Đelaludina Šukalić
Edina Rizvić-Eminović
Adnan Bujak

A CORPUS-BASED STUDY OF ANGLICISMS
ACROSS DIFFERENT TEXT TYPES OF ONLINE NEWS

Abstract

The study focuses on the analysis of anglicisms in the Bosnian/Croatian/Serbian (B/C/S) language on the basis of the compiled corpus of online news. The corpus includes two hundred and fifty online news articles, fifty per each of the five chosen text types, published during one month, the month of January 2022, on online news sites Klix.ba and ba.N1info.com. The chosen text types are: SciTech, News, Sport, Lifestyle, and Business. The task was to determine the extent of presence of anglicisms in these text types, that is, their high frequency of occurrence along with any differences across text types. The qualitative analysis of the aforementioned corpus includes investigation of unique anglicism tokens, total frequency of anglicism occurrence as well as part of speech classification of the noted anglicisms. The analysis reveals that anglicisms show different frequencies of occurrence across the analysed text types. The SciTech and Sport sections exhibit the highest frequency of anglicisms compared to the News section with the lowest anglicism frequency.

Key words: anglicisms, B/C/S, corpus, online news, text types.

Introduction

Lexical borrowing entails taking words from one language, the source language, and incorporating them into the lexicon of another language, the target language. Samardžija defines it simply as "every taking of a lexical unit from one lexicon into another" (Samardžija, 1995, p. 45). The process of lexical borrowing is influenced by a number of various factors such as cultural, economic, and political influences, territorial closeness of the areas employing the said languages, and more (Samardžija, 1995; Čedić, 2008, p. 22). In other words, apart from linguistic factors, e.g. structural similarities between languages under consideration, sociolinguistic factors play an important role as the driving force behind the process of lexical borrowing. Recently, the English language in particular has seized the forefront as the language of globalisation (Crystal, 2006; Čedić, 2008; Prćić, 2019, p. 17). Consequently, it rightly became the main source language for lexical borrowings. In this sense, its use within certain thematic domains, in particular the ones more susceptible to the aforementioned factors that facilitate lexical borrowing, is not surprising. For example, English, as a global Lingua Franca (Crystal, 2003), is commonly used as a linguistic answer for the new, emerging terminology, neologisms, and is therefore recognized as the predominant language in the technology field (Čedić, 2008, p. 23; Kaplan, 2001). In fact, even in general, anglicisms tend to emerge as the predominant type of lexical borrowing for new words (Muhvić-Dinamovski, 2005, p. 40).

Rosenhouse and Krowen (2008) list three key motives behind lexical borrowing in general: need to coin new terminology and concepts, tendency to emulate a dominant group, and tendency to create a special jargon in closed groups. The first motive behind lexical borrowing can be frequently reflected in the texts related to science and technology (e.g. *haker*), as there is

often a lack of equivalent words in the host language. The next motive of lexical borrowing can be seen in texts related to, for example, sport, as we might be influenced by the use of well-established anglicisms, borderline internationalisms (e.g. *baseball*). Lastly, anglicisms may appear in texts related to Lifestyle and Business, simply as more stylistic and fashionable choices (e.g. *outfit, look*), or typical lexical choices representing a special jargon, respectively (e.g. *menadžer, partner*).

