

Kaja Mandić

Faculty of Health Studies, University of Mostar, Bosnia and Hercegovina
kaja.mandic@fzs.sum.ba

Nursing Corpus and the Academic Collocation List

Abstract

Knowledge of frequent collocations is very important in the scientific discourse and writers should be familiar with collocations valued in a particular scientific discipline. Collocations enable English language learners to put words commonly used by native speakers to appropriate use in appropriate contexts. The study is based on the comparison of collocations from a corpus of articles written by native speakers of English and collocations from a corpus of English articles written by native speakers of Croatian, Bosnian, and Slovenian with the Academic Collocation List (ACL). The objective of the study is to extract the most frequent two-word collocations from the two nursing corpora and compare them with the ACL. The study will produce two corpora that are classified as specialized corpora. Both native and non-native corpora will include only English scientific articles from the field of nursing science. The results will be useful to generate a field-specific academic vocabulary list and teaching materials in order to strengthen learners' academic reading and writing proficiency.

Keywords: collocations, corpus linguistics, non-native speakers, data analysis, nursing

1. Introduction

English is the language of dissemination of academic texts and therefore of academic knowledge in all parts of the world. Its dominant role as the language of higher education is reflected in the number of scientific articles published in English. Using English as the main medium to transfer new research is the only way for scientists to become recognized and successful in their work and be cited by their peers. It is challenging to define the exact number of indexed articles published in English, but it is now estimated that more than 90% of scientific articles are written in English, regardless of the author's native language (Hamel 2007; Lillis, and Curry 2010; Montgomery 2013). According to Baji et al. (2022), nearly 80% of all indexed journals and the world's top fifty journals are in English. However, non-native writers face difficulties in the production of fluent academic texts and adequate vocabulary knowledge (Crystal 2012; Montgomery 2013). English for Specific Academic Purposes (ESAP) plays a very important role in educating existing and new users

who will join the international network of scientific English (Wood 2001: 71-73). The corpus-based approach is recognized to be particularly suitable for research and teaching of English for Specific, Professional, or Academic Purposes. Lindquist (2009) points out that if one is interested in the workings of a particular language, one efficient way to do this is by using corpus methodology. This means that corpus linguistics allows us to see how language is used today in different contexts, thus enabling us to teach and learn the language in a different way (Bennett 2010).

2. Definition of Collocation and Review of Previous Studies

Collocations are seen as a necessary component of foreign language spoken and written competence and are considered an essential component in developing native-like fluency (Nesselhauf 2005). This means that they strongly contribute to language efficacy, namely language comprehension and production. Collocations are of particular importance for learners striving for a high degree of competence in the second or foreign language; however, they are also important for less ambitious learners, as they enhance both accuracy and fluency (Nesselhauf 2003).

It is a challenging task to define the term collocation. Sinclair as a pioneer of corpus-based research defines collocation as: “The occurrence of two or more words within a short space of each other in text” (1991: 170). A simple definition of the concept of collocation is of arbitrarily restricted lexeme combinations (Nesselhauf 2005). Stubbs (2001) sees a collocation as a relationship of habitual co-occurrence of words, either lemmas or word-forms. Collocations are one type of a group of expressions whose importance in language has been increasingly recognized in recent years. This group of expressions has been variously called prefabricated units, prefabs, phraseological units, (lexical) chunks, multi-word units, or formulaic sequences (Najafi and Talebinezhad 2018). Lexical collocations are the subject of this study, more precisely two-word lexical collocations, noun-noun and adjective-noun combinations. The corpora constructed are specialized corpora consisting of scientific articles from the field of nursing.

Collocations play an essential role in language learning and seem to be the basis of creative language development. They are essential for spoken and written fluency, and the availability of a large number of prefabricated units makes fluent language possible. The use of collocations supports comprehension, as the reader can understand the meaning without having to attend to every word (Aitchison 2003; Hunston and Francis 2000; Najafi and Talebinezhad 2018).

In English for Specific Purposes (ESP), learners have to learn high-priority vocabulary items, which need to be selected and included into learning materials and class activities. What is important in order to ensure their effective learning is that students turn a high proportion of input to which they are exposed into intake (Kavaliauskienė and Janulevičienė 2001). In terms of specific genre analysis and ESP, there

is a wide range of evidence on the relationship between the notion of collocation and scientific writing and style (Gledhill 2000). Knowledge of collocations is essential for English as a Foreign Language (EFL) learners since they take up a large portion of language. According to Partington (1998), registers rely largely on prefabrication, and in many genres of writing, pre-cooked expressions are vital elements. Collocations are a big concern for non-native speakers as they are challenged with the task of producing proficient texts (Seretan et al. 2003: 91), and they can become an issue for them due to the interference of the mother tongue. Hyland (2008) observes that professional collocations are familiar to readers and writers who regularly participate in a particular discourse or a given language community. This conclusion was reached by observing the absence of such clusters and lack of fluency in specific registers in novices or newcomers in these communities. According to Hill (2000), collocations estimate up to 70% of everything we say, hear, read, and write.

