

MOST FREQUENT MULTI-WORD EXPRESSIONS IN ENGLISH FOR BANKING: A CORPUS-BASED DIACHRONIC STUDY

Borislav **Marušić**, University of Applied Sciences "Lavoslav Ružička" in Vukovar, bmarusic@vevu.hr

Original scientific paper
DOI: 10.31902/fil.46.2023.3
UDC: 811.111:336.71

Abstract: The aim of the paper is to perform diachronic research in one segment of Business English – English for Banking to check the diachronic stability of its vocabulary. Therefore, an analysis of multi-word expressions (MWEs) in 15 syntactic categories in annual reports of four biggest US banks for fiscal years 2006 and 2021 was made. More precisely, the diachronic changes comprising the span of time of 15 years were analysed in both the frequency of the occurrence of MWEs across the 15 analysed syntactic categories and in their percentage relationship to the scope of the corpora. The corpus analysis made it possible to extract real language in use, enabling the author to determine diachronic changes in vocabulary across categories, thus bridging the gap between the existing knowledge of vocabulary in English for Banking and the current situation. The research showed that there is a high statistical significance of changes in both the frequency of the occurrence of MWEs across the 15 analysed categories and in their percentage relationship to the scope of the corpus in the period over 15 years. The research findings might be useful for both Business English lecturers and their students, due to the various possible pedagogical implications of the results. The research might function as an impetus for linguists doing research in corpus linguistics to undertake further research in other areas of Business English to determine and describe possible changes in the language used. These areas could include: international trade, logistics, wholesale, retailing, industry, business media, leadership, management, marketing or investing.

Keywords: ESP, Business English, English for banking, vocabulary, teaching BE at a tertiary level

1. Introduction

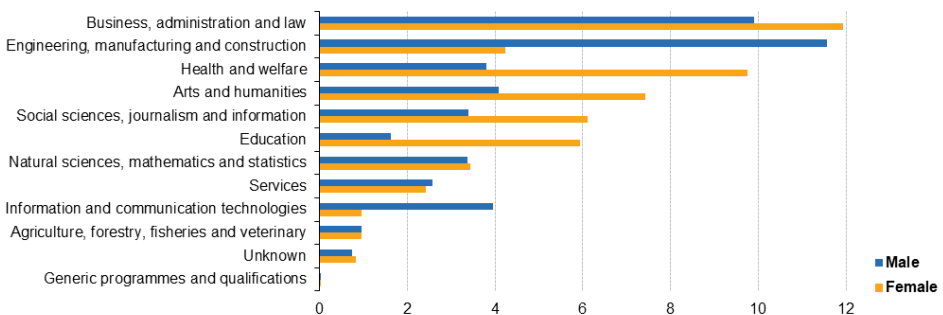
The expansion of English as a *lingua franca* and a language of globalization gathered speed after the WW2, when the USA arose as the world's only superpower. One of the important factors in becoming a global superpower was the status of U.S. Dollar as the world's reserve currency, which status it achieved by the Bretton Woods Agreement in 1944. The Congressional Research Service that informs the US Congress Members the legislative debate since 1914 states in the text "The U.S. Dollar as the World's Dominant Reserve Currency" from

15th September 2022 the following advantages of having a world's reserve currency:

Because many central banks and financial institutions around the world want to hold U.S. dollars and dollar-backed securities like U.S. Treasury bonds, there is strong demand for U.S. dollars. That demand, in turn, allows the United States to borrow more cheaply (at lower interest rates) than it would otherwise. Strong demand for dollars also allows the U.S. government, firms, and consumers to borrow from foreign creditors in dollars rather than foreign currencies. As a result, the value of that debt does not depend on fluctuations in exchange rates. When other governments, firms, and individuals borrow in foreign currencies, they incur the risk that swings in exchange rates will cause their real debt level (the size of the debt in the borrower's national currency) to increase, potentially quickly and significantly. U.S. firms and consumers also benefit by saving on transaction costs.

Consequently, the importance of Business English (BE) standing for language of communication in international business has grown even more than that of General English (GE) in the several recent decades. According to the Eurostat, Business English has achieved the undisputable position Nr. 1 among all Languages for Specific Purposes, since business is omnipresent and taught, learnt and used much more than any other major in the European Union:

Distribution of tertiary education students by broad field and sex, EU, 2020
(%)

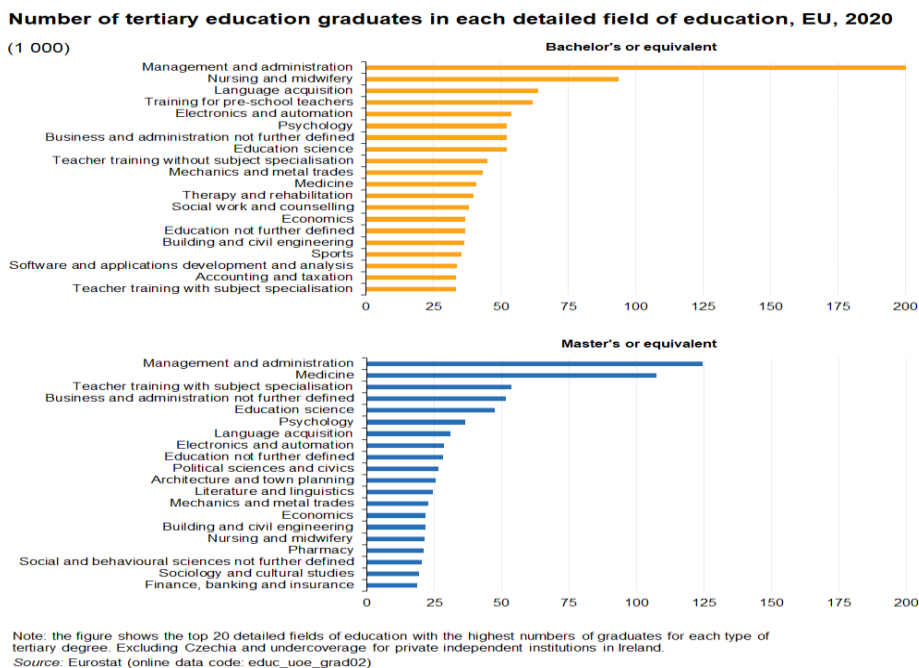


Note: ranked on the total (male and female) share of students in each broad field.
Source: Eurostat (online data code: educ_uae_enrt03)

eurostat

Graph 1.

The research reveals that in 2020, 21.8 % of tertiary education students in the EU were studying business, administration or law vs. 15.8 % of students studying the second frequent group of majors including engineering, manufacturing and construction. More clear picture can be obtained from the same research depicting the number of graduates:



eurostat

Graph 2.

As it can be seen from the Graph 2., there are around 200 graduated students in management and administration plus more than 50 graduated students in business and administration not further defined out of each 1.000 students in the EU Member States. These numbers reveal the dominant number of students learning BE reflecting the need of BE users to communicate and gain information when doing business intentionally. Banking is one of the important topics in BE classes and it is therefore useful to analyse its language in the authentic environment. One of the most efficient scientific methods in this regard is corpus analysis.

Annual reports of the biggest US banking corporations for fiscal years 2006 and 2021 – *JP Morgan Chase & Co.*, *Wells Fargo & Co.*, *Citigroup Inc.* and *U.S. Bancorp* were used as two corpora in the research. The diacronic study enabled the author to identify the most frequent MWEs in 15 syntactic categories and

compare the most frequent from 2006 and 2021 in each category. The results gave insights into the diachronic position of MWEs in this segment of BE. The results of the research are of practical importance for BE lecturers, students with majors in business and economics and all those people dealing with financials around the world with their mother tongue other than English. Finally, the corpus findings can be used to compile lists of phrases suitable for teaching and practising on various exercises at faculties and colleges for business and economics.

2. Theoretical background

2.1. Financial sector and the American banking system

The financial sector represents one of the biggest sectors under the 11 S&P 500 sectors: Communication Services, Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Information Technology, Materials, Real Estate and Utilities. The importance of the sector derives from the fact that banks finance other businesses and thus enable and generate growth of the whole economy.

S&P 500 sector	Operating profit margins in percent	S&P 500 sector	Operating profit margins in percent
Information technology	22	S&P 500 total	11.9
Real estate	21.8	Utilities	11
Financials	15.4	Healthcare	9.1
Communication services	15.3	Industrials	8.8
Energy	13.4	Consumer staples	6.7
Materials	12.5	Consumer discretionary	6.5

Table 1.

The Financials sector is very profitable too. Table 1. gives the insight into operating profit margins across the S&P 500 sectors¹. Commercial banks function as a part within the wider financial system. Between 1933 and 1999 as a consequence of NYSE crash in 1929, American banks were clearly divided into two large sectors: the commercial one and the investment one. The division was

¹ <https://www.yardeni.com/pub/sp500margin.pdf>

made by the Glass-Steagall Act², passed in 1933 which prohibited the commercial banks to conduct the investment business in order to prevent repetition of events such as the 1929 stock market crash. The mix of commercial and investment banking activities was regarded as a potential risk of high degree and as a possible trigger for prospective stock market crisis. However, during the Clinton Administration in 1999, it was annulled by the Gramm-Leach Bliley Act³ due to concerns that such division was making American industry less competitive on a global scale. Moreover, it was thought that lifting the division between the sectors would reduce risks of running banking operations. Since then, according to investopedia.com⁴ the commercial banks have been taking on the following business roles: accepting deposits, lending money, processing payments, issuing bank drafts and checks and offering safety deposit boxes for items and documents. The chapter on the roles of commercial banks in economy ends with the statement that "...the common theme among these activities is that they are aimed at providing a financial service to an individual or business". As such the banks are a very important part of wider financials and the economy as a whole. This gives the reason why English for banking is a very important part of BE curricula at a tertiary level of education.

2.2. English for banking

When it comes to *English for banking*, it is a part of almost each BE curriculum at a tertiary level of business education. This fact can be confirmed by checking the contents pages of BE textbooks for courses at intermediate and upper-intermediate level by British publishers Oxford University Press, Cambridge University Press and Longman Pearson. These textbooks are used as mandatory (or optional) teaching materials in many BE curricula at a tertiary level.

