

*E. L. Jiménez-Navarro, PhD (Philology)*  
*Universidad de Córdoba, Córdoba, Spain*  
*e-mail: lucia.jimenez@uco.es*

*Э. Л. Хименес-Наварро, д. философии (лингв.)*  
*Университет Кордовы, Кордова, Испания*  
*e-mail: lucia.jimenez@uco.es*

## THE SEMANTIC POTENTIAL OF LEXICAL BUNDLES IN THE RESEARCH ARTICLE GENRE

The current work presents a hybrid (manual and automatic) study of the meaning of sequences of words frequently repeated in a corpus of research articles.

*Key words:* corpus; genre; lexical bundle; research article; semantic frame.

## СЕМАНТИЧЕСКИЙ ПОТЕНЦИАЛ ЛЕКСИЧЕСКИХ СЛОВСОЧЕТАНИЙ В ЖАНРЕ НАУЧНОЙ СТАТЬИ

В работе представлены результаты комплексного исследования (выполняемого человеком и компьютером) значения лексических цепочек слов, часто повторяемых в корпусе исследовательских статей.

*Ключевые слова:* корпус; жанр; лексическая цепочка; исследовательская статья; семантический фрейм

A research article (henceforth RA) is “a written text <...> that reports on some investigation carried out by its author or authors” [1, p. 93]. This type of text is usually written in English and published in high-impact journals for two sound reasons. First, English has become a *lingua franca* [2], therefore, it is used by both native and non-native speakers to interact with others in professional contexts. Second, publication of research findings is of the utmost importance if one purports to reach a wider audience. In order to facilitate the dissemination of knowledge, the language used by scientists must be deeply cared for so that it can become comprehensible to anyone. In this context, the RA genre has attracted discourse analysts’ attention for the last years. More specifically, the phraseology (i. e., combinations of words) included in these papers has sparked the interest of a large number of authors, probably because “each register employs a distinct set of lexical bundles, associated with [its] typical communicative purposes” [3, p. 265].

The purpose of this paper is to contribute to the study of the language of RAs by exploring the semantic potential of 4-element lexical bundles. At this point, it must be remembered that lexical bundles were first identified by Douglas Biber and his colleagues [4; 5] and are defined as “sequences of words that tend to co-occur, irrespective of their idiomaticity” [5, p. 59]. Their analysis has become possible thanks to the advancements that have revolutionized corpus and computational linguistics since the beginning of the twenty-first century. For this reason, my study is based on a specialized corpus and uses computer software. On the one hand, the UCOSCIENCOR corpus was semi-automatically compiled, for the texts included were manually selected and the tagging process was automatically completed by Sketch Engine [6; 7]. On the other hand, Sketch Engine was also the resource used for the automatic extraction of lexical bundles.

Regarding the methodology employed for the semantic analysis of these units, I relied on the theory of Frame Semantics developed by Charles Fillmore [8; 9; 10; 11]. According to this author, words and phrases do not relate to each other, but they belong to frames which provide speakers with the conceptual base to determine the concept that the word or phrase encodes. To put it another way, the conceptual structures underlying the meanings of linguistic entities are produced by semantic frames which contain elements of the kind of situation described, such

as its participants. These frame elements can be classified as core (essential to the meaning of the frame) and non-core (they can be dispensed with). The idea behind this theory is that “people have in memory an inventory of schemata for <...> interpreting experiences” [8, p. 25] and exposure to a specific linguistic form in a given context “evokes” [11, p. 378] a frame in the perceiver’s mind.

The methodological steps followed to achieve my goal were the following: (1) the specialized corpus UCOSCIENCOR was built; (2) 4-element lexical bundles were automatically extracted using the ‘N-grams’ function of Sketch Engine; (3) manual verification to suppress inaccurate multiword units was performed; (4) the classification of the results into semantic frames was determined according to the most prominent word of the lexical bundles (in some cases, the ‘Concordance’ function of Sketch Engine was used to check the combinations in context). Preliminary results show that 18 semantic frames were evoked by the keywords contained in the 160 4-element lexical bundles selected. As an illustration, STATEMENT was the semantic frame activated by a higher number of lexical bundles (23), followed by MEDICAL PROFESSION (15). Further research may target distinct lines. For instance, a study of lexical bundles might be performed in different sections of the RAs separately (e. g., abstract, results, discussion) and a similar research to the one presented here might be conducted in other types of genres, such as law, the environment or tourism.

## REFERENCES

1. Swales J. M. Genre analysis. English in academic and research settings. Cambridge : Cambridge University Press, 1990.
2. Seidlhofer B. English as a lingua franca // *ELT Journal*. 2005. № 59 (4). P. 339–341.
3. Biber D., Barbieri F. Lexical bundles in university spoken and written registers // *English for Specific Purposes*. 2007. № 26 (3). P 263–286.
4. Biber D., Conrad S. Lexical bundles in conversation and academic prose // *Language and Computers*. 1999. № 26. P. 181–190.
5. Biber D. et al. The Longman grammar of spoken and written English / D. Biber, S Johansson, G. Leech, S. Conrad, E. Finegan. London : Longman, 1999.
6. Sketch Engine [Software]. Brno : Lexical Computing Limited, 2021.
7. Kilgariff A. The Sketch Engine: Ten years on // *Lexicography*. 2014. № 1 (1). P. 7–36.
8. Fillmore C. J. Frame semantics and the nature of language // *Annals of the New York Academy of Sciences*. 1976. № 280 (1). P 20–32.
9. Fillmore C. J. Frame semantics. In *Linguistic Society of Korea // Linguistics in the morning calm*. Seoul : Hanshin Publishing Company, 1982. P. 111–137.
10. Fillmore C. J. Frames and the semantics of understanding // *Quaderni di Semantica*. 1985. № 6 (2). P 222–254.
11. Fillmore C. J. Frame semantics. In D. Geeraerts, *Cognitive linguistics: Basic readings*. Berlin/New York : Mouton de Gruyter, 2006. P. 373–400.