A number of corpus-based studies suggests that non-native writers underuse collocations when compared to native peers (Demir 2017; Durrant and Schmitt 2009; Liu and Shaw 2001). Studies on academic word lists have a longer history and greater popularity in corpus linguistics than those on academic collocation lists, but findings and methodological methods from those studies provide a tremendous support for collocation studies. These types of lists entail frequent lexical combinations that enhance English language learners' proficiency in a specialized field. The Academic Word List (AWL) by Coxhead (2000) was the first list that applied frequency, range, and specialized occurrence. It used a corpus of 3.5 million words from four disciplines: commerce, science, humanities, and law (Coxhead and Nation 2001: 255). The AWL inspired numerous studies on word lists that later proved many flaws of Coxhead's list. One of its flaws was that it did not include texts from medicine and used only a limited number of short texts from the law subcorpora (Hyland and Tse 2007). In 2015, Yang also conducted a study closely related to the current one, and created a Nursing Academic Word List. This list contains the most frequent words found in nursing research articles. These findings suggest that it is necessary to generate field specific lists for EFL nursing students and professionals in order to strengthen their academic written and spoken comprehension.

A review of collocation studies and non-native speakers of English indicates that there is a lack of collocational knowledge in EFL learners and lack of sufficient exposure to specific types of collocations (Hill 1999, Henriksen 2013, Nesselhauf 2003, 2005). An example is a corpus of 150 cancer research articles, the results showing that new science is founded on a system of preferred expressions, and that collocation is a fundamental mechanism that allows for new formulations to take place throughout the text (Gledhill 2000). Fan (2009) examined a native and non-native corpus composed of sixty essays written by Hong Kong secondary school students and sixty British essays collected in a comprehensive school in northern England. The results revealed that Hong Kong learners used a limited number of collocations compared to British students. Durrant and Schmitt (2009) investigated the

use of collocations in native and non-native student essays. Non-native essays were written by Turkish and Bulgarian first-year undergraduate and postgraduate students. The findings confirm that native writers use significantly more low-frequency collocations than their non-native peers. Demir (2017) investigated research articles from leading journals on English Language Teaching. On two small corpora, he found that native writers used a higher percentage of collocations than Turkish authors. Two Arabic authors explored the production of English adjective-noun collocations by Arab ESL (English as a Second Language) writers as compared to English native writers. The findings revealed that Arab ESL writers use collocations with a greater frequency, but in a context often judged as “not appropriate” by English language instructors (Qureshi and Nurmukhamedov 2018). Chen and Baker (2010) compared the use of frequent word combinations in academic writing by Chinese EFL university students, native English-speaking university students, and native expert writers. The study revealed differences and similarities between native and learner academic writing. The use of word combinations in non-native and native student essays was very similar, but native expert writers used a wide range of word combinations when compared to others. Navarro Gil and Martinez Caro (2019) explored the use of lexical bundles in a corpus of bachelor dissertations from linguistics and medicine written in English by Spanish native students and compared it with published research articles. The product of the study was a list of 218 different bundles with the intention of assisting ESL and EFL learners in their academic writing. Pavičić Takač and Lukač (2013) confirm that collocations are indeed a problematic area for non-native users of medical English and that teaching of vocabulary through collocations can be very useful. The study indicates that teaching of medical collocations has a significant effect on vocabulary retention, and instructors should strive towards incorporating collocations into their teaching content. According to the literature review, we can conclude that EFL and ESL writers’ use of collocations equals the native writers in terms of frequency and accuracy.

3. Assembly of the Native and Non-Native Nursing Corpora

The two corpora in this study are composed entirely of scientific articles from the field of nursing and are classified as specialized. The argument is that we should use specialized corpora in order to understand academic and professional language instead of general corpora (Connor and Upton 2004). The spread of novel nursing knowledge, which stems from scientific research and is published in nursing journals, is vital for the development of the nursing profession, and can have a much higher quality and be provided much faster compared to what students and working professionals can find in nursing textbooks (Campbell-Crofts 2012: 6). We also believe that an entirely article-based corpus better supports the study and its purposes.

The corpus of English native nursing scientific articles (NNSAC) comprises

1,119,441 words from articles of high-quality journals. The articles were chosen on the basis of the NHARS Selected List of Nursing Journal from 2016. We collected 262 nursing articles from ten nursing journals (see Appendix 1 for a complete list of journals and number of articles). Accessibility and availability in electronic form were an important criterion; therefore, the number of articles from journals varies, since not all journals and articles are open access and available for download. Our second corpus, the corpus of non-native English nursing scientific articles (NNNSAC), consists of 249 nursing scientific articles from seven nursing journals and includes 930,786 words written by Croatian, Bosnian, and Slovenian authors (see Appendix 2 for a complete list of journals and number of articles). This corpus of non-native articles can be classified as learner corpora. The non-native journals represented here are the only ones available in this field of research in countries where English is not the native language of the speakers who are the subjects of this study. We collected articles from four journals from Croatia, two journals from Bosnia and Herzegovina, and one journal from Slovenia. There is no clear classification of these types of journals in the region; therefore, we selected all that were available and in which nursing scientific articles are published in the English language.

After corpora assembly, the first step was to extract and compare collocations produced by native and non-native speakers of the English language, and compile a list of English nursing collocations. The focus was on noun-noun and adjective-noun collocations. This particular combination has also been confirmed as the most frequent by the founders of the Academic Collocation List (ACL), our reference corpus. Further, we compare the list of nursing collocations (see Appendix 3 for a complete list of collocations) and the ACL (<https://www.pearsonpte.com/teachers/academic-collocation>) in order to establish whether there is a need for a specialized list of collocations for nurses.