Publisher	Title of the textbook	Intermediate level	Upper-intermediate level
OUP ⁵	Business Result	Performance	Performance
	International Express	-	-

² <https://www.goldmansachs.com/our-firm/history/moments/1933-glass-steagall-act.html>

³ <https://www.govinfo.gov/content/pkg/PLAW-106publ102/pdf/PLAW-106publ102.pdf>

⁴ <https://www.investopedia.com/articles/investing/062513/role-commercial-banks-economy.asp>

⁵ See https://elt.oup.com/learning_resources/subjects/businessenglish/?cc=global&selLanguage=en&mode=hu

Cambridge University Press ⁶	English for Business Studies	Finance Banking	Finance Banking
	Business Advantage	Accounting Finance	Accounting Microfinance
	New International Business English	-	Money matters
Longman Pearson ⁷	New Insights into Business	Banking Company Performance	no upper-intermediate textbook
	Business Leader	Money	Raising Finance
	Business Partner	-	Finance Performance

Table 2.

There is an overview over banking topics in the most frequently used textbooks in BE classes given in the Table 2. There is at least one lesson / unit committed to *English for banking* at intermediate and upper-intermediate level in each textbook with the exception of *International Express* published by the Oxford University Press. Moreover, having full awareness of the importance of financial sector in the modern economy and BE classes, the Cambridge University Press published *Professional English in Use – Finance* within its well-known series *English in Use*, as well as *English for the Financial Sector*, a material covering "a wide range of financial topics, including retail and investment banking, accounting, trade finance, and mergers and acquisitions"⁸. Although there are many textbooks published by big publishers, Stojković (2023) states that they "produce material that is of a specific purpose, yet far too general at the same time." As main problems, she mentions such books almost never reaching beyond intermediate level and the segmentation of texts and exercises not allowing "the overall aim of university studies – to study thoroughly." (ibid.) That is why corpus analysis can be used as a source of additional teaching materials, that can be created according to individual needs of BE lecturers and their students.

There are many ways how the finance industry has changed over more than a decade. According to the Graft on Banks Finance⁹ there have been following

⁶ See <https://www.cambridge.org/is/cambridgeenglish/catalog/business-professional-and-vocational>

⁷ See <https://www.pearson.com/english/catalogue/business-english.html>

⁸ See <https://www.cambridge.es/en/catalogue/business-english/other-titles/english-for-the-financial-sector>

⁹ <https://graftonbanksfinance.co.uk/news/news.php?news=70>

changes in the banking sector: financial services have become stronger regulated and supervised after the financial crisis, trust in finance industry has newly been built mostly in "...organisations that operate with transparency and strong values at the forefront", gender and ethnic diversity have been promoted, as well as inclusivity and diversity. Moreover, "...this surge in new technology has shaken the finance industry up, breaking down the barriers for new competition to enter the market and made some of the bigger players sit up and listen. The explosion of new technology and availability of mobile functionality means it has become an expectation from most consumers today." These changes, particularly the ones regarding new technology and its use, within the finance industry could have furthered potential changes in its vocabulary too. The analysis of expected diachronic changes in the language of Annual Reports of four biggest American banks was the initial intention of this scientific research.

An annual report is according to *the Longman Business English Dictionary* (2007) "...a report presented each year by the directors of the company to the members and shareholders, containing financial information about the company's trading activities and the documents the company must produce by law, which are the BALANCE SHEET, the PROFIT AND LOSS ACCOUNT and the AUDITOR'S and DIRECTORS' REPORT". *The Oxford Learner's Pocket Dictionary of Business English* (2015) gives this brief definition of an annual report "...a financial report that a company must by law present each year to its shareholders". From the linguistic standpoint, an annual report is not only a complex financial document containing business information on all aspects of a company's performance in a given fiscal year, but also written material appropriate for linguistic analysis. The results of such analysis can be applied and used in BE classes.

2.3. Corpus-based studies

Corpus-based studies enable analysis of the use of language in context. They have come into foreground of linguistic research having huge research potential. Stojković (2023) claims that „The potential of corpora is such that Conrad (2000) spoke of them as a means that will thoroughly change the teaching of foreign languages and the overall language education, to include both what is taught and how it is taught“. As such, they are the best way to explore the functionality of language on lexical, semantic and syntactic levels. Mockienė (2023) names lexis as the most changeable level of the language: "New words constantly appear to name new phenomena; thus, language change is never ending (Meyer 2002)". The most comprehensive insight into the history and importance of corpus-based studies in contemporary linguistics gives Bennett (2020), so that for the purpose of the theoretical background, we convey his whole information: "The

principles of corpus linguistics have been made around for almost a century. Lexicographers, or dictionary makers, have been collecting examples of language in use to help accurately define words since at least the late 19th century. Before computers, these examples of language were essentially collected on small slips of paper and organized in pigeon holes. The advent of computers led to the creation of what we consider to be modern-day corpora. The first computer-based corpus, the Brown corpus, was created in 1961 and comprised about 1 million words. Today, generalized corpora are hundreds of millions of words in size, and corpus linguistics is making outstanding contributions to the fields of second language research and teaching“. Additionally, Römer (2008) states “that corpus resources and methods have a great potential to improve pedagogical practice and that corpora can be used in a number of ways, indirectly to inform teaching materials and reference works, or directly as language learning tools and repositories for the design of data-intensive teaching activities“.

Custom-made corpora research in LSP pose a great challenge for lecturers as such research enable analysis of authentic LSP and consequently creation of teaching materials for their classes. It is the so-called *indirect* way of their use according to Römer (2011). The researchers prepare a custom-made corpus, analyse it and incorporate the corpus findings into their classes. When it comes to the *direct* way of using custom-made corpora, it is primarily used in the process of *data-driven learning (DDL)*. The use of data-driven learning in LSP is confined in at least two ways: it can be done only with students at higher levels of language knowledge and the students should be technically skilled and trained for such learning. Đurović & Bauk (2022) conducted a research in order “to test the possible application of contemporary corpus linguistics methods through the Data-driven learning method in the process of teaching/learning teaching vocabulary. The method applied exposed the students to authentic technical material and provided them with a specifically developed word list from the corpus of selected ship instruction books and manuals“. According to the authors, the students showed positive reactions regarding experimental and innovative teaching models. Anđelković et al. (2018) state that “In recent years, corpus-based studies have recently become the core areas of LSP research. LSP and professional communication have benefited from both direct and indirect applications of specialized corpora. For LSP teachers and learners, the most obvious advantage of the use of specialized corpora is that they provide authentic examples of specialized language, which are, in turn, beneficial in developing reading and writing skills and in understanding and producing particular texts and types of texts (Aston, 2001; Flowerdew, 2001; Connor & Upton, 2004; Römer, 2011)“. According to the same authors Anđelković et al. (2018:12) “Several studies have also reported the effectiveness of using corpora for creating LSP teaching materials (indirect application of corpora in LSP teaching and learning), especially those aimed at teaching and learning

collocations (Howarth, 1998), lexical acquisition (Cobb, 1999), writing and grammar (Tribble, 2001)“.

In addition to various corpora related to general and specific areas of English such as Sketch Engine containing “600 ready-to-use corpora in 90+ languages, each having a size of up to 60 billion words to provide a truly representative sample of language¹⁰“, the British National Corpus (BNC) “a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of British English, both spoken and written, from the late twentieth century¹¹“ and one of the most widely-used English corpora – Corpus of Contemporary American English (COCA) – containing more than 1 billion words¹², there are numerous corpus analysis tools for custom-made corpora available online. They are convenient for large-scale analysis of the compiled corpora. Some of the widely-used are AntConc¹³, a freeware corpus analysis toolkit for concordancing and text analysis created by Laurence Anthony, LancsBox¹⁴, a Lancaster University corpus toolbox, Corpus Workbench¹⁵ and WordSmith Tools¹⁶. With the aim of doing this research we used TermoStat Web 3.0 corpus analysis tool.

3. Methodology

3.1. Research aim and hypotheses

The primary aim of the research was to analyse the two separate corpora extracted from annual reports of four biggest American banks for fiscal years 2006 and 2021. The extracted MWEs were analysed in 15 syntactic categories: 1) noun + noun; 2) adjective + noun; 3) adjective + adjective + adjective + noun; 4) noun + preposition + noun; 5) adjective + noun + noun; 6) noun + preposition + adjective + noun; 7) noun + noun + noun + noun; 8) adjective + adjective + noun; 9) noun + noun + noun; 10) adjective + noun + noun + noun; 11) noun + noun + noun + noun + noun; 12) noun + preposition + noun + noun; 13) noun + preposition + 4 nouns; 14) noun + adjective + noun; 15) noun + adjective + noun + noun. Further diachronic analysis made possible to compare the most frequent MWEs in the given categories and years and spot changes having come into existence over a decade and a half.

The secondary aim of the research was to check the frequency of occurrence of the above categories in relation to the total scope of corpus (total

¹⁰ See <https://www.sketchengine.eu/>

¹¹ See <http://www.natcorp.ox.ac.uk/>

¹² See <https://www.english-corpora.org/coca/>

¹³ See <https://www.laurenceanthony.net/software/antconc/>

¹⁴ See <http://corpora.lancs.ac.uk/lancsbox/download.php>

¹⁵ See <https://cwb.sourceforge.io/cwb.php>

¹⁶ See <https://www.lexically.net/wordsmith/>

number of words) in both corpora to determine possible diachronic changes in the occurrence of each category.

The final aim of the research was to raise awareness about possible application of gained results in BE classes, where BE lecturers must teach only authentic and up-to-date vocabulary.

Prior to the research two hypotheses were formulated:

1. The language in annual reports has experienced diachronic changes related to the frequency of the most frequent MWEs in 15 analysed syntactic categories. Even though the span of time encompassing a decade and a half does not represent a very long period in a language diachronically observed, the above change was expected. As financials in general are becoming more and more complex, the head-to-head competition in the banking sector is becoming more and more fierce, the processes within it are speeding up and the banks are continuously becoming more and more global, the above mentioned change is to be expected.

2. The total share of 15 syntactic categories as a percentage in relation to the scope of the corpora (total number of words) has diachronically remained the same. Although the author infers that the language in annual reports has diachronically changed on a lexical level over 15 years, he assumes a ratio of the syntactic categories to the total scope of the corpora to have remained stable. The texts in both corpora are of the same type and no significant numerical deviations were expected when it comes to the percentage of a particular category in relation to the total scope of corpus.