It seems possible to anticipate a higher number of anglicisms in text types related to a particular subject matter. An investigation into various text types has an added benefit of revealing more about the potential motivation behind the inclusion of anglicisms in certain texts based on the noted differences. Furthermore, such research focusing on language contact and reasons or factors behind lexical borrowing is quite prolific in many languages, but not in B/C/S. Influence of English on many other languages, such as German, French and Spanish, has been heavily investigated and analysed (Gorlach, 2001; 2002; Chesley, 2010; Núñez Nogueroles, 2017). Research studies focused on a number of related issues, from the need to better understand the borrowing process and adaptation of the loan words, to the raised question of language purity and attitudes toward lexical borrowings. For instance, various works related to some of these questions already exist in Croatian and Serbian, respectively (e.g. Bosnar-Valković, et al., 2008; Prćić, 2019). The Croatian language seems to exhibit a considerable resistance towards the incorporation of new words (Čedić, 2008, p. 18). On the other hand, Prćić (2019, pp. 29-30) cautions that some English words are gradually taking over the existing words in the Serbian language, and that the influence of the English language has created problems with understanding the English-Serbian false friends. Furthermore, Prćić (2019, pp. 57-58) notes the existence of what he terms *AngloSerbian*, a type of language that increasingly deviates from the Serbian language, approaches the English language, and is abundant with anglicisms. While in both Croatian and Serbian the topic of anglicisms was mentioned early on and analysed over the years, with the results presented in a form of various dictionaries (e.g. Anić, et al., 2006; Klajn & Šipka, 2006), in B/C/S that is mostly not the case. Ibrahim Čedić's *Rječnik anglicizama u bosanskom jeziku* (2008) is one of the rare examples of a lexicographic and reference work that deals with anglicisms in the B/C/S language. Čedić (2008, p. 29) directly states that the issue of anglicisms in B/C/S, especially in terms of their meanings and orthographic rules, will definitely be important and will have to be dealt with in the future. Similarly, focusing on anglicisms in B/C/S, Šehović (2007, 2009, 2019) stresses the need for inclusion of anglicisms in the lexicographic sources of that language and recognizes the importance of anglicisms in terms of their stylistic potential for Bosnian speakers.

Given the fact that research on language contact and lexical borrowing between English and B/C/S is scarce, a study that would shed more light on the frequency of anglicism use in B/C/S along with their distribution across text types is needed. It creates possibilities for further research, especially in terms of predominant factors influencing the adoption of anglicisms into B/C/S and the history of contact between the two languages under consideration. Consequently, this paper is a result of the noted need for further, corpus-based research in the field of anglicisms in B/C/S. The focus on their dissemination and frequency across text types is particularly useful as similar studies investigating anglicisms in some other languages have uncovered interesting results indicative of text type differences (e.g. Núñez Nogueroles). Thus, the current research can be a step towards bridging the noticed informational gap in B/C/S as well. The main goal is to clearly demonstrate different frequency of anglicism occurrence across the analysed text types. An additional objective is to explore the use of anglicisms in terms of parts of speech as it is certainly

expected that nouns will be the predominant part of speech that can serve as an agent for further adaptation of borrowed words.

Theoretical background

The present paper explores anglicisms as a linguistic phenomenon within the framework of languages in contact. As such, anglicism can be defined as a word taken from the English language and used to define an object, idea or a concept as constituent parts of the English civilization, and while the said word does not have to be of English origin, it must be adapted according to the system and lexicon of the English language (Filipović, 1990, p. 17). It can also be defined as a word of English origin, but one already used in the target receptor language, after the adaptation to the rules of the said language (Opačić, 2012, p. 5).

In terms of the adaptation of the foreign lexical elements, it is possible to differentiate between orthographic, phonological, morphological, and semantic types of adaptation (Filipović, 1990, p. 24; Samardžija, 1995). The mentioned adaptation levels are immensely important as they can be indicative of the diachronic changes taking place and, in turn, tendencies of a particular host language, or even text type within the said language, to accept or resist the influx of foreign words. To illustrate, Mihaljević (2006, p. 42) notes how older computer terminology has undergone at least some level of adaptation, while modern terminology seems to rely more heavily on completely unadapted foreign words. Čedić (2008, p. 22) concludes that anglicisms are often used in their original form and will likely fully enter the target language, or in other words be fully adapted to the system of B/C/S, later on. Prćić (2005, p. 21) makes a distinction between primary and secondary adaptation, that is, adaptations and the words that enter the host language, and adaptations after the words have been accepted in the host language.