3.1. Methods

The selection criterion for the differentiation of native and non-native authors was done according to Wood. In current literature, there is no exact way or agreed criteria to ascertain the first language of a writer, especially in published research articles. As Wood suggests, first authors must have names that are native to the country in question and be affiliated with an institution where this language (English) is spoken as the first language. Wood does not claim that every scientist subjected to this criterion is in fact a native or non-native speaker of English, but that the overall proportion of native and non-native speakers will be approximated by such an operational definition (2001: 79). In our non-native corpus for any author to be regarded as an L1 (native language) Croatian, Bosnian, or Slovenian writer, they must be affiliated with an institution in their home country and have the first and last name considered native to the same country. The same criteria were applied for L1 English writers. We removed all scientific research papers that did not meet Wood's criteria.

We used Shin and Nation's (2008) criteria for identifying collocation but adapted it to serve the purpose of the study. We adapted two out of six criteria: the third criterion on the reference corpus was eliminated, since our reference corpus is the Academic Collocation List, as well as the fourth criterion for frequency, in which we also adopted the normed frequency from Ackerman and Chen (2013), stating that a collocation had to occur 0.2 times per one million words.

Three software programs were used in the present study. The ABBY Fine Reader, for the conversion of PDF files into simple text files. The second software is the Corpus Builder v.2.3., which is an online software program that assembles text, html, or Word files into a single combined file. The third and most important software used in the study is TermeX version 1.0. It is a tool for automatic collocation extraction and terminology lexical construction. It is based on statistical measures called association measures (Delač 2009: 2). What differs it from the rest is that it provides a much wider range of association measures to choose from and outperforms the majority of tools in terms of processing speed.

4. Results and Discussion

The corpus-based study retrieved a total of 480 collocation combinations from the NNSAC and 272 collocations from the NNNSAC. The final product of the comparison of the two corpora is a list of 629 nursing collocations.

Adjective-noun combinations of collocations are the most frequent form of two-word collocations used by both native and non-native speakers of the English language in our study. This particular combination has also been confirmed as the most frequent one by the founders of the ACL. Since the percentages for the use of adjective-noun combinations are similar in both corpora, 53% for the NNSAC and 56% for the NNNSAC, we can conclude that the results oppose those previously found by Durrant and Schmitt (2010). In their study on the exposure to collocations non-native learners receive, with an emphasis on adjective-noun combinations and their retention, they found that there is a shortfall in the use of adjective-noun collocations by non-native learners, and that the use is limited to a certain set of collocations and is not native-like. Our study shows that both native and non-native writers use adjective-noun combinations to a larger extent than other combinations. Since the highest number of our non-native articles comes from Croatian nursing journals, the results on the most frequent collocation combinations might be under the influence of the first language and the fact that the most common collocation pairs in the Croatian language are formed by an adjective and a noun (Hudeček and Mihaljević 2012: 6). Here are some examples in Croatian: *duljina djelovanja, vijek trajanja, stopa rasta, stanje mirovanja, mjera sigurnosti*.