3.2. Corpus and corpus analysis tool

The two comparable corpora (those for the fiscal years 2006 and 2021) were compiled from the annual reports of four biggest US banks by assets: *JP Morgan Chase & Co.*, *Wells Fargo & Co.*, *Citigroup Inc.* and *U.S. Bancorp*. The two corpora are monolingual, representative, well-designed and large enough to enable the diachronic study. The corpora have the following counts:

	Corpus of 2006	Corpus of 2021
Words	259.491	334.373
Sentences	7.745	16.295
tag	63	63
pos	9	9
Documents	4	4

Table 3.

The number of words in the corpus of 2021 is larger because the content of annual reports has become bigger over time. In order to make diachronic comparison possible, all necessary statistical data were calculated as a percentage either of a single syntactic category or of a total scope of each corpus. The composition of the corpora with multiple authors of the analysed documents and its total scope encompassing 593.864 words prevent any type of idiolect that would influence the results in a negative way. The MWEs of the two sub-corpora were analysed in 15 syntactic categories:

	Categories
1.	Noun + noun
2.	Adjective + noun
3.	Adjective + adjective + adjective + noun
4.	Noun + preposition + noun
5.	Adjective + noun + noun
6.	Noun + preposition + adjective + noun
7.	Noun + noun + noun + noun
8.	Adjective + adjective + noun
9.	Noun + noun + noun
10.	Adjective + noun + noun + noun
11.	Noun + noun + noun + noun + noun
12.	Noun + preposition + noun + noun
13.	Noun + preposition + 4 nouns
14.	Noun + adjective + noun
15.	Noun + adjective + noun + noun

Table 4.

The corpus analysis was carried out by means of TermoStat Web 3.0¹⁷, a corpus analysis tool created by professor Patric Drouin from the University of Montreal, Canada. It is a corpus analysis tool developed for quantitative analysis of custom-made corpora. It is an open-source software, that requires initial registration and is thereafter free for research purposes. It enables analysis of texts in French, English, Spanish, Italian and Portuguese. A list of results with terms that can be simple (one word) or complex (a sequence of words) is made after a term gets a score based on term's frequency in the analysed corpus and its frequency in the reference corpus consisting of about 8.000.000 occurrences. The uploaded corpus, the so-called technical corpus (Drouin 2003) is compared with the non-technical one in this way. Since the tool does not analyse adjectives in postpositive position, such constructions were excluded from the ADJECTIVE

¹⁷ http://termostat.ling.umontreal.ca/index.php?lang=en_CA

+ NOUN category. Figure 1. shows the partial results of the analysis of JP Morgan Chase & Co. Annual Report 2021 listed by the frequency of terms.

Candidate (grouping variant)	Score		Variants	Pattern
	Frequency	(Specificity)		
firm	1680	204.64	firm firms	Common_Noun
loan	1192	187.39	loan loans	Common_Noun
credit	954	151.28	credit credits	Common_Noun
fair value	488	139.09	fair value fair values	Adjective Common_Noun
asset	657	134.14	asset assets	Common_Noun
risk	862	129.46	risk risks	Common_Noun
o	478	105.26	o	Common_Noun
value	796	101.66	value values	Common_Noun
derivative	262	96.99	derivative derivatives	Common_Noun
liability	308	96.04	liability liabilities	Common_Noun
security	649	95.91	security securities	Common_Noun
credit loss	215	92.3	credit loss credit losses equity	Common_Noun Common_Noun

Figure 1. Corpus analysis tool Termostat Web 3.0

When it comes to statistical part of the research, the software IBM SPSS 22 S (IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.) was used in the analysis. Statistical significance level (p -value) was set to 0,05 mirroring less than 5 % chance that the results would occur if the null hypothesis is true.

To compare the frequency of the occurrence of a particular MWE within a particular syntactic category, a Chi-square test was used. The results are shown in graphs in which the share of each MWE is shown in relation to the total number of MWEs in that category in a given year – 2006 and 2021. The share of each MWE is shown as a percentage of the total number of MWEs in the category for both years to determine the occurrence of the MWE in a given category. Additionally, the total share of MWEs for the whole categories are calculated in relation to the total scope of the corpus (number of words) for both fiscal years to determine diachronic category to scope of the corpus ratio.

The tools enabled the author to test his hypothesis ie. whether there is a statistically significant difference in frequency of occurrence of the MWEs across 15 analysed categories from a diachronic point of view. Another point of scientific analysis was to determine whether the share of particular categories as a percentage of total number of words in the corpus had remained the same over period of 15 years.

4. Corpus analysis and results

Corpus analysis of the annual reports of the biggest US banking corporations for fiscal years 2006 and 2021 – *JP Morgan Chase & Co., Wells Fargo & Co.,*

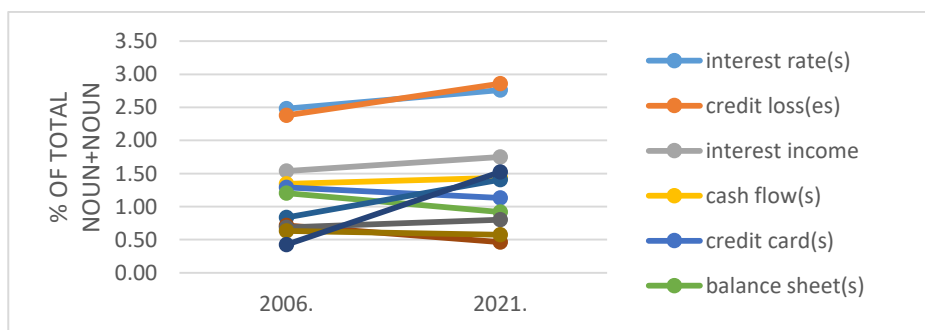
Citigroup Inc., U.S. Bancorp was made on 593.864 words. The separate sub-corpora were created out of Annual Reports 2006 and 2021 so that diacronic study of MWEs in 15 years could be done. The corpus analysis enabled the autor to determine the frequency of the occurrence of a particular syntactic category in the annual reports. The frequency of the occurrence itself is a very useful information for teachers of LSP / BE, since it reveals real BE in use. In this light, Stojković (2023) claims that "Corpora help lecturers indirectly, in deciding what to teach, but also in their direct use, regarding how to teach".

The most frequent MWEs for a respective category and fiscal year are as follows:

NOUN + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	interest rate(s) ¹⁸	326	credit loss(es)	671
2.	credit loss(es)	313	interest rate(s)	648
3.	interest income	202	interest income	411
4.	cash flow(s)	177	debt security	357
5.	credit card(s)	170	cash flow(s)	337
6.	balance sheet(s)	158	credit risk(s)	330
7.	credit risk(s)	110	credit card(s)	266
8.	noninterest expense	94	balance sheet(s)	216
9.	mortgage loan(s)	90	mortgage loan(s)	188
10.	non-interest income	83	non-interest income	135
	TOTAL NN CORPUS FINDINGS	13.146	TOTAL NN CORPUS FINDINGS	23.467

Table 5.



Graph 3.

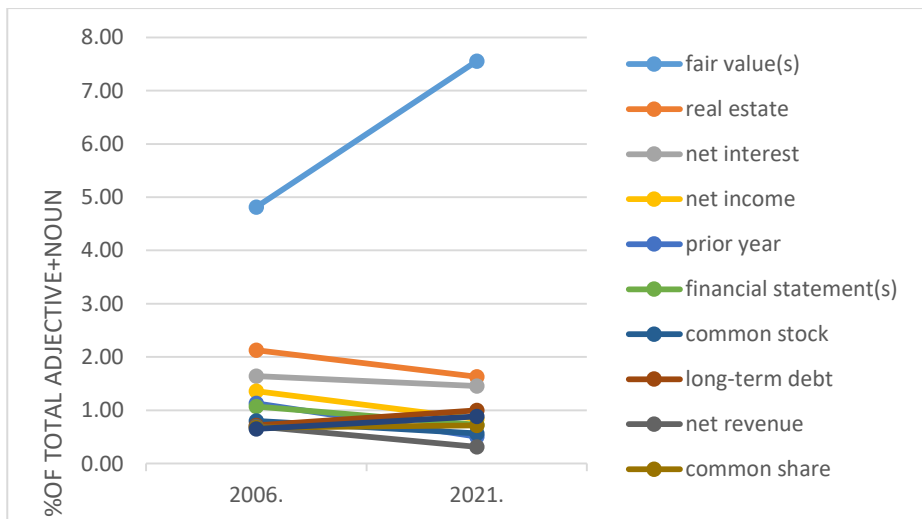
¹⁸ Terms highlighted in yellow indicate the presence of the same MWEs among the 10 most frequent ones in both corpora.

There are 5 MWEs of a type NOUN + NOUN with moderate diachronic rise ($\leq 25\%$) in the occurrence in this category – *interest rate(s)*, *credit loss(es)*, *interest income*, *cash flow(s)* and *mortgage loan(s)*. Two MWEs – *credit risk(s)* and *debt security* show sharp diachronic rise ($\geq 25\%$) in the occurrence within the category. There are 3 MWEs showing slight diachronic fall ($\leq 25\%$) in the occurrence – *non-interest income*, *credit card(s)* and *balance sheet(s)* and 1 MWE *noninterest expense* showing sharp diachronic fall in the occurrence within the category. The category NOUN + NOUN is the most frequent one (43.8 % in the corpus of 2006; 41.8 % in the corpus of 2021) The most prototypical MWEs in the category are *interest rate(s)* and *credit loss(es)*.

ADJECTIVE + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	fair value(s)	613	fair value(s)	1.675
2.	real estate	271	real estate	361
3.	net interest	209	net interest	323
4.	net income	173	long-term debt	221
5.	prior year	144	residential mortgage	196
6.	financial statement(s)	136	net income	186
7.	common stock	102	financial statement(s)	162
8.	long-term debt	91	common share	159
9.	net revenue	89	common stock	126
10.	common share	86	prior year	112
	TOTAL AN CORPUS FINDINGS	12.728	TOTAL AN CORPUS FINDINGS	22.181

Table 6.



