UDK 811.111'373.7:33
Izvorni naučni rad

Dr. sc. Edina Rizvić-Eminović
Nirvana Koluh, MA

TYPES AND FREQUENCY OF COMPOUNDS IN A SELECTION OF TEXTS IN ECONOMICS

Abstract

Compounding in English is a complex linguistic phenomenon. Analyses of compounds can be quite demanding as it is sometimes necessary to analyze word classes comprising compounds in order to distinguish between composite nominals and compound nouns. Due to the complexity of their form, it may be rather difficult to interpret their meaning appropriately, especially in the absence of proper contextual clues. The present paper investigates the frequency and types of compounds in a selection of texts in economics collected from the 1995 and 2015 journal volumes. The analysis includes frequency data and comparisons between different types of compounds used in the two-decade period. The results show that the frequency of compounds in economic texts has increased, along with their complexity. The most frequent type were the nominal compounds, neoclassical being the least used.

***Keywords**: compounds, economic texts, nominal, neoclassical, verbal*

Introduction

Compounding in English is a complex linguistic phenomenon. While some linguists (Lieber et al., 2011) do not consider compounding a separate word formation process, other linguists do (Plag, 2003). Further, linguists disagree with respect to the terminology they use to define the building blocks of compounds. While Bauer (1983) defines compounds as containing two or more stems, Fabb (1998) considers them to be combinations of two or more words. Matthews, in turn, (1991) uses the term lexemes. According to Lieber et al. (2011), compounding is "a linguistic phenomenon that might at first glance seem straightforward" and is defined as "the formation of a new lexeme by adjoining two or more lexemes". In contrast to complex words, which are made up of roots and affixes, compounds, according to Plag (2003) are combinations of bases.

The relationship between compound elements can be explained in grammatical terms, as those elements can be nouns, verbs or words derived from a verb or an adjective (Greenbaum, 1996). Nonetheless, the analysis of compounds is not straightforward and is represents the most controversial type of word-formation processes in English (Plag, 2003). The English language is considered to be an open compounding language as most compound elements are written as separate words rather than single or hyphenated units (Ziering, 2018).

Compound words and their usage across different registers and genres of the English language have been the subject of many studies. Secord (2015) investigated compound words

within literary works – the Anglo-Saxon vernacular literature. Pisoschi (2012) conducted a research into characteristics of adjectival compounds in Shakespeare's sonnets (Pisoschi, 2012), while Brady (1982) investigated poetic use of compounds.

In addition to their role in literature, compounds have also been heavily researched in many studies related to language learning. Gleitman et al. (1970) focus their research on the interpretation of complex three-part compounds and the ability of adults to interpret such compounds. Additionally, investigations into compound words have often focused on specific registers. Agathapolou (2011, p. 11) refers to Salager's analysis of compounds in medical and general English as well as Horsella and Perez's (1991) corpus study of compounds in chemical English. The latter study also has pedagogical implications, suggesting that the most frequent semantic relations in nominal compounds should be taught from the early stages. A corpus study of English scientific texts by Moskowich et al. (2006) focuses on the word-formation process in compounds as a valuable marker of the levels of vernacularisation. Compounds have also been studies in the context of computational linguistics, based on noun compound semantics (Rosario et al., 2001; Kim and Baldwin, 2005; O'S´eaghdha, 2008; Tratz and Hovy, 2010) and the relationship between the nouns (Vanderwende, 1994; Lapata, 2002; Nakov and Hearst, 2006; Butnariu and Veale, 2008). The first group of studies investigated the semantics of nominal compounds and the semantics of their component parts, while the latter focused on the direct relationship between the nouns. Studies with the focus on the morphological developments include O Seaghdha and Copestake's (1998) study noting that with time, older, accepted compounds get somewhat lexicalized and many novel compounds develop, and Mattiello's (2008) research on slang compounds, which he considers to be a sign of morphological complexity.

