish History”. Routledge, 1998, 510 p.
12. <https://www.laurenceanthony.net/software/antconc/releases/AntConc405/help.pdf>
13. Janda L. A. “Quantitative Methods in Cognitive Linguistics. An Introduction”. 32 p. URL: <https://munin.uit.no/bitstream/handle/10037/5582/article.pdf?sequence=1&isAllowed=y>
14. McEnery T., Wilson A. “Corpus Linguistics”. Edinburgh: Edinburgh University Press, 2001, 2009p.
15. SketchEngine tool. URL: <https://www.sketchengine.eu/>
16. SketchEngine: POS tags. URL: <https://www.sketchengine.eu/pos-tags/>
17. Wagner Ch. A. “Harvard: Four Centuries and Freedoms (American College and University Series)”. Boston: Dutton, 1950, 326 p.
18. YouTube (video sharing platform). URL: <https://www.youtube.com/>
19. YouTube Users Statistics by GMI Blogger, 2022. URL: <https://www.globalmediainsight.com/blog/youtube-users-statistics/>