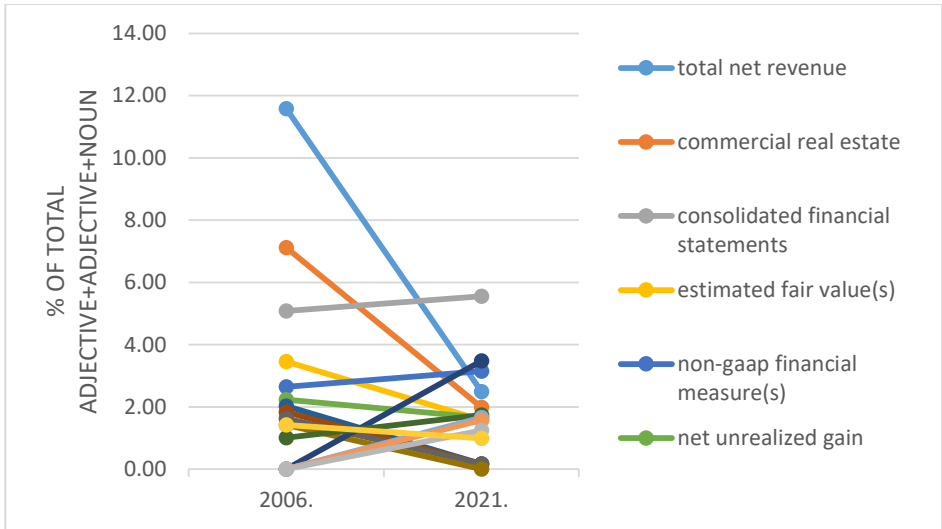
Graph 4.

This category has seen more dynamic diachronic change than the first one. There are 3 MWEs of a type ADJECTIVE + NOUN with sharp diachronic rise (≥ 25 %) in the occurrence in the category – *long-term debt*, *fair value(s)*, *residential mortgage*. There is 1 MWE with moderate diachronic rise (≤ 25 %) in the occurrence in this category – *common share(s)*. There are 5 MWEs showing sharp diachronic fall (≥ 25 %) in the occurrence within the category – *net revenue*, *common stock*, *financial statement(s)*, *prior year* and *net income*. The remaining 2 MWEs – *net interest* and *real estate* have slight diachronic fall (≤ 25 %) in their occurrence. The category ADJECTIVE + NOUN is the second most frequent one in both corpora (42.4 % in the corpus of 2006; 39.5 % in the corpus of 2021). The most prototypical MWEs in the category are *fair value(s)* and *real estate*.

ADJECTIVE + ADJECTIVE + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	total net revenue	57	consolidated financial statements	67
2.	commercial real estate	35	tangible common equity	42
3.	consolidated financial statements	25	non-gaap financial measure(s)	38
4.	estimated fair value(s)	17	total net revenue	30
5.	non-gaap financial measure(s)	13	commercial real estate	24
6.	net unrealized gain	11	net realizable value	21
7.	net actuarial loss	10	total retained loans, net unrealized gain	20
8.	hybrid financial instrument(s)	9	significant unobservable input(s), estimated fair value(s)	19
9.	material legal proceedings	8	non-cumulative perpetual class	15
10.	average total asset(s), gross unrealized losses	7	gross unrealized losses	12
	TOTAL AAN CORPUS FINDINGS	492	TOTAL AAN CORPUS FINDINGS	1.206

Table 7.



Graph 5.

The syntactic category ADJECTIVE + ADJECTIVE + NOUN is the sixth most productive category in the analysed corpora.

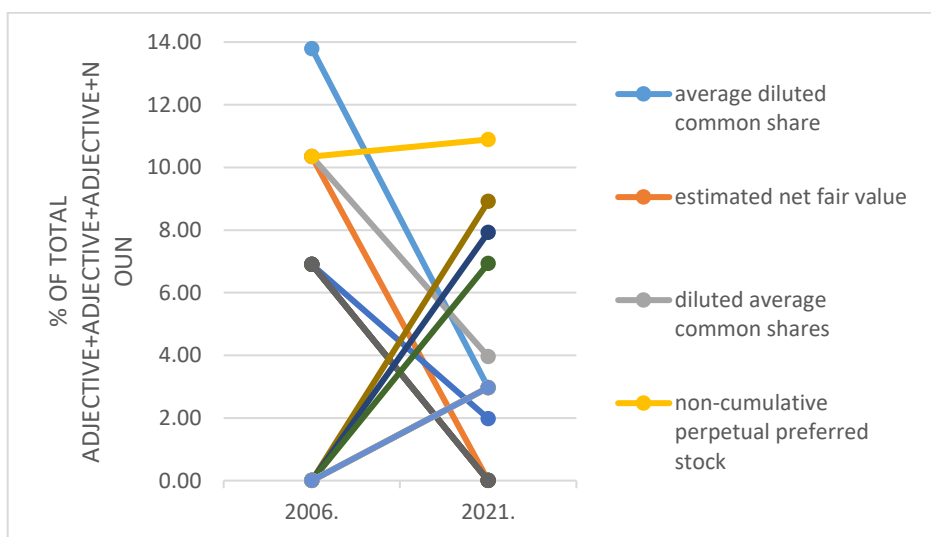
There are 2 MWEs of this type with moderate diachronic rise ($\leq 25\%$) in occurrence in this category – *consolidated financial statements* and *non-gaap financial measure(s)* and 1 MWE with sharp diachronic rise ($\geq 25\%$) – *net realizable value*. There are no MWEs showing slight diachronic fall ($\leq 25\%$) in the occurrence and 7 MWEs – *total net revenue*, *commercial real estate*, *estimated fair value(s)*, *net unrealized gain*, *hybrid financial instrument(s)*, *material legal proceedings* and *gross unrealized losses* – showing sharp diachronic fall in the occurrence in the category. The most prototypical MWEs in this category are *total net revenue* and *consolidated financial statements*.

ADJECTIVE + ADJECTIVE + ADJECTIVE + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	average diluted common share	4	non-cumulative perpetual preferred stock	11
2.	estimated net fair value	3	total commercial real estate	9
3.	diluted average common shares	3	average tangible common equity	8
4.	non-cumulative perpetual preferred stock	3	direct outstanding funded loans	7

5.	total junior subordinated debt	2	diluted average common shares	4
6.	total average interest-bearing deposits	2	total net derivative receivables	3
7.	total average common equity	2	total estimated fair value	3
8.	granted long-term stock-based awards	2	total long-term unsecured funding	3
9.	total net managed revenue	2	real gross domestic product	3
10.			total non-accrual corporate loans	3
	TOTAL AAAN CORPUS FINDINGS	29	TOTAL AAAN CORPUS FINDINGS	101

Table 8.



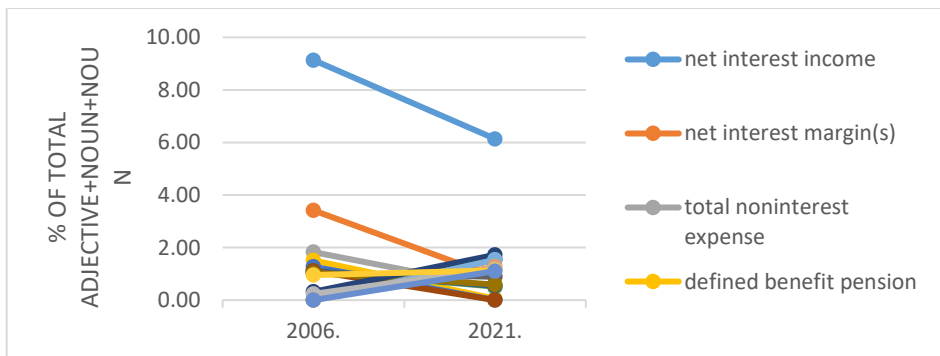
Graph 6.

The syntactic category ADJECTIVE + ADJECTIVE + ADJECTIVE + NOUN is among unproductive categories in the analysed corpora. There is 1 MWE of this type with moderate diachronic rise ($\leq 25\%$) in the occurrence in this category – *non-cumulative perpetual preferred stock*. There are 3 MWEs showing sharp diachronic fall ($\geq 25\%$) in the occurrence within this category – *average diluted common share*, *diluted average common shares* and *total junior subordinated debt*. The most prototypical MWEs in this category are *non-cumulative perpetual preferred stock* and *total commercial real estate*.

ADJECTIVE + NOUN + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	net interest income	115	net interest income	228
2.	net interest margin(s)	43	consolidated balance sheet	64
3.	total noninterest expense	23	nonmarketable equity security(ies)	57
4.	defined benefit pension	19	fair value option	57
5.	net charge-off rate(s)	16	net credit losses	47
6.	real estate mortgage	14	fair value measurement(s)	44
7.	real estate construction	14	residential mortgage loans	42
8.	corporate trust business(es)	14	fair value hierarchy	41
9.	unfunded credit commitments	13	unfunded credit commitments	33
10.	marketable equity securities	13	net interest margin(s)	33
	TOTAL ANN CORPUS FINDINGS	1259	TOTAL ANN CORPUS FINDINGS	3715

Table 9.



Graph 7.

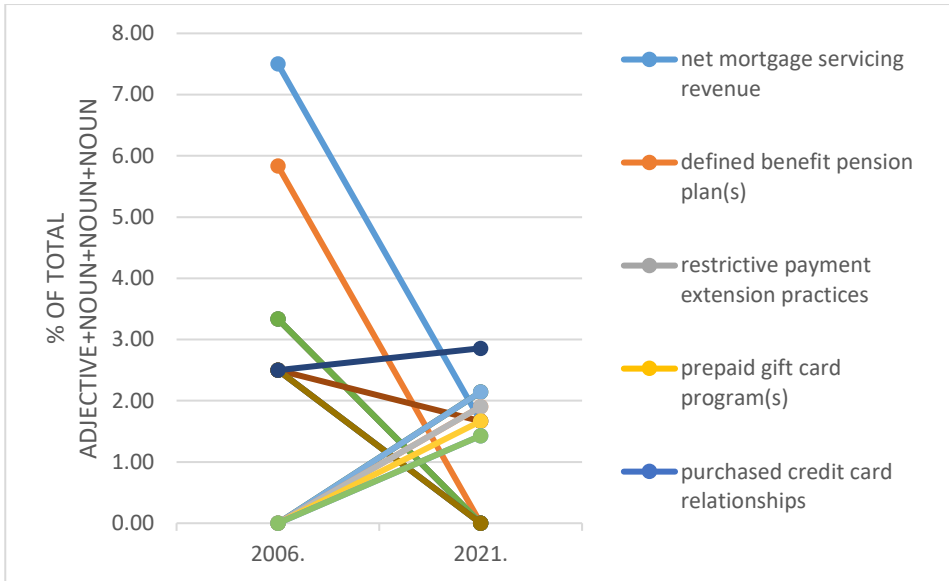
The syntactic category ADJECTIVE + NOUN + NOUN is the third frequent one in both corpora (4.2 % in the corpus of 2006; 6.6 % in the corpus of 2021). Two MWEs – *consolidated balance sheet* and *fair value measurement(s)* show sharp diachronic rise (≥ 25 %) in the occurrence within the category. The following

MWEs show sharp diachronic fall ($\geq 25\%$) – *net interest income*, *net interest margin(s)*, *total non interest expense*, *defined benefit pension*, *net charge-off rate(s)* in the occurrence in this category. The most prototypical MWEs in the category are *net interest income* and *net interest margins*.