According to the mentioned levels of adaptation, there are varied approaches to not just the classification of anglicisms but also to the language borrowing in general. Filipović (1986, p. 38) proposes classification into three levels of lexical borrowing: switching code, interference, and integration. All three types can be described as moving on a scale starting from raw, foreign words to adapted ones. Examples within certain fields under observation that clearly demonstrate this process and various resulting types of borrowing, are plentiful. For example, in the field of technology, especially computing, it is possible to encounter examples such as *display vs displej*, or s*tick vs stik*, and in the field of sports, examples such as *nokaut*, a fully integrated word, and *triple-double* as a raw anglicism. It is also important to mention pseudo anglicisms as words that do not exist in the English language in the form that they employ in the target language but they are clearly influenced by English, such as *teniser*, through -er suffixation (Filipović, 1990, p. 19).

While it is important to note that lexical borrowings can be classified in a number of different ways and that there are various proposed classification systems, most classification systems in general contain at least tripartite classification into loan words, loan translations, and foreign words. To exemplify, Samardžija (1995) distinguishes between a number of categories including loan words, foreign words, borrowings, adaptations, and loan translations. Prćić (2005; 2019, pp. 122-123) makes a distinction between obvious, hidden, and raw anglicisms. In this classification, raw anglicisms are anglicisms that are directly adopted from English with no changes, e.g. *triple-double*. Obvious anglicisms are the ones showing multiple levels of adaptation to the host language such as *drajver*. Finally, there are also hidden anglicisms, that is, words created through the influence of English (e.g. through translation). As such, they are not easily recognized as anglicisms (e.g. *e-mail vs elektronska pošta)*. Čedić (2008, p. 4) differentiates between two types of foreign words based on their phonological and morphological characteristics

and those are direct loan words and lexical calques. Čedić (2008, p. 26) also notes that some of the anglicisms are used as original foreign words, some are gradually adapted, and some have been adapted completely and have become productive in the target language. Therefore, based on the levels of adaptation, a distinction can be made between a completely foreign word, one that has been completely adapted according to the system of the host language, and numerous types in between the two extremes, depending on the employed classification system.

As the paper focuses on anglicisms across text types of online news, it is also necessary to elucidate upon the language of online news, and the two chosen corpus sources. Firstly, the corpus choice was influenced by the nature of online media as they already have an established link to the process of dissemination of modern vocabulary and as the Internet allows the English language community to exert its cultural influence (Rosenhouse & Kowen, 2008).

Fundamentally, the Internet allows quick and easy access to a great deal of different information, including news in general. In fact, the Internet is quickly becoming the predominant news source as news sites experience considerable site traffic (Newman, et al., 2021). Online news websites are naturally one of the primary sources of news online, which is reflected in the number of monthly visitors, e.g. the BBC website had 1.1 billion visits during the month of December 2021 (Majid, Pressgazette, 2022). Hence, analysis of online text types in terms of the employed language, is definitely a productive avenue for further research. That fact coupled with the one that from the very beginning of the internet development the English language was, and still is, the predominantly used language online (Young, 2013), increases the influence of English on the online content in other languages as well. Consequently, studying anglicisms across online text types can be a great contribution to this research area.

The general language of online digital journalism is characterised by a set of unique features including hypertextuality, multimediality and interactivity (Deuze, 2001). Online news websites usually demonstrate internal hypertextuality, as they offer connections to other news on the website. In terms of multimediality, they often include videos and photos along with the text. Finally, when it comes to interactivity, online news websites are mainly navigational (Deuze, 2001). They allow their users to navigate the web page and news articles easier (e.g. go to the top). All of the aforementioned characteristics are, of course, evident in the online news from the chosen corpus sources, Klix.ba[1] and ba.N1info.com[2].