Analyses of compounds can be quite complex as it is sometimes necessary to analyze word classes comprising compounds in order to distinguish between composite nominals (*some new cars, two London colleges*) and compound nouns (*shortbread, two ice-creams*). For example, according to Huddleston and Pullum (2002, p. 449), *New* and *London* are separate words, adjective and noun respectively, functioning as attributive modifiers, while *short* and *ice* do not themselves have the status of words; they are bases, adjective and noun respectively, forming parts of compound words. Analysis of compounds within a specific register can be even more challenging due to their frequency and complexity. Citing Halliday, Agathapolou (2011, p. 11) indicates that, among other features which contribute to the high lexical density of scientific English, compounds may be *'the hardest examples to process'.* Furthermore, some studies suggest that learners face difficulties with the structural form of compounds and the selection of proper contextual clues to interpret their meaning appropriately (D.Ó.Séaghdha, 2008).

While compounding in English is well researched (Lieber 2011), research into the frequency and types of compounds in economics is lacking. Economics as well as many other social sciences is characterized by multidisciplinarity (Scarpa, 2001, p. 4-5) as it combines more than one academic discipline. Economic texts are characterized by the frequent use of nominalization as the intent to avoid relative clauses and instead favor the use of paratexis, adding clauses independently one after the other (Brugnano, 2015). One of the main characteristics of

economic texts, similar to other specialized texts, is the abundant use of terminology (Cabré 1999). Because of the frequent nominalization within economic texts and a great deals of terminology, compounds are frequent in this type of texts. Former analyses have shown significant differences between the compounds in spoken and written registers. "Salager's (1984) small corpus analysis reveals that compounds are considerably more frequent in medical than in general English texts and in Biber and Gray's (2010) large corpus analysis, compounds occurred about three times more in English research articles than in oral discourse" (Agathopolou, 2011).

In view of the above, there is a need to analyze compound words in the register of economics, their frequency and types in particular. Such an analysis can prove to be of use in EFL teaching and translation studies in Bosnia and Herzegovina. The aim of this research paper, therefore, is to analyze the frequency and types of compounds occurring in a two-decade period in a selection of economic texts in order to test the hypothesis that there is a trend of increased use of compounds in economic texts, particularly the nominal compounds.

Methodology

The corpus created for the analysis of compounds is a selection of texts in economics based on two journals: "The Quarterly Journal of Economics", Vol. 110, 1995 and "The Quarterly Journal of Economics", Vol. 130, 2015. The Quarterly Journal of Economics is rated as the first out of 347 journals in the 'economics' category with the impact factor of 6662. The very first Quarterly Journal of Economics was published in 1886. It is one of the most appreciated journals in economics and it is edited by native speakers of English language at Harvard University. Thus, the influence of translation on number or type of compounds has been eliminated.

The selection of texts consists of eight texts in the field of both empirical and theoretical macroeconomics, that is, four texts for each volume. As the journal is published quarterly, the chosen sample contained a total yearly publication, that is, four texts published in February, May, August and November. The 2015 sample was taken as the newest available sample at the time and 1995 sample was taken as an older sample for comparison purposes. The two chosen samples were compared based on the number of compounds and their type and complexity.

For each text, the first step was to identify all of the compounds. However, the criteria aforementioned (ortographic, phonological, morphological, syntactic or semantic) were insufficient and there was a need for syntactic tests in many cases. Therefore, compounds were determined by following the noted criteria as well as applying syntactic tests when necessary, particularly Bauer' pronoun substitution test (2006). The test shows that phrases contain separate elements and a pronoun can replace one of them, however, with compounds that is not be the case:

e.g. She borrowed a *red dress*, not a blue one. PHRASE

I wrote it on the *blackboard*, not a white one* COMPOUND

After the texts had been analyzed and all the compounds determined, all the examples of compounds were counted. The second count was performed as well, which excluded compounds repeating throughout the chosen period. This was done in order to obtain frequency of compounds in total as well as based on single examples. The results relating to the compound count including

and excluding repetitions were presented separately for the two periods, as well as in a single table comparing the two journal volumes.