ADJECTIVE + NOUN + NOUN + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	net mortgage servicing revenue	9	foreign currency translation adjustment(s)	12
2.	defined benefit pension plan(s)	7	federal income tax rate	9
3.	restrictive payment extension practices	4	net credit loss rate(s)	9
4.	prepaid gift card program(s)	4	primary credit quality indicator	8
5.	purchased credit card relationships	4	corporate payment products revenue	8
6.	total income tax expense	4	effective income tax rate	7
7.	global equity market appreciation	3	net mortgage servicing revenue	7
8.	selected balance sheet data	3	selected balance sheet data	6
9.	fair value measurement method	3	off-balance sheet credit exposure(s)	6
10.	residential mortgage servicing rights	3	nonconforming mortgage loan securitizations	6
	TOTAL ANNN CORPUS FINDINGS	120	TOTAL ANNN CORPUS FINDINGS	420

Table 10.



Graph 8.

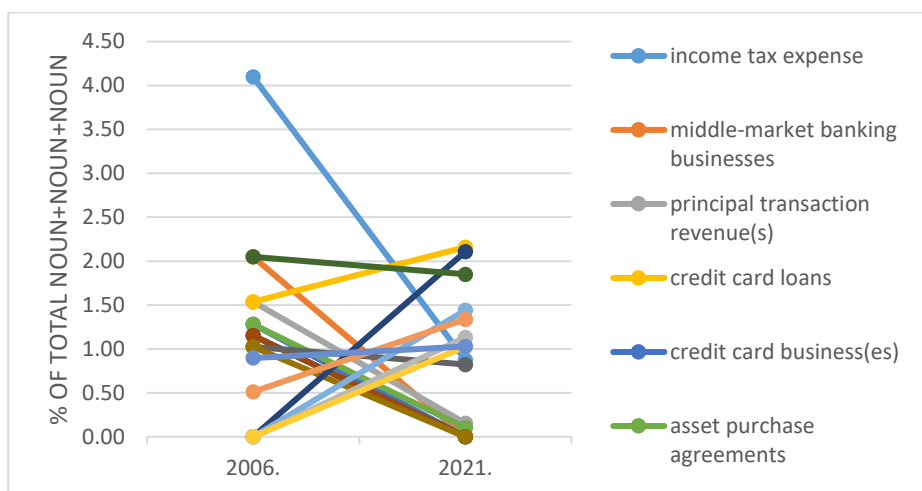
There are 2 MWEs with sharp diachronic fall ($\geq 25\%$) in their occurrence in the category ADJECTIVE + NOUN + NOUN + NOUN – *net mortgage servicing revenue* and *selected balance sheet data*. There is 1 MWE showing slow diachronic rise ($\leq 25\%$) in the given category – *foreign currency translation adjustment(s)*. The majority of the most frequent corpus findings in this category only occur either in the corpus of 2006 or in the corpus of 2021, so that they could not be compared with each other. The most prototypical MWEs in the category are *net mortgage servicing revenue* and *foreign currency translation adjustment(s)*.

NOUN + NOUN + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	income tax expense	32	credit card loan(s)	42
2.	middle-market banking businesses	16	consumer banking business(es)	41
3.	principal transaction revenue(s)	12	cash flow hedge(s)	36
4.	credit card loans	12	trading account assets	28
5.	credit card business(es)	10	investment securities portfolio	26
6.	asset purchase agreements	10	trading account liabilities	22

7.	document management services	9	capital transition provision(s)	20
8.	stock repurchase program	9	mortgage banking revenue	20
9.	bank holding company	8	income tax expense	17
10.	investment management fees	8	bank holding company	16
	TOTAL NNN CORPUS FINDINGS	781	TOTAL NNN CORPUS FINDINGS	1.945

Table 11.



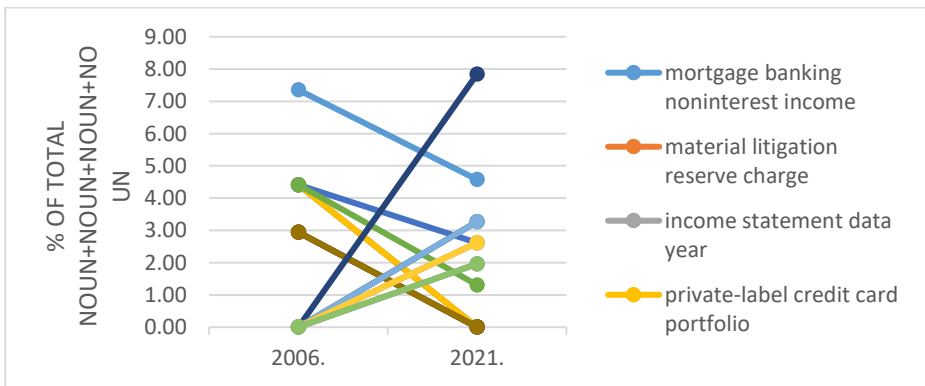
Graph 9.

The syntactic category NOUN + NOUN + NOUN is the fifth most productive in the analysed corpora. Only one MWE has a slight diachronic rise ($\leq 25\%$) in this category – *mortgage banking revenue*. There are 2 MWEs showing sharp diachronic rise ($\geq 25\%$) in the given category – *credit card loan(s)* and *investment securities portfolio*. The following 2 MWEs have a slight diachronic fall ($\leq 25\%$) in their occurrence – *cash flow hedge(s)* and *bank holding company*, whereas the following MWEs show sharp diachronic rise ($\geq 25\%$) in their occurrence – *income tax expense*, *principal transaction revenue(s)* and *asset purchase agreements*. The most prototypical MWEs in the category are *income tax expense* and *credit card loan(s)*.

NOUN + NOUN + NOUN + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	net mortgage servicing revenue	9	foreign currency translation adjustment(s)	12
2.	defined benefit pension plan(s)	7	federal income tax rate	9
3.	restrictive payment extension practices	4	net credit loss rate(s)	9
4.	prepaid gift card program(s)	4	primary credit quality indicator	8
5.	purchased credit card relationships	4	corporate payment products revenue	8
6.	total income tax expense	4	effective income tax rate	7
7.	global equity market appreciation	3	net mortgage servicing revenue	7
8.	selected balance sheet data	3	selected balance sheet data	6
9.	fair value measurement method	3	off-balance sheet credit exposure(s)	6
10.	residential mortgage servicing rights	3	nonconforming mortgage loan securitizations	6
	TOTAL NNNN CORPUS FINDINGS	120	TOTAL NNNN CORPUS FINDINGS	420

Table 12.



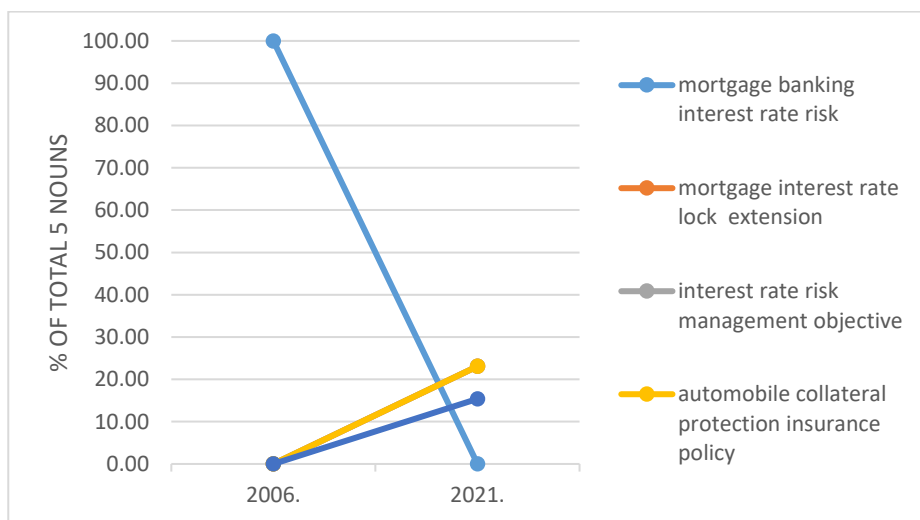
Graph 10.

The following MWEs show sharp diachronic fall ($\geq 25\%$) in this syntactic category – *credit risk management process*, *merchant processing services revenue* and *mortgage banking noninterest income*. Other MWEs could not be compared with each other since they are non-existent either in the corpus of 2006 or of 2021. The most prototypical MWEs in the category are *net mortgage servicing revenue* and *foreign currency translation adjustment(s)*.

5 NOUNS

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	mortgage banking interest rate risk	3	mortgage interest rate lock extension	3
2.			interest rate risk management objective	3
3.			automobile collateral protection insurance policy	3
4.			cash balance interest crediting rate	2
	TOTAL 5N CORPUS FINDINGS	3	TOTAL 5N CORPUS FINDINGS	13

Table 13.



Graph 11.