Methodology

The corpus used for the purpose of this research paper was created using fifty news articles from each of the five sections of the prominent online news sources in Bosnia and Herzegovina. The online sources used are Klix (https://www.klix.ba/) and N1info (https://ba.n1info.com/). Fifty of the news articles belonging to the text types of SciTech, News, Lifestyle, Business and Sport, published during the month of January, were copied to create the corpus. It is important to note that for some of the sections on the N1info website, it was also necessary to include articles from the very end of December 2021, as there were not enough published articles during the month of January. For the purpose of this research, texts were assigned to five text types according to the

---

[1] Klix, available at https://klix.ba, is a leading online news website in Bosnia and Herzegovina with an average of 550.000 monthly visitors (Klix.ba, n.d.). The website interface allows users to choose between several sections of news: News, Business, Sport, Magazine, Lifestyle, Scitech, Auto.

[2] N1info, available at https://ba.n1info.com/, is a 24/7 online platform for dissemination of news with almost 400.000 unique visitors over a ten day period (Sekulić, 2020). Similar to Klix, their website also contains various news sections such as Business, Culture, Science etc.

thematic domains: News, Business, Sport, Lifestyle and SciTech. Each one of the text types includes fifty news articles from the related thematic domains. For example, the News section includes articles related to International, Regional, and Local News. Business section focuses on finance and economy. The Sports section covers various types of sports including basketball, football, tennis, etc. The Lifestyle section covers texts on travel, food, pets, home, relationships, fashion and beauty. SciTech, of course, deals with texts in the domain of science and technology. The corpus was compiled in the same way for all of the thematic domains and that is by copying fifty news articles published chronologically within the target text type. For all the articles, titles and by lines were also noted. Within the corpus, all existing images or inserted video clips were removed along with their captions. Proper names and names of various companies, organisations, and so on, were also excluded as they do not belong within the classical classification of anglicisms (Prćić, 2019, p. 123).

Following the compilation, the texts were analysed to first note and then quantitatively determine the number of anglicisms they contain. The noted anglicisms were confirmed with the help of *Rječnik anglicizama u bosanskom jeziku* by Ibrahim Čedić (2008), as well as *Rječnik stranih riječi hrvatskog jezika* by Anić et al. (2002), where necessary. Čedić's (2008) Dictionary did not contain as many examples of anglicisms as the other dictionary so, at times, it was necessary to refer to this second source. The first aim was to determine the overall total number of anglicisms in the compiled corpus per each of the five text type sections, along with the ratio of anglicisms per number of words. Therefore, after all the anglicisms were noted, the number of anglicisms within sections was calculated along with the total number of words within those sections. The following quantitative analysis included calculating ratios of total and unique anglicisms per each of the five text type sections. The ratio, that is the percentage of anglicisms, was calculated by dividing the total numbers (tokens) of unique and total anglicisms with the total number of words and then multiplying the result by 100. The same was repeated for individual sections and sources. After the initial quantitative analysis, anglicisms were counted again with only the very first instance of an anglicism within the individual texts being noted. In other words, repeated examples of anglicisms within the same article were not counted. This was followed by a further quantitative analysis in terms of ratios of these anglicisms per number of words, across all five text type sections as well. This allows us to see if there are any differences with the inclusion of repeated anglicisms with respect to the total frequency ratios across the individual text type sections. It is important to note that it was necessary to determine not only the number of anglicisms but also the ratio of anglicisms per number of words as the chosen text type samples were not of the same length. Each of the unique anglicism tokens was also analysed in terms of its parts of speech type and labelled as belonging to nouns, verbs, adjectives, adverbs or other parts of speech.

Results and discussion

The total number of unique anglicisms, that is, anglicism tokens, along with the ratio of anglicisms across text types is presented below in Table 1 for Klix, and Table 2 for N1info.

Text type	Words	Number of anglicisms	Ratio of anglicisms
SciTech	12,604	114	0.90%
Lifestyle	15,634	77	0.49%
Sport	12,046	62	0.51%

Business	19,049	52	0.27%
News	13,940	37	0.27%

Table 1. Number and ratio of Anglicisms with respect to article's subject matter for Klix.