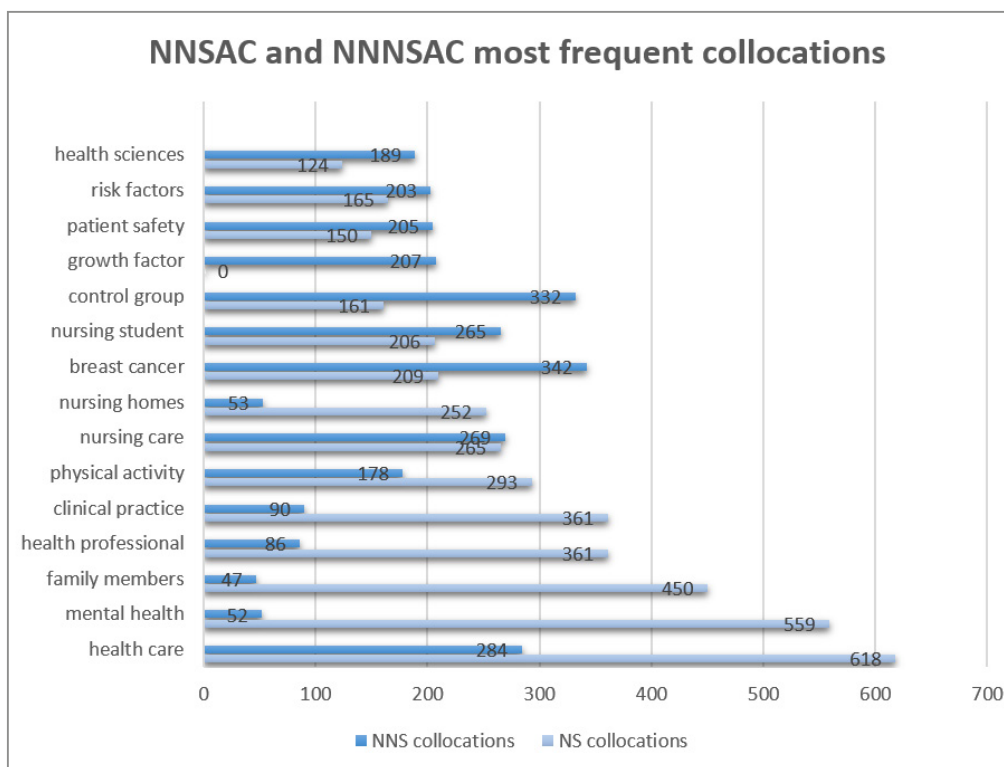


Figure 1. Frequency of the most common collocations from the NNSAC and NNNSAC

Figure 1 shows the most frequent collocations from native and non-native nursing corpora. We can see 10 collocations from the NNSAC and 10 from the NNNSAC as well normed frequency for each extracted collocation. Collocations extracted from the non-native corpus have lower frequencies than those found in the native corpus. The most frequent collocation from the native corpus, *health care*, has the normed frequency of 618, while the most frequent collocation from the non-native corpus, *breast cancer*, has the normed frequency of 342. Collocations with the highest frequencies from the NNSAC are exclusively professional nursing collocations, and three from the NNNSAC (*risk factors*, *growth factor*, and *control group*) are general collocations used across different scientific fields. Some other examples of professional nursing collocations from the native corpus include the following: *blood sugar*, *care facility*, *care staff*, *care team*, *clinical education*, *family caregivers*, *healthcare team*, *nurse leader*, *nurse burnout*.

These results could also be under the influence of topics and particular matters being studied in the selection of native and non-native issues and articles. They could be a by-product of the topic of breast cancer studied in the selection of non-native scientific research articles. The same argument could be true for the native corpus and its second most frequent collocation, *mental health*. This particular collocation could have a high frequency due to the topic studied, since its normed frequency in the non-native corpus is only 52.

We compared two very different corpora regarding size. When combined, our nursing corpora consists of 2,050,227 words from nursing scientific articles written

by native and non-native speakers of English. The corpus of the ACL comprises over 37 million words of academic written and spoken texts from five major English-speaking countries. The list of nursing collocations is assembled on the basis of overlapping collocations from the two corpora of nursing articles, with the addition of all other extracted collocations. The number of overlapping collocations with the ACL is only 151 or to be more accurate only 6% of collocations from the Academic Collocation List appear on the nursing list. Figure 2 represents the proportion of collocations from the ACL and the list of nursing collocations.

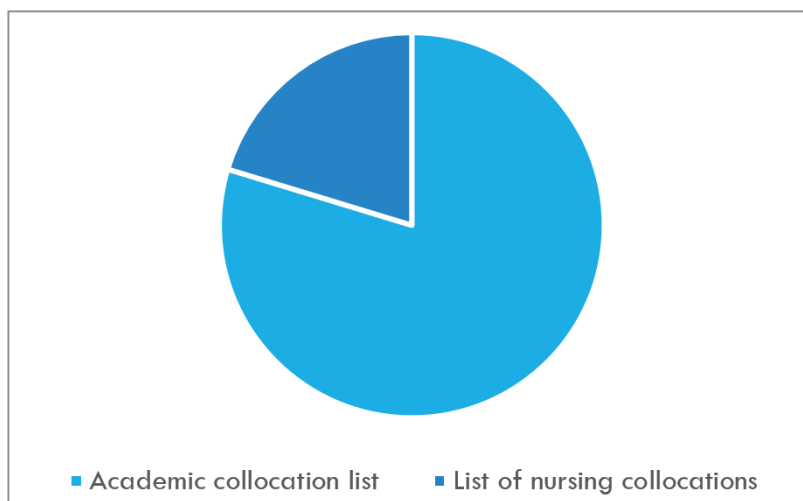


Figure 2. Proportion of overlapping collocations from the list of nursing collocations and the ACL

An interesting finding is that nursing collocations have significantly higher frequencies on the list of nursing collocations than on the ACL. Out of the 151 overlapping collocations, only six collocations (*mental health, mental illness, physical activity, physical symptom, and adverse effect*) are classified as professional nursing collocations. All other overlapping collocations are mostly general ones used in various scientific fields, for example: *available data, research methodology, and study results*. Other examples of overlapping collocations with the ACL include the following: *blood samples, clinical characteristics, community health, diagnostic test, family medicine, food intake, low dose, nursing team, speech disabilities*. We can conclude that nursing collocations have a low occurrence rate on the Academic Collocation List.

5. Limitations

A limitation to the representativeness of the study is that we only looked at noun combinations of collocation (noun-noun and noun-adjective), and our reference list, the ACL, consists of several other collocation combinations. The comparison of the two lists could be more comprehensive if we also included other types of lexical collocations found on the ACL. If this had been the case, the statistical results of this

study would have a much broader framework. Another limitation is that our corpus is assembled only from scientific articles, and the Pearson International Corpus of Academic English contains written and spoken language from lectures, seminars, textbooks, and journal papers. The two nursing corpora of this study could be additionally verified with larger corpora of other medical and healthcare genres.

6. Conclusion

Both the ACL and the list of nursing collocations contain collocations that English learners are likely to encounter during their academic education. The ACL represents the most important cross-disciplinary collocations that can help learners increase their collocation competence and thus their proficiency in academic English. The list created as part of this study comprises collocations purposely tailored for EFL nurses in order for them to better acquire English for specific and academic purposes.