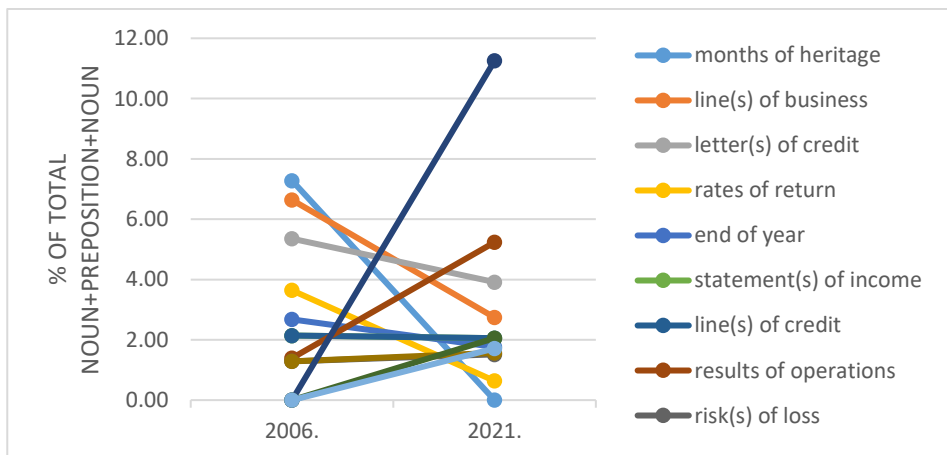
The syntactic category 5 NOUNS is the second least productive among all analysed categories with totally 16 corpus findings in both corpora). The few

MWEs from one corpus do not overlap with the findings from another one, so they could not be statistically compared.

NOUN + PREPOSITION + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	months of heritage	68	millions of dollars	230
2.	line(s) of business	62	results of operations	107
3.	letter(s) of credit	50	letter(s) of credit	80
4.	rates of return	34	line(s) of business	56
5.	end of year	25	billions of dollars	42
6.	statement(s) of income	20	line(s) of credit	42
7.	line(s) of credit	20	statement(s) of income	42
8.	results of operations	13	end of year	36
9.	risk(s) of loss	12	line(s) of defense	35
10.	course of business	12	course of business	32
	TOTAL NPN CORPUS FINDINGS	934	TOTAL NPN CORPUS FINDINGS	2.044

Table 14.



Graph 12.

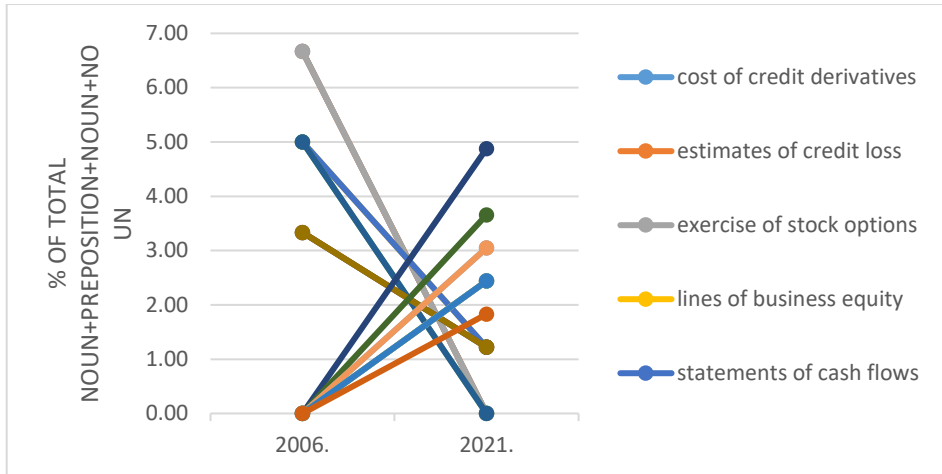
There is one collocation – *course of business* of a type NOUN + PREPOSITION + NOUN with moderate diachronic rise ($\leq 25\%$) in the occurrence in this category. Two MWEs – *consolidated balance sheet* and *fair value measurement(s)* show sharp diachronic rise ($\geq 25\%$) in the occurrence within the category. The following MWEs show sharp diachronic fall ($\geq 25\%$) – *net interest income*, *net interest margin(s)*, *total non interest expense*, *defined benefit*

pension, net charge-off rate(s) in occurrence in this category. The syntactic category NOUN + PREPOSITION + NOUN is the fourth most frequent one in both corpora (3.1 % in the corpus of 2006; 3.6 % in the corpus of 2021) The most prototypical MWEs in the category are *millions of dollars* and *results of operations*.

NOUN + PREPOSITION + NOUN + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	cost of credit derivatives	4	netting of cash collateral	8
2.	estimates of credit loss	4	impact(s) of fx translation	6
3.	exercise of stock options	4	reduction of interest revenue	5
4.	lines of business equity	3	charge-off of peak delinquencies	5
5.	statements of cash flows	3	reduction of card income	4
6.	timing of repurchase activity	3	market comparable pricing techniques	4
7.	rate of compensation increase	3	trust preferred security vies	4
8.	level of market volatility	2	recognition of credit losses	4
9.	calculation of credit risk	2	transfer of investment securities	4
10.	percent of credit exposure	2	components of card income	3
	TOTAL NPNN CORPUS FINDINGS	60	TOTAL NPNN CORPUS FINDINGS	164

Table 15.



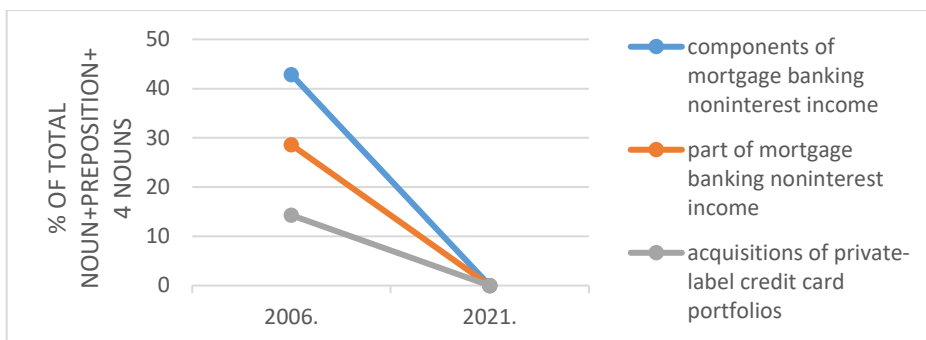
Graph 13.

The upper syntactic category is the unproductive one with few corpus findings related to the scope of both corpora. When taking into consideration the most frequent MWEs, they are completely different in both corpora (some of them do not occur in the first corpus, other ones in the second one). Still, there are two MWEs existent in both corpora – *statement of cash flows* and *percent of credit exposure* showing sharp diachronic fall ($\geq 25\%$) in the occurrence in the given category.

NOUN + PREPOSITION + 4 NOUNS

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	components of mortgage banking noninterest income	3		
2.	part of mortgage banking noninterest income	2		
3.	acquisitions of private-label credit card portfolios	1		
	TOTAL NP + 4N CORPUS FINDINGS	7	TOTAL NP + 4N CORPUS FINDINGS	0

Table 16.



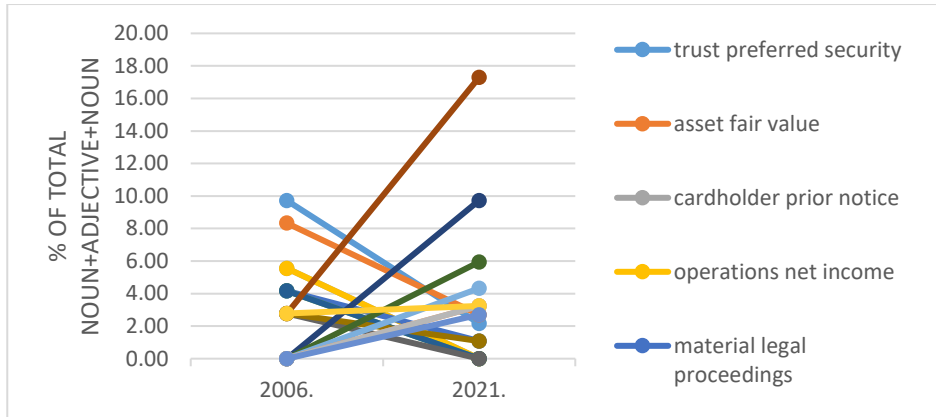
Graph 14.

The syntactic category NOUN + PREPOSITION + 4 NOUNS is the least productive among all analysed categories. There are altogether 7 corpus findings in the corpus of 2006 and none corpus findings in the corpus of 2021.

NOUN + ADJECTIVE + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	trust preferred security	7	deposit related fees	32
2.	asset fair value	6	agency mortgage-backed securities	18
3.	cardholder prior notice	4	credit derivative contract(s)	11
4.	operations net income	4	credit risk-weighted assets	8
5.	material legal proceedings	3	securities loaned agreements	6
6.	employees stock-based incentives	3	business retail banking	6
7.	employees stock-based plans	3	option adjusted spread	6
8.	deposit related fees	2	market comparable pricing	5
9.	payment due date	2	asset fair value	5
10.	consumer real estate	2	trust preferred securities	4
	TOTAL NAN CORPUS FINDINGS	72	TOTAL NAN CORPUS FINDINGS	185

Table 17.



Graph 15.