Text type	Words	Number of anglicisms	Ratio of anglicisms
SciTech	17,243	81	0.47%
Lifestyle	15,148	44	0.29%
Sport	10,652	58	0.54%
Business	21,928	69	0.31%
News	14,492	36	0.25%

Table 2. Number and ratio of Anglicisms with respect to article's subject matter for N1info.

As Table 1 demonstrates, the SciTech section for Klix contains the highest total frequency of unique anglicism tokens (0.90%) with the Sport section following it (0.51%). The Lifestyle section follows closely behind (0.49%) while the Business and News sections contain the lowest total frequency of anglicisms with 0.27% each. For N1info, as Table 2 indicates, the Sport section contains the highest total frequency of unique anglicism tokens (0.54%) with the SciTech section following closely behind (0.47%). The third place is occupied by the Business section (0.31%), while the Lifestyle and News sections show the lowest frequencies with 0.29% and 0.25% respectively.

Overall, both sources show similar total frequency percentages with the most noticeable difference being the high frequency of anglicisms within the Klix SciTech section. This can potentially be explained by the nature of the text sources. N1info SciTech texts often contain articles from other sources, mostly Croatian ones, so the number of anglicisms is lower as those texts use Croatian words instead of anglicisms (e.g. *link vs poveznica*). Therefore, this could explain why the Sport section occupies the first place for N1info compared to the SciTech section for Klix.

Same can be said for the difference between the order of Lifestyle and Business sections. N1info again leads with the Business section as opposed to the Lifestyle section for Klix, and once again, these texts are often taken from external sources, as indicated within bylines or even directly within the texts. Those text types that show fewer differences across the two corpus sources, such as News or Sport are predominantly written directly for N1info as their bylines indicate. This is a great insight into the response of the Croatian language, compared to Bosnian, towards the incorporation and use of anglicisms.

Interestingly, both sources indicate higher frequencies within the expected text types of SciTech and Sport sections, along with the lowest frequencies within the News section. As the News sections deal with mostly local and regional information, it is expected that the number of anglicisms is lower which clearly indicates, along with the other higher frequencies within other sections, that text types can indeed be indicative of the frequency of anglicism occurrence.

As suggested earlier, these differences between text types in terms of anglicism occurrence and distribution, can be linked to the reasons behind the inclusion of anglicisms and factors facilitating their inclusion. For example, the Sport section, due to the nature of this particular text type, is easily influenced by most factors of the inclusion of anglicisms such as territorial or cultural ones, due to the global nature of sports and their uniformed rules. The SciTech section,

of course, incorporates numerous neologisms so there is a need to name new things and concepts, along with a tendency to create specialised jargon per lexical borrowing motives suggested earlier.

Tables 3 and 4 show the total number of unique anglicisms within individual news articles, along with the total number of all, even repeated anglicisms. The results are also presented in the form of ratios of anglicisms per number of words.

Text type	Words	Unique anglicisms within individual texts	Total number of anglicisms	Percentage of unique anglicisms	Percentage of total anglicisms
Sport	12,046	182	377	1.51%	3.13%
SciTech	12,604	187	359	1.48%	2.85%
Lifestyle	15,634	124	201	0.79%	1.29%
Business	19,049	109	241	0.57%	1.27%
News	13,940	49	76	0.35%	0.55%

Table 3. Total number and ratio of Anglicisms with respect to the article's subject matter for Klix.

Text type	Words	Unique anglicisms within individual texts	Total number of anglicisms	Percentage of unique anglicisms	Percentage of repeated anglicisms
Sport	10,652	167	290	1.57%	2.72%
SciTech	17,243	151	287	0.88%	1.66%
Lifestyle	15,148	55	79	0.36%	0.52%
Business	21,928	105	184	0.48%	0.84%
News	14,492	51	64	0.35%	0.44%

Table 4. Total number and ratio of Anglicisms with respect to the article's subject matter for N1info.