There is a need for an independent list of nursing collocations that will be available for nursing students, working nurses, and other medical professionals who are in need to use English terminology in their writing and specific job situations. This is supported by Gilquin et al. (2007), who state that if we rely solely on native corpus data, the learning materials fail to provide non-native learners with the information that is most valuable to them. According to Hyland (2008), writers need a familiarity with both the clusters which characterize their disciplines and those which are valued in the particular genres of those disciplines. The existence of an independent list of nursing collocations is further supported by Mackin (1978), who claims that collocations are countless, and it is difficult for non-native speakers of English to rule out those unnecessary for them. Our results also support Hyland and Tse's (2007) claim that general vocabulary lists do not reflect the real needs of English for Specific or Academic Purposes students. Providing learners with a list of most frequent academic nursing collocations helps them to overcome the deficiency of academic nursing vocabulary competence. Another reason for an independent list of collocations is the increasing emphasis of researchers on the importance of collocation acquisition in foreign language teaching. Pawley and Syder state that the inventory of multiword combinations known to the mature speaker of English should account for hundreds of thousands (1983: 92).

The development of a discipline specific collocation list hopes to inspire and draw attention of teachers, students, learners, and nurses on the importance of this type of English vocabulary. In order to save time and improve acquisition of specific academic vocabulary, mastery of the list of nursing collocations would be a more valuable decision, since these collocations proved to have higher frequency in nursing scientific articles, which are the base for any future work and professional development in an evidence-based healthcare profession such as nursing. The study results can be used for future design of English nursing materials for reading, writing, and comprehension purposes. They are also beneficial and helpful for nursing

students, nurses, teachers, ESP material designers, and all researchers in applied linguistics and health sciences. The list can serve as a basis for developing English teaching materials for English for Academic Specific Purposes and English for (Professional) Nursing Purposes. Material developers of professional English textbooks should consider using collocations, since multi-word vocabulary items are very frequent in professional academic texts. English foreign language learners who are interested in continuing nursing studies can use the list to help them expand their vocabulary in size and practice, and use the list of nursing collocations as an aid in their academic and scientific writings.

References

- Ackermann, Kirsten; Chen, Yu-Hua. 2013. "Developing the Academic Collocation List (ACL) – A corpus-driven and expert-judged approach." *Journal of English for Academic Purposes* 12(4). 1235-247.
- Bahji, Aness; Acion, Laura; Laslett, Anne-Marie; Adinoff, Bryon. 2022. "Exclusion of the non-English-speaking world from the scientific literature: Recommendations for change for addiction journals and publishers." *Nordic Studies on Alcohol and Drugs* 40(1). 6-13.
- Bennett, Gena. 2010. *Using Corpora in The Language Learning Classroom: Corpus Linguistics for Teachers*. Michigan ELT.
- Biber, Douglas. 2006. *University language: A corpus-based study of spoken and written registers*. Amsterdam: John Benjamins.
- Byrd, Pat; Coxhead, Averil. 2010. "On the other hand: Lexical bundles in academic writing and in the teaching of EAP." *Papers in TESOL* 5. University of Sydney. 31-64.
- Campbell-Crofts, Sandra. 2012. The future of nursing journals. *Renal Society of Australasia Journal* 8(1). 6-8.
- Chen, Yu-Hua; Baker, Paul. 2010. "Lexical bundles in L1 and L2 academic writing." *Language, Learning and Technology* 14(2). 30-49.
- Connor, Ulla; Upton, A. Thomas. 2004. *Discourse in the Profession: Perspectives from the corpus linguistics*. Amsterdam: John Benjamins Publishing Company
- Coxhead, Averil; Nation, Paul. 2001. "The specialised vocabulary of English for academic purposes." In Flowerdew, J. & Peacock, M. (Eds.) *Research perspectives on English for academic purposes*. Cambridge University Press. 252-267.
- Coxhead, Averil. 2000. "A New Academic Word List." *TESOL Quarterly* 34(2). 213-238.
- Crystal, David. 1992. *The Cambridge encyclopaedia of language*. Cambridge: Cambridge University Press.
- Delač, Davor; Krleža, Zoran; Šnajder, Jan; Dalbelo Bašić, Bojana. 2009. "TermeX: A Tool for Collocation Extraction". In Glebukh, A. (Ed.) *Computational Linguistics and Intelligent Text Processing*. 149-157.