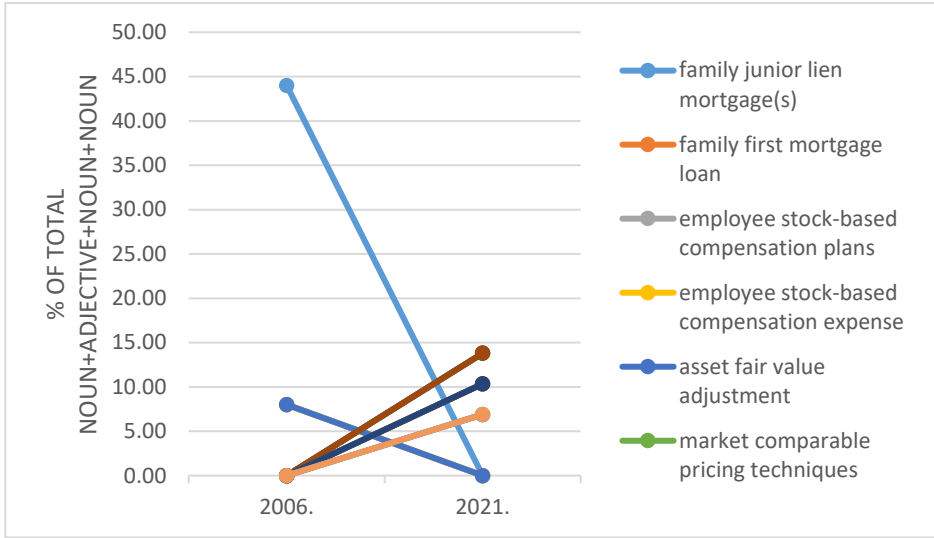
The category NOUN + ADJECTIVE + NOUN is the unproductive one. There are 4 MWEs showing sharp diachronic fall ($\geq 25\%$) in the occurrence in the given category – *trust preferred security*, *asset fair value*, *material legal proceedings* and *consumer real estate*. There is 1 MWE having a slight diachronic rise ($\leq 25\%$) in the occurrence in this category – *option adjusted spread* and 1 collocation showing a sharp diachronic rise ($\geq 25\%$) – *deposit related fees*. The most prototypical MWEs in the category are *deposit related fees* and *agency mortgage-backed securities*.

NOUN + ADJECTIVE + NOUN + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	family junior lien mortgage(s)	11	market comparable pricing techniques	4
2.	family first mortgage loan	2	trust preferred security vies	4
3.	employee stock-based compensation plans	2	multi-seller commercial paper conduits	4
4.	employee stock-based compensation expense	2	assets total impairment losses	3
5.	asset fair value adjustment	2	revenue net interest income	3
6.			company adopted accounting guidance	3
7.			government supported loan programs	2
8.			amount accumulated amortization net	2

9.			deposit net interest income	2
	TOTAL NANN CORPUS FINDINGS	25	TOTAL NANN CORPUS FINDINGS	29

Table 18.



Graph 16.

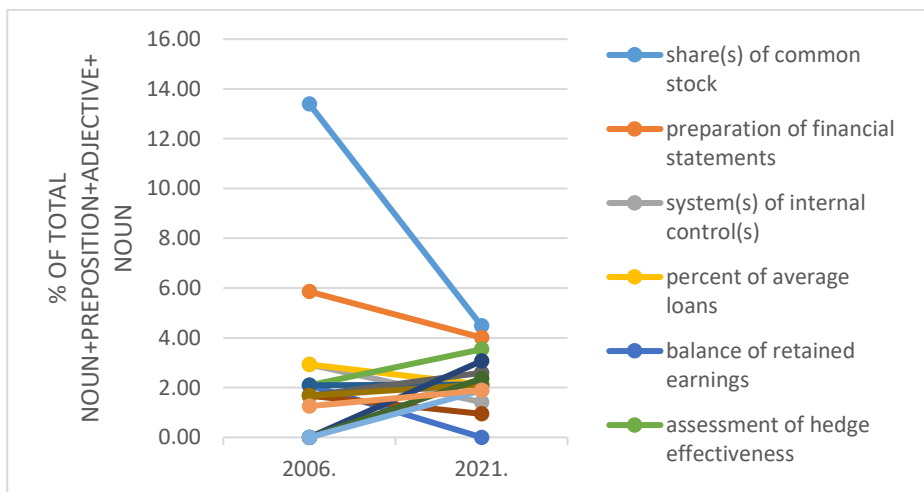
The syntactic category NOUN + ADJECTIVE + NOUN + NOUN is the third least productive among all analysed categories (totally 54 corpus findings in both corpora). There are 5 corpus findings with sharp diachronic fall ($\geq 25\%$) in occurrence in the given category.

NOUN + PREPOSITION + ADJECTIVE + NOUN

	Annual Reports 2006	Findings	Annual Reports 2021	Findings
1.	share(s) of common stock	32	shares of common stock	19
2.	preparation of financial statements	14	preparation of financial statements	17
3.	system(s) of internal control(s)	7	assessment of hedge effectiveness	15
4.	percent of average loans	7	significance of unobservable inputs	13
5.	balance of retained earnings	5	issuance(s) of long-term debt	11

6.	assessment of hedge effectiveness	5	percent of risk-weighted assets	10
7.	redemption of preferred stock	5	percent of average loans	9
8.	issuance of common stock	4	redemption of preferred stock	9
9.	issuance(s) of long-term debt	4	determination of fair value	8
10.	percent of average loans	4	estimate(s) of fair value	8
	TOTAL NPAN CORPUS FINDINGS	239	TOTAL NPAN CORPUS FINDINGS	424

Table 19.



Graph 17.

The syntactic category NOUN + PREPOSITION + ADJECTIVE + NOUN is in the middle category regarding the frequency of its occurrence (the seventh most productive one). There are 4 MWEs of a type NOUN + PREPOSITION + ADJECTIVE + NOUN with sharp diachronic rise ($\geq 25\%$) in their occurrence – *assessment of hedge effectiveness*, *issuance(s) of long-term debt*, *percent of average loans* and *estimate(s) of fair value*. There is 1 MWE with moderate diachronic rise ($\leq 25\%$) in the occurrence in this category – *redemption of preferred stock*. There are 4 MWEs showing sharp diachronic fall ($\geq 25\%$) in the occurrence in the given category – *share(s) of common stock*, *preparation of financial statements*, *issuance of common stock* and *system(s) of internal control(s)*. The most prototypical MWEs in the category are *share(s) of common stock* and *preparation of financial statements*.

At the beginning of the research two hypothesis were put forward:

1. The language in English for banking has experienced diachronic changes related to the frequency of the most frequent MWEs in 15 analysed syntactic categories. The Chi-square test was made for all 15 categories. Only in the category NOUN + PREPOSITION + 4 NOUNS the Chi-square test could not be made since there are not any MWEs in that category for 2021. There is a p-value < 0,0001 in each syntactic category proving high statistical significance that the occurrence of the most frequent MWEs has diachronically changed when comparing corpora from 2006 and 2021. That confirmed our first hypothesis.
2. The total share of 15 syntactic categories as a percentage of scope of the corpus (total number of words) has diachronically remained the same.

Category	Category share in 2006	Category share in 2021	2021/2006
NN	5.07	7.02	1.38
AN	4.90	6.63	1.35
AAN	0.19	3.61	19
AAAN	0.01	0.03	3
ANN	0.49	1.11	2.27
ANNN	0.05	0.13	2.6
3N	0.30	0.58	1.93
4N	0.03	0.05	1.66
5N	0.001	0.004	4
NPN	0.36	0.61	1.69
NPNN	0.02	0.05	2.5
NP4N	0.003	0	-
NAN	0.03	0.06	2
NANN	0.01	0.008	0.8
NPAN	0.09	0.13	1.44
			Ø 1.90

Table 20.

After two extreme values were sorted out that would substantially influence the statistics, we became 2021/2006 ratio of 1.90 which shows that there are 90 % more of the analysed categories represented in the Annual Reports in 2021. The set hypothesis proved false. The analysed texts have become more syntactically structured over 15 years which mirrors their higher grade of complexity. All syntactic categories have seen rise with an exception of two categories: NOUN + PREPOSITION + 4 NOUNS (a syntactic category not present in the corpus from 2021) and NOUN + ADJECTIVE + 2 NOUNS (2021/2006 ratio = 0.8). The 2021/2006 ratio of the category ADJECTIVE + ADJECTIVE + NOUN

was 19 and was sorted out from calculation as an extreme value. The percentage diachronic rise in syntactic categories is as follows: 5N – 300 %, AAAN – 200 %, ANNN – 160 %, NPNN – 150 %, ANN – 127 %, NAN – 100 %, 3N – 93 %, NPN – 69 %, 4N – 66 %, NPAN – 44 %, NN – 38 %, AN – 35 %.

5. Discussion

Due to shortage of authentic teaching materials, the BE lecturers are often doomed to create their own ones. The research was started with the awareness in mind that only words / phrases with confirmed frequency of occurrence are worth teaching / learning in LSP, in this case BE classes. Therefore, corpus analysis was chosen as an appropriate method for the extraction of authentic MWEs that can be analysed and transformed into valuable teaching material. If BE lecturers want to create such kind of teaching material, they must put emphasis on the most important, useful and frequent in language. The results of skillfully done corpus analysis enables precisely that. The diachronic component of the study enabled getting the up-to-date results, showing on the one hand changes in the language and on the other the newest vocabulary suitable for application in BE classes. Since the teaching hours are limited, the teaching should be as efficient as possible. Teaching students the most frequent vocabulary makes teaching more productive – there is no use in teaching general forms and made up (randomly constructed) MWEs.

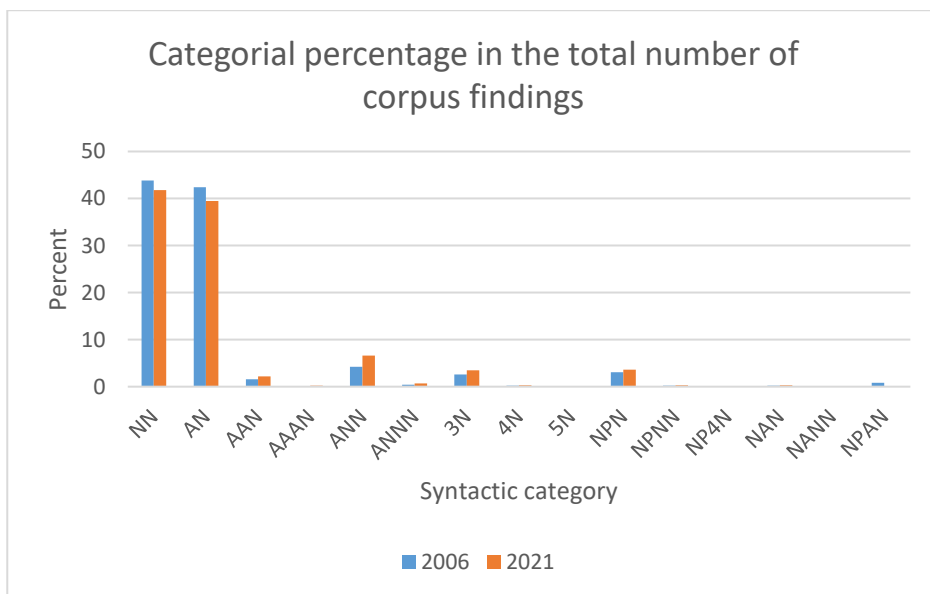
The diachronic corpus analysis of Annual Reports of biggest US banking corporations for fiscal years 2006 and 2021 – *JP Morgan Chase & Co.*, *Wells Fargo & Co.*, *Citigroup Inc.*, *U.S. Bancorp* with its p -value $< 0,0001$ across all syntactic categories revealed diachronic changes in analysed language of high statistical significance. One of the characteristics of these financial documents diachronically regarded is that the annual reports have become more complex over the last 15 years. This applies both to their structure and their scope. The following table gives an overview of syntactic categories sorted by the number of corpus findings:

Category	Corpus findings 2006	Corpus findings 2021	Total number of findings in two corpora
NN	13.146	23.467	36.613
AN	12.728	22.181	34.909
AAN	492	1.206	1.698
AAAN	29	101	130
ANN	1.259	3.715	4.974
ANNN	120	420	540
3N	781	1.945	2.726

4N	68	153	221
5N	3	13	16
NPN	934	2.044	2.978
NPNN	60	164	224
NP4N	7	0	7
NAN	72	185	257
NANN	25	29	54
NPAN	239	424	663
	29.963	56.047	86.010

Table 21.