When analysing the total number and ratio of anglicisms across text types, as shown in Table 4, it is possible to note that not much changes for N1info apart from the differences between text types becoming even more prominent. For example, the News section shows indeed the very lowest frequency percentage with 0.44% compared to the Sport section with 2.72%, which is more than six times as much. For Klix, as Table 3 indicates, the results do change, that is, they become even closer to the results for N1info. For example, with the inclusion of the total number of anglicisms, the Sport section of Klix contains the highest ratio of anglicisms and the same is true for N1info. Even the difference between the Lifestyle and Business sections becomes less significant and more in line with the results for N1info with the earlier frequency being 0.49% and 0.27% compared to 1.29% and 1.27%, respectively. Once again, the Sport section far outnumbers the News section with respect to the number and frequency of anglicisms with almost six times more anglicisms in the Sport section (3.13%) compared to the News section (0.55%).

The consistency of the results across the two sources of corpus material reaffirms that certain text types do contain different frequencies of anglicisms. With respect to the noted frequent repetition of anglicisms within certain text types, once again, the reasons can be found in the nature of the said text types and the way that they are influenced by the factors facilitating the process of lexical borrowing as well as the reasons behind it. For example, as the News text types deal with such diverse topics, there is less possibility and need for the repetitive use of the same anglicisms.

When we compare this section to Sports, it is possible to see the discrepancy even more clearly. For example, for the News section of N1info, there are eleven examples of repeated anglicisms as opposed to 123 for the Sport section, which is more than ten times as many. To illustrate, anglicism *klub* is used extremely frequently (48 times Klix, 51 times N1info) meaning that the total number of tokens of this single anglicism within the Sport sections almost exceeds the total number of all unique anglicisms within the News section.

Lastly, Table 5 and Table 6 below show the corresponding parts of speech of anglicisms across the analysed text types for Klix and N1info, respectively.

Part of speech	Number of tokens				
	SciTech	Lifestyle	Sport	Business	News
Nouns	98	67	49	41	30
Adjectives	14	9	7	10	5
Verbs	2	1	6	2	2
Adverbs	0	0	0	0	0
Other POS	0	0	0	0	0

Table 5. Number of Anglicisms according to part of speech for Klix.

Part of speech	Number of tokens				
	SciTech	Lifestyle	Sport	Business	News
Nouns	66	33	49	57	29
Adjectives	11	7	7	10	6
Verbs	4	4	2	2	1
Adverbs	0	0	0	0	0
Other POS	0	0	0	0	0

Table 6. Number of Anglicisms according to part of speech for N1info.

As can be seen, the predominant part of speech of anglicisms for both Klix and N1info is, expectedly, nouns. Nouns are followed by significantly fewer adjectives, and just a few verbs while there are no adverbs or any other parts of speech. Considering that lexical borrowings are

predominantly used to refer to the new concepts, ideas, and things, it is expected that the nouns will be the most common part of speech of anglicisms. Interestingly, the existence of other types of speech can also be indicative of the overall anglicisms saturation within the text type as those text types containing higher frequency also show greater numbers of other parts of speech compared to text types of lower frequency that mainly contain nouns. This can perhaps be an added evidence as to the level of adaptation of the anglicisms within the various text type as common and repeated use of anglicisms leads to their adaptation at various levels, e.g. morphological level, and, through that, formulation of other types of parts of speech of the anglicisms in question (e.g. *internet* and *internetski, or šut* and *šutirati*).

Conclusion

A number of factors led the English language to have a tremendously influential position and, through that, heightened its effects and influence on the other languages as well. This resulted in the continuous process of lexical borrowing and, therefore, to English being one of the leading sources of the said borrowing in other languages. Furthermore, as new discoveries and developments around the world lead to the need of neologisms, it is expected that the English language will act as the predominantly used source language yet again. Taking this into consideration, investigations as to the current state of anglicisms, and any future developments related to the process of lexical borrowing, are immensely important.