- Delač, Davor. 2009. *TermeX v1.0*. University of Zagreb, Faculty of Electrical Engineering and Computing.
- Demir, Cuneyt. 2017. "Lexical Collocations in English: A Comparative Study of Native and Non-native Scholar of English". *Journal of Language and Linguistic Studies* 13(1). 75-78.
- Durrant, Phillip; Schmitt, Norbert. 2009. "To what extent do native and non-native writers make use of collocations?" *International Review of Applied Linguistics in Language Teaching* 47(2). 157-177.
- Fan, May. 2009. "An exploratory study of collocational use by ESL students – A task based approach". *System* 37(1). 110-123.
- Gilquin, Gaëtanelle; Granger, Sylviane; Paquot, Magali. 2007. "Learner Corpora: The Missing Link in EAP Pedagogy". *English for Academic Purposes* 6(4). 1-26.
- Gledhill, Christopher. 2000. "Collocations in Science Writing". *Language in Performance Series* 22. 7-20.
- Granger, Sylviane; Meunier, Fanny. 2008. *Phraseology, An Interdisciplinary Perspective*. John Benjamins Publishing Company.
- Hamel, Rainer Enrique. 2007. "The dominance of English in the international scientific periodical literature and the future of language use in science." *AILA Review* 20(1). 53–71.
- Henriksen, Brigit. 2013. *Research on L2 learners' collocational competence and development – a progress report*. Copenhagen University.
- Hill, Jimmie. 1999. "Collocational competence". *English Teaching Professional* 11. 3-6.
- Hill, Jimmie. 2000. "Revising priorities: From grammatical failure to collocational success." In Lewis, M. (Ed.) *Teaching collocation: Further developments in the lexical approach*. Hove: LPT. 47-67.
- Hoey, Michael. 2005. *Lexical Primings: A new theory of words and language*. Oxon: Routledge.
- Hsu, Wenhua. 2013. "Bridging the vocabulary gap for EFL medical undergraduates: the establishment of a medical word list". *Language Teaching Research* 17(4). 454-484.
- Hudeček, Lana; Mihaljević, Milica. 2012. *Hrvatski terminološki priručnik*. Institut za hrvatski jezik i jezikoslovlje, Zagreb.
- Hyland, Ken; Tse, Polly. 2007. "Is there an academic vocabulary?" *TESOL Quarterly* 41(2). 235-253.
- Hyland, Ken. 2008. "Academic clusters: text patterning in published and postgraduate writing". *International Journal of Applied Linguistics* 18(1). 41-62.
- Kavaliauskienė, Galina; Janulevičienė, Violeta. 2001. "Using the Lexical Approach for the Acquisition of ESP Vocabulary." *The Internet TESL Journal* 7(3). 1-6.
- Kosem, Iztok; Krek, Simon; Gantar, Polona. 2020. "Defining Collocation for Slovenian Lexical Resources." *Slovenščina* 20(2). 1-27.
- Lillis, Theresa; Curry, Mary Jane. 2010. *Academic writing in a global context: The politics and practices of publishing in English*. London, UK: Routledge

- Lindquist, Hans. 2009. *Corpus Linguistics and the Description of English*. Edinburgh University Press.
- Mackin, Ronald. 1978. *On collocations: Words shall be known by the company they keep*. In Strevenes, P. (Ed.) In honour of A.S. Hornby. Oxford University Press. 149-165.
- Montgomery, Scott. 2013. *Does science need a global language? English and the future of research*. Chicago, IL: University of Chicago Press.
- Nesselhauf, Nadja. 2003. The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics* 24(2). 223-242.
- Nesselhauf, Nadja. 2005. *Collocations in a Learner Corpus*. John Benjamins Publishing Company.
- Partington, Alan. 1998. *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam: John Benjamins.
- Pawley, Andrew; Syder, Frances Hodgetts. 1983. Two puzzles for linguistic theory native like selection and nativelike fluency. In Richards, J. C. & Schmidt, R. W. (Eds.) *Language and communication*. London: Longman. 191-230.
- Qureshi, Muhammas; Nurmukhamedov, Ulubek. 2018. "Use of Collocations in Freshman Composition: Implications for L1 English and Arabic ESL Writers." *International Journal of Lexicography* 30(4). 454-482.
- Seretan, Violeta; Nerima, Luka; Wehrli, Eric. 2003. Multi-Word Collocation Extraction by Syntactic Composition of Collocational Bigrams. In Nicolov N., Bontcheva, K., Angelova, G. & Mitkov, R. (Eds.) *Recent Advances in Natural Language Processing III*. John Benjamins Publishing. 91-100.
- Shin, Dongkwang; Nation, Paul. 2008. "Beyond single words: the most frequent collocations in spoken English." *ELT Journal* 62(4). 339-348.
- Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stubbs, Michael. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Wang, Jing; Liang, Shao-Lin; Ge, Guang-Chun. 2008. "Establishment of a medical academic word list." *English for Specific Purposes* 27(4). 442-458.
- Wood, Alisatir. 2001. International scientific English: The language of research scientists around the world. In Flowerdew J. & Peacock M. (Eds.) *Research perspectives on English for academic purposes*. Cambridge: Cambridge University Press. 71-83.
- Yang, Ming-Nuan. 2015. "A nursing academic word list." *English for Specific Purposes* 37(1). 27-38.