The syntactic categories with the most corpus findings are NOUN + NOUN – 36.613 corpus findings, ADJECTIVE + NOUN – 34.909, ADJECTIVE + NOUN + NOUN – 4.974, NOUN + PREPOSITION + NOUN – 2.978 and NOUN + NOUN + NOUN – 2.726. Graph 18. shows the percentage of each syntactic category in the total number of corpus findings in 15 categories:



Graph 18.

6. Conclusion

This research gives a diachronic insight into real English for banking in use, which is a part of BE curriculum at a tertiary level of education. The results have proved high statistical significance of changes in both frequency of the occurrence of MWEs across the 15 analysed categories and in their percentual relation to the scope of the corpus. The corpus analysis helps both BE lecturers

and their students to become more educated in both English and major they teach / learn – business. Research findings are relevant to BE lecturers in many ways: they enable them to follow recent vocabulary trends in this segment of BE. Moreover, they enable them to create their own working material based on the recently determined vocabulary and to improve the quality and efficiency of their teaching in this way. When it comes to students, the contribution to more effective learning is undoubtedly present too, as vocabulary is quite specific and an average BE student's time at disposal for their learning is limited, regarding the fact that precisely the time is actually the most expensive asset. The research can function as a driving force for further corpus analysis research in BE, which should result in optimizing and improving teaching techniques based on the results of the same. The prospective corpus analysis research could use either diachronic or contrastive approach to the categories of vocabulary and parts of speech. The corpus analysis research could include other segments of BE such as international trade, logistics, wholesale, retailing, industry, business media, leadership, management, marketing and investing.

CORPUS AND STATISTICAL ANALYSIS TOOL

TermoStat Web 3.0 http://termostat.ling.umontreal.ca/index.php?lang=en_CA
IBM SPSS 22 S (IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.)

WORKS CITED:

- Anđelković, Jelena, et al. "Using Parallel Corpora for Designing ESP Teaching Materials in Higher Education". *Inter Alia*, Volume 4, 9-20, 2018.
- Bennett, Gena R. *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*. University of Michigan, 2010.
- Brown, Keith, et. al. *Oxford Learner's Pocket Dictionary of Business English*. Oxford: Oxford University Press, 2015.
- Bullon, Stephen, et al. *Longman Business English Dictionary*. Harlow: Pearson Education Limited, 2007.
- Cobb, Tom. "Breadth and depth of lexical acquisition with hands-on concordancing". *Computer Assisted Language Learning*, 12(4), 345–360, 1999.
- Drouin, Patrick. "Term Extraction Using Non-Technical Corpora as a Point of Leverage". *Terminology* 9 (1): 99-117, 2003.
- Đurović, Zorica, and Bauk, Sanja. "Testing Corpus Linguistics Methods in ESP Vocabulary Teaching". *The Journal of Teaching English for Specific and Academic Purposes*. Vol 10, N° 2, 319-337, 2022.
- Howarth, Peter. "Phraseology and second language proficiency". *Applied linguistics*, 19(1), 24–44, 1998.

- Mockienė, Liudmila. "Application of Corpus Analysis of Terminology in Language Studies at University Level". *Corpus Analysis for Language Studies at the University Level*. Cambridge: Cambridge Scholars Publishing, 21-82, 2023.
- Römer, Ute. "Corpora and language teaching". *Corpus Linguistics – An International Handbook Volume 1*. Berlin, New York: De Gruyter Mouton, 112-130, 2008. <https://doi.org/10.1515/booksetHSK29>
- Römer, Ute. "Corpus Research Applications in Second Language Teaching". *Annual Review of Applied Linguistics*, 31, 205–225, 2011.. <https://doi.org/10.1017/S0267190511000055>
- Stojković, Nadežda. "Theoretical Considerations for Teaching Foreign Languages in University Education Settings". *Corpus Analysis for Language Studies at the University Level*. Cambridge: Cambridge Scholars Publishing, 8-20, 2023.
- Tribble, Christopher. "Small corpora and teaching writing". *Small corpus studies and ELT: theory and practice*. Amsterdam: John Benjamins, 381–408, 2001.

INTERNET RESOURCES

- <https://crsreports.congress.gov/product/pdf/IF/IF11707>
- https://ec.europa.eu/eurostat/statistics-explained/images/4/40/Number_of_tertiary_education_graduates_in_each_detailed_field_of_education%2C_EU%2C_2020_%281_000%29_ET2022.png
- [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Distribution_of_tertiary_education_students_by_broad_field_and_sex,_EU,_2020_\(%25\)_ET2022.png#filelinks](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Distribution_of_tertiary_education_students_by_broad_field_and_sex,_EU,_2020_(%25)_ET2022.png#filelinks)
- https://elt.oup.com/learning_resources/subjects/businessenglish/?cc=global&elLanguage=en&mode=hu
- <https://graftonbanksfinance.co.uk/news/news.php?news=70>
- <https://www.cambridge.es/en/catalogue/business-english/other-titles/english-for-the-financial-sector>
- <https://www.cambridge.org/is/cambridgeenglish/catalog/business-professional-and-vocational>
- <https://www.goldmansachs.com/our-firm/history/moments/1933-glass-steagall-act.html>
- <https://www.govinfo.gov/content/pkg/PLAW-106publ102/pdf/PLAW-106publ102.pdf>
- <https://www.investopedia.com/articles/investing/062513/role-commercial-banks-economy.asp>
- <https://www.pearson.com/english/catalogue/business-english.html>
- <https://www.yardeni.com/pub/sp500margin.pdf>

ANNUAL REPORTS

- <https://ir.usbank.com/static-files/b85c3163-a83c-44db-87ee-e5214afb6896>
- <https://ir.usbank.com/static-files/effe82f7-ec42-4f18-84e1-43504b33d239>

https://www.annualreports.com/HostedData/AnnualReportArchive/j/NYSE_JPM_2006.pdf

https://www.citigroup.com/rcs/citigpa/storage/public/ar06c_en.pdf?ieNocache=325

https://www.citigroup.com/rcs/citigpa/storage/public/ar11c_en.pdf?ieNocache=325

<https://www.jpmorganchase.com/content/dam/jpmc/jpmorgan-chase-and-co/investor-relations/documents/annualreport-2021.pdf>

<https://www.wellsfargohistory.com/assets/pdf/annual-reports/2006-annual-report.pdf>

<https://www08.wellsfargomedia.com/assets/pdf/about/investor-relations/annual-reports/2021-annual-report.pdf>

ZUSAMMENFASSUNG: THE MOST FREQUENT MULTI-WORD EXPRESSIONS IN ENGLISH FOR BANKING: A CORPUS-BASED DIACHRONIC STUDY

Das Ziel der Arbeit bestand darin, eine diachrone Untersuchung in einem Segment des Wirtschaftsenglischen – Englisch für Bankenwesen – durchzuführen, um die diachrone Stabilität seines Wortschatzes zu überprüfen. Daher wurde eine Analyse von Begriffen in 15 syntaktischen Kategorien in Geschäftsberichten der vier größten US-Banken für die Geschäftsjahre 2006 und 2021 durchgeführt. Genauer gesagt, wurden die diachronen Veränderungen, die sich über einen Zeitraum von 15 Jahren erstreckten, sowohl hinsichtlich der Häufigkeit des Auftretens der Begriffe in den 15 analysierten syntaktischen Kategorien als auch in ihrer prozentualen Beziehung zum Umfang der Korpora analysiert. Die Korpusanalyse ermöglichte es, die tatsächlich verwendete Sprache zu extrahieren, was es dem Autor ermöglichte, diachrone Veränderungen im Wortschatz über Kategorien hinweg zu bestimmen und so die Lücke zwischen den vorhandenen Kenntnissen des englischen Wortschatzes für Bankenwesen und dem neuesten Stand der Dinge zu schließen. Die Untersuchung zeigte, dass Änderungen sowohl in der Häufigkeit des Auftretens von Begriffen in den 15 analysierten Kategorien als auch in ihrem prozentualen Verhältnis zum Umfang des Korpus im Zeitraum von 15 Jahren eine hohe statistische Signifikanz aufweisen. Forschungsergebnisse können sowohl für Business-Englisch-Dozenten als auch für ihre Studierenden nützlich sein, da die Ergebnisse verschiedene mögliche pädagogische Implikationen haben. Die Forschung kann als Anregung für Linguisten dienen, die in der Korpuslinguistik forschen, weitere Forschungen in anderen Bereichen des Wirtschaftsenglischen durchzuführen, um mögliche Veränderungen in der Sprache zu ermitteln und zu beschreiben. Zu diesen Bereichen könnten gehören: internationaler Handel, Logistik, Großhandel, Einzelhandel, Industrie, Wirtschaftsmedien, Führung, Management, Marketing und Investieren.

Schlüsselwörter: Englisch für spezielle Zwecke, Wirtschaftsenglisch, Englisch für Bankenwesen, Wortschatz, Unterrichten vom Wirtschaftsenglischen auf tertiärer Ebene