This paper tried to offer some of the answers to the existing questions related to the existence and use of anglicisms in B/C/S. It contributes to the noted lack of research of this particular type within B/C/S, especially compared to other languages. The investigation into the frequency of anglicisms and their distribution across text types can have several further applications. As the results revealed, certain text types, such as SciTech, are expected to contain a higher number of anglicisms compared to some other text types. Therefore, knowing which areas to focus on can help linguists in a number of different ways, from creating more comprehensive, or even specialised, dictionaries of anglicisms, to dealing with the overwhelming number of anglicisms being adapted into the recipient language, especially from the standpoint of language purism. This is particularly important from the perspective of offering possible alternative host words, and maintaining the host language as much as possible. To exemplify and reiterate, the Croatian language has proven to be quite successful in this sense (e.g. *preglednik, datoteka*), while B/C/S seems to prefer anglicisms (e.g. *browser, fajl*).

A comparative study between exclusively online text types and more traditional text types that originate offline is just one of the possible avenues for further research. Diachronic studies of anglicisms use and frequency over time can also be helpful. The frequency of anglicisms can be reflective of the ongoing influence of the English language and can provide insights as to the history of language contact between English and B/C/S.

To conclude, further research of anglicisms in the B/C/S language represents an urgent task for several aforementioned reasons including the important endeavour of dictionary creation, resolution of the question of linguistic purism and so forth. This paper offers an insight as to the current use, frequency and distribution of anglicisms across text types of online news, and as such, represents a necessary step towards further needed research in this area.

References

1. Anić, Š., Klaić, N., & Domović, Ž. (2002). *Rječnik stranih riječi: tuđice, posuđenice, izrazi, kratice i fraze.* Sani-plus. Zagreb.

2. Bosnar-Valković, B., Blažević, N., & Gjuran-Coha, A. (2008). The Americanization of the Croatian Language. *Coll. Antropol. 32*(4), pp. 1035–1042.

3. Chesley, P. (2010). Lexical borrowings in French: Anglicisms as a separate phenomenon. *Journal of French Language Studies, 20*(3), pp. 231-251.

4. Crystal, D. (2003). *English as a global language*. Cambridge: Cambridge University Press.

5. Crystal, D. (2006). Into the twenty-first Century. In L. Muggustone (Ed.), *The Oxford History of English.* New York: Oxford university Press.

6. Deuze, M. (2001). Online Journalism: Modelling the First Generation of News Media on the World Wide Web. *First Monday, 6*(10).

7. Čedić, I. (2008). *Rječnik anglicizama u bosanskom jeziku*. Institut za jezik u Sarajevu.

8. Filipović, R. (1986). *Teorija jezika u kontaktu*. Zagreb, Školska knjiga.

9. Filipović, R. (1990). *Anglicizmi u hrvatskom ili srpskom jeziku: porijeklo – razvoj – značenje*. Zagreb: Jugoslavenska akademija znanosti i umjetnosti.

10. Görlach, M. (ed.) (2001). *A Dictionary of European Anglicisms. A Usage Dictionary of Anglicisms in Sixteen European Languages*. Oxford: Oxford University Press.

11. Görlach, M. (2002). *English in Europe.* Oxford: Clarendon Press.

12. Kaplan, R. B. (2001). The dominance of English as a language of science: Effects on other languages and language communities. In U.Ammon (Ed.), *The dominance of English as a language of science: Effects on other languages and language communities* (pp. 3–26). New York: Mouton de Gruyter.

13. Klajn, I., & Šipka, M. (2006). *Veliki rečnik stranih reči i izraza.* Prometej. Novi sad.

14. Klix.ba. (n.d.). Retrieved January, 2022, from https://www.klix.ba/.

15. Majid, A. (2022, January 17). December: a better month for many global news sites, top 50 ranking shows. *Pressgazette.* Retrieved January 28, 2022, from https://pressgazette.co.uk/most-popular-websites-news-world-monthly/.

16. Mihaljević, M. (2006). Hrvatsko i englesko računalno nazivlje. *Jezik, 53*(2), pp. 41-50.

17. Muhvić-Dimanovski, V. (2005). *Neologizmi: problemi teorije i primjene.* Zavod za lingvistiku Filozofskoga fakulteta Sveučilišta u Zagrebu.