Appendix 1

Table 1. Journals and number of articles for The corpus of English native nursing scientific articles

| No. | Journal | Number of articles |
|--------------|--|--------------------|
| 1 | Journal of Clinical Nursing | 83 |
| 2 | Journal of Advanced Nursing | 34 |
| 3 | Scandinavian Journal of Caring Science | 28 |
| 4 | Cancer Nursing | 26 |
| 5 | International Journal of Nursing Practice | 20 |
| 6 | Journal of Nursing Management | 17 |
| 7 | Journal of Nursing Education | 11 |
| 8 | Journal of Psychiatric and Mental Health Nursing | 13 |
| 9 | Oncology Nursing Forum | 8 |
| 10 | Journal of Women's Health | 22 |
| Total | | 262 |

Appendix 2

Table 2. Journals and number of articles for The corpus of English non-native nursing scientific articles

| No. | Journal | Number of articles |
|--------------|---|--------------------|
| 1. | Acta Clinica Croatica | 38 |
| 2. | Bosnian Journal of Basic Medical Sciences | 98 |
| 3. | Croatian Nursing Journal | 22 |
| 4. | Journal of Applied Health Sciences | 6 |
| 5. | Journal of Health Sciences | 35 |
| 6. | Slovenian Nursing Review | 15 |
| 7. | Southeastern European Medical Journal | 31 |
| Total | | 249 |

Appendix 3

Table 3. List of nursing collocations

| | | | | |
|--------------------------|--------------------------|--------------------|------------------------|-----------------------------|
| academic life | academic year | action research | acute care | acute illness |
| adverse effect | adverse events | age group | age range | alcohol abuse |
| alcohol consumption | alcohol use | annual review | anxiety levels | assessment tools |
| available data | average age | average score | average value | behavior pattern |
| beneficial effect | bipolar disorder | blood clot | blood donor | blood flow |
| blood pressure | blood samples | blood sugar | blood vessels | body composition |
| body fat | body image | body language | body mass | body temperature |
| body weight | bone marrow | bone mineral | bowel disease | breast cancer |
| brief review | broad range | caesarean section | cancer care | cancer cell |
| cancer center | cancer diagnosis | cancer nurse | cancer nursing | cancer patients |
| cancer screening | cancer survivor | cancer treatment | cardiovascular disease | care activities |
| care environment | care facility | care home | care management | care plans |
| care providers | care provision | care recipient | care services | care staff |
| care systems | care team | care unit | case report | case study |
| cell growth | cell proliferation | cell tumor | cervical cancer | chemotherapy administration |
| chest injuries | children's hospice | chronic condition | chronic disease | chronic illness |
| chronic pain | chronic wounds | cigarette smoke | clinical care | clinical characteristics |
| clinical community | clinical decision-making | clinical education | clinical examination | clinical experience |
| clinical features | clinical nurse | clinical nursing | clinical outcomes | clinical placement |
| clinical practice | clinical reasoning | clinical setting | clinical skills | clinical staff |
| clinical study | clinical trials | cognitive deficits | cognitive functions | cognitive impairments |
| cognitive performance | cohort study | collective action | colon cancer | communication skills |
| communication strategies | community care | community health | community nurses | community setting |

| | | | | |
|---------------------------|--------------------------|----------------------|-----------------------------|------------------------|
| comparative analysis | comparative study | conceptual framework | conservative management | consultation models |
| continence care | continued use | continuous variable | control group | conventional methods |
| coping strategy | coronary artery | coworker support | critical analysis | critical thinking |
| cruciate ligament | cultural background | cultural differences | current research | current study |
| daily activities | daily living | data analysis | data collection | data extraction |
| data saturation | delivery room | dementia care | demographic characteristics | demographic data |
| dental health | dependent variable | depression scale | descriptive statistics | diabetes patients |
| diagnostic test | direct care | eating disorders | education level | educational programme |
| educational qualification | elderly cancer | elderly patients | eligibility criteria | emergency department |
| emergency room | emotional distress | emotional exhaustion | emotional intelligence | emotional reaction |
| emotional support | empirical research | empirical study | erectile dysfunction | essential role |
| ethical approval | ethics committee | ethnic group | exclusion criteria | exercise program |
| experimental design | expert nurse | eye contact | facial expression | family care |
| family caregivers | family medicine | family members | family support | fatty acids |
| final analysis | final version | focus groups | foetal abnormalities | food intake |
| foot care | future research | future study | gastric cancer | gastrointestinal tract |
| gender differences | general anesthesia | general health | general hospital | general population |
| general practitioner | germ cells | global health | graduate nurse | great impact |
| growth factor | gynecological cancer | health behavior | health care | health condition |
| health education | health information | health insurance | health issues | health organization |
| health outcomes | health policy | health problems | health professionals | health promotion |
| health research | health sciences | health service | health status | health system |
| health wards | healthcare assistant | healthcare costs | healthcare encounters | healthcare environment |
| healthcare institutions | healthcare organizations | healthcare personnel | healthcare professional | healthcare providers |
| healthcare services | healthcare settings | healthcare staff | healthcare system | healthcare team |