18. ba.N1info.com. (n.d.). Retrieved January, 2022, from https://ba.n1info.com/.

19. Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C., & Nielsen, R. K. (2021). Reuters Institute Digital News Report 2021. Reuters Institute for the Study of Journalism

20. Núñez Nogueroles, E. (2017). The Use of Anglicisms in Various Thematic Fields: An Analysis Based on the Corpus de Referencia del Español Actual. *Anglica. An International Journal of English Studies. 26*(2). pp. 123-149.

21. Opačić, N. (2012). *Hrvatski ni u zagradama.* Zagreb: Hrvatska sveučilišna naklada.

22. Prćić, T. (2005). *Engleski u srpskom*. Novi Sad: Zmaj.

23. Prćić, T. (2019). *Engleski u srpskom*. Filozofski fakultet, Novi Sad. online edition. http://digitalna.ff.uns.ac.rs/sadrzaj/2019/978-86-6065-512-9 .

24. Rosenhouse, J., & Kowner, R. (2008). *Globally Speaking: Motives for Adopting English Vocabulary in Other Languages.* Bristol, Blue Ridge Summit: Multilingual Matters.

25. Samardžija, M. (1995). *Leksikologija s poviješću hrvatskoga jezika*. Zagreb: Školska knjiga.

26. Sekulić, T. (2020). *Hvala vam na povjerenju u N1!* ba.N1info.com. March 23, 2020, retrieved January, 31, 2022, from https://ba.n1info.com/vijesti/a419121-hvala-vam-na-povjerenju-u-n1/.

27. Šehović, A. (2007). Neologizmi u razgovornome stilu bosanskog jezika. *Strani jezici. 36*(1). pp. 21-29.

28. Šehović, A. (2009). Mjesto i funkcija anglicizama u savremenom bosanskom jeziku. *Pismo, časopis za jezik i književnost. 7*(1), pp. 122-138.

29. Šehović, A. (2019). Anglicizmi : motivatori u tvorbi glagolskih izvedenica u savremenom bosanskom jeziku. *Bosanskohercegovački slavistički kongres. 2*(1), pp. 229–242.

30. Young, Holly. (2013). The digital language divide. *The Guardian.* Retrieved January 30, 2022, from http://labs.theguardian.com/digital-language-divide/.

KORPUSNA ANALIZA ANGLICIZAMA
U RAZLIČITIM VRSTAMA TEKSTOVA ONLINE VIJESTI

Sažetak

Ovaj rad se fokusira na analizu anglicizama u bosanskom/hrvatskom/srpskom (B/H/S) jeziku na osnovu prikupljene građe korpusa online vijesti. Korpus uključuje dvije stotine i pedeset online novinskih članaka, pedeset za svaku izabranu rubriku, objavljenih tijekom jednog mjeseca, mjeseca januara 2022. godine, na online stranicama Klix.ba i ba.N1info.com. Izabrane rubrike tekstova su: Nauka i tehnologija, Vijesti, Sport, Lifestyle i Biznis. Zadatak je bio utvrditi nivo prisustva anglicizama u ovim rubrikama, tj. njihovu veliku učestalost pojavljivanja, kao i bilo kakve razlike po vrstama tekstova. Kvalitativna analiza ranije pomenutog korpusa uključuje istraživanje jedinstvenih primjera anglicizama, ukupnu učestalost pojave anglicizama kao i podjelu po vrstama riječi zabilježenih anglicizama. Analiza otkriva da anglicizmi pokazuju različitu učestalost pojave kroz analizirane rubrike vrsta tekstova. Rubrike Nauke i tehnologije te Sporta pokazuju veću učestalost anglicizama u poređenju sa Vijestima gdje je učestalost anglicizama najniža.

Ključne riječi*: anglicizmi, B/H/S, korpus, online vijesti, vrste tekstova.*