| | | | | |
|---------------------------------|------------------------|------------------------|-------------------------|-----------------------------|
| hearing loss | heart disease | heart failure | heart rate | high concentration |
| home care | hormone therapy | hospital admission | hospital care | hospital environment |
| hospital nurse | hospital setting | hospital staff | hospital stay | hot flushes |
| human research | human resource | immune system | important aspects | important factor |
| inclusion criteria | independent variable | individual level | individual needs | informal caregivers |
| information sharing | information sheet | insertion pain | intellectual disability | intensive care |
| interdisciplinary collaboration | internal conflict | international journal | international study | interpersonal relationships |
| ischemic stroke | joint pain | key concept | key findings | key role |
| kidney disease | knee injury | knee joint | labour induction | lacrimonal sac |
| large percentage | leading cause | learning environment | learning strategy | left ventricle |
| life quality | limited information | literature review | local anesthetic | logistic regression |
| longitudinal study | low dose | low level | low score | lower frequency |
| lower rates | lung cancer | magnetic resonance | main category | major cause |
| major change | male caregivers | male infertility | malignant tumor | management strategies |
| managerial competence | manual restraint | marital status | maternity leave | mean age |
| mean score | median age | medical care | medical centers | medical comorbidity |
| medical condition | medical imaging | medical record | medical research | medical sciences |
| medical staff | medical student | medical treatment | medical unit | medical university |
| medication adherence | medication preparation | medication safety | melanoma cells | mental disorder |
| mental health | mental healthcare | mental illness | mental retardation | metabolic activity |
| metabolic syndrome | mortality rate | multidisciplinary team | muscular dystrophy | myocardial infarction |
| negative attitude | negative effect | negative impact | nervous system | normal distribution |
| numerous studies | nurse assessment | nurse assistant | nurse association | nurse burnout |
| nurse care | nurse competence | nurse educator | nurse intervention | nurse leaders |
| nurse managers | nurse practitioner | nurse report | nurse research | nurse specialist |

| | | | | |
|---------------------------|-----------------------|---------------------------|---------------------------|-------------------------------|
| nursing activities | nursing assessment | nursing care | nursing curriculum | nursing documentation |
| nursing education | nursing homes | nursing interventions | nursing journals | nursing management |
| nursing practice | nursing profession | nursing research | nursing roles | nursing schools |
| nursing staff | nursing student | nursing study | nursing team | observational study |
| occasional tiredness | oncology nurse | oncology nursing | oncology patients | online version |
| open access | organ transplantation | original work | outcome measures | outpatient service |
| outpatient utilization | ovarian cancer | overall level | pain assessment | pain management |
| pain rehabilitation | palliative treatment | pancreatic cancer | parental cancer | parental role |
| pathological features | patient care | patient characteristics | patient education | patient experiences |
| patient needs | patient outcomes | patient population | patient safety | patient satisfaction |
| pediatric care | peer support | permanent teeth | personal care | pharmacological interventions |
| phenomenological analysis | physical activity | physical care | physical environment | physical exercise |
| physical health | physical restraints | physical symptom | pilot study | poor communication |
| poor prognosis | positive attitude | positive correlation | positive effect | positive experience |
| positive impact | positive outcome | positive value | postmenopausal women | postoperative care |
| postpartum depression | potential risk | prenatal care | preoperative care | preoperative protocol |
| present study | preterm infants | previous research | previous study | primary care |
| primary data | primary healthcare | professional body | professional care | professional development |
| professional experience | professional practice | professional status | professional support | professional training |
| professional work | prognostic factors | prognostic value | prospective study | prostate cancer |
| psychiatric hospital | psychiatric nurse | psychological distress | psychological empowerment | psychological symptoms |
| public health | pulmonary disease | qualitative analysis | qualitative approach | qualitative data |
| qualitative design | qualitative method | qualitative methodologies | qualitative research | qualitative study |
| quality appraisal | quality assessment | quality indicators | quality outcomes | quantitative data |
| random sample | rapid response | rating scale | recent decades | recent study |

| | | | | |
|--------------------------|--------------------------|---------------------|---------------------|-------------------------|
| rehabilitation clinic | relevant factors | replacement therapy | research article | research assistant |
| research ethics | research evidence | research fellow | research findings | research method |
| research methodology | research project | research topic | residential care | resonance imaging |
| respiratory depression | respiratory rate | retrospective study | risk assessment | risk factors |
| role model | room temperature | safety vests | sample size | sampling methods |
| scientific research | secondary care | secondary data | secondary education | secondary school |
| sectional study | sedentary behavior | service delivery | service providers | service user |
| sexual activity | sexual health | sexual intercourse | sexual orientation | shared decision |
| shared experience | shift work | side effects | significant amount | significant change |
| significant correlation | significant difference | significant effect | significant impact | significant improvement |
| significant increase | significant relationship | similar results | sleep behaviors | sleep conditions |
| sleep disorder | sleep disturbance | sleep duration | sleep patterns | sleep quality |
| small percentage | smooth muscle | social activity | social care | social contact |
| social function | social interaction | social organization | social status | social welfare |
| social worker | socioeconomic status | soft tissue | speech disabilities | spinal cord |
| sport activities | staff members | staff training | standard deviations | statistical analysis |
| statistical significance | stem cells | stress level | stroke patient | student nurse |
| study design | study findings | study group | study limitations | study methods |
| study participants | study period | study programme | study report | study results |
| study sample | substance use | support programme | surgical patients | surgical procedure |
| surgical treatment | survival rate | systematic study | target audience | team members |
| theoretical framework | time frame | time period | total income | total number |
| total score | tract infections | training course | training programme | training session |
| traumatic stress | treatment options | undergraduate nurse | university hospital | urinary catheter |
| urinary incontinence | urinary tract | usual care | vaginal delivery | vast number |

| | | | | |
|----------------------|---------------------|--------------------|--------------------|---------------|
| verbal communication | vital signs | vocational nurse | weight gain | weight loss |
| wide range | wide variation | work experience | work overload | work practice |
| working conditions | working environment | workplace violence | younger generation | |