The next step was to determine the compound types according to Plag's model. The Plag's model classifies compounds into four types: nominal, adjectival, verbal and neoclassical. Further tables were then created to present compounds according to their type within both samples. Tables include the number count as well as percentages for each type. The internal structure of the analyzed compounds was presented by bracketing, while at the same time a note was made on the length of compounds in question. The length of compounds was presented in tables containing counts and percentages to illustrate the conclusion supporting the hypothesis that the recent period contains more complex compounds. Once again, tables included percentages based on the total compound count as well as the ones based on the second count excluding repetitions. This was done in order to show the effect of the inclusion of repeated compounds on the percentages indicating the complexity of compounds. Additionally, any endocentric compound types were also noted. The samples contained only a single exocentric example *a red tape* (Text 3, 1995).

Results and Discussion

The number of compounds noted in the two chosen journal volume samples is presented in Table 1.

Total count		Frequency
1995 SAMPLE	**1281**	**30.77%**
2015 SAMPLE	**2882**	**69.23%**
TOTAL	**4163**	**100.00%**

Table 1. Total count and frequency of compounds in the (1995) and (2015) samples

The results presented in Table 1 show that the number of compounds is higher in the 2015 sample, which was to be expected according to the hypothesis. The compound count for the 2015 sample is more than double compared to the compound count for the 1995 sample. When the compounds which repeat throughout a single text are excluded the results are somewhat different. Table 2 shows the total count of compounds excluding repetitions.

Total count		Frequency
1995 SAMPLE	**420**	**38.78%**
2015 SAMPLE	**663**	**61.22%**
TOTAL	**1083**	**100.00%**

Table 2. The total count of compounds excluding repetitions

Once again, the compound count is higher in the 2015 sample but the difference is smaller. This suggests that the 2015 sample contains more examples of repeating compounds. Table 3 shows the increase of compound count with the inclusion of repetitions and 2015 sample clearly shows higher increase.

	Compounds 1995	Compounds 2015	Increase %
Single count	420	663	158 %
Including repetitions	1281	2882	225 %

Table 3. Compound count increase

These figures add to the hypothesis that the compound count is higher in the more recent periods compared to the past as the compounds are becoming more regularly used and accepted and therefore more repeated. This is also supported by the previously referenced works by O Seaghdha and Copestake (1998) and Mattiello (2008).

When the focus is turned to the individual texts from the two sample periods, each text from the 2015 sample contains higher compound count compared to the texts from the 1995 and this is shown in Table 4, containing the total compound count and Table 5 containing the compound count excluding repeating examples. Moreover, based on the total count presented by Table 4, there seems to be similar frequency of compounds across texts in the 2015 with the biggest difference being 10.23% (Text 3 – 2015 – 20.44%/Text 4 – 2015 – 30.67%). In the 1995 sample the biggest difference is 26.62% (Text 3 – 1995 – 13.43%/Text 1 – 1995 – 40.05%), which is more than double between the two texts. This suggests that compounds are becoming more stable over the years so their numbers do not differ that much.

Compounds 1995			Compounds 2015		
TOTAL	100.00%	1281	TOTAL	100.00%	2882
Text 1 (February)	40.05%	513	Text 1 (February)	24.53%	707
Text 2 (May)	16.08%	206	Text 2 (May)	24.36%	702
Text 3 (August)	13.43%	172	Text 3 (August)	20.44%	589
Text 4 (November)	30.44%	390	Text 4 (November)	30.67%	884

Table 4. Total compound count across texts in the (1995) and (2015) samples

Compounds 1995			Compounds 2015		
TOTAL	100.00%	420	TOTAL	100.00%	663
Text 1 (February)	33.33%	140	Text 1 (February)	21.72%	144
Text 2 (May)	25.24%	106	Text 2 (May)	23.98%	159
Text 3 (August)	15.24%	64	Text 3 (August)	28.36%	188
Text 4 (November)	26.19%	110	Text 4 (November)	25.94%	172

Table 5. Compound count across texts in the (1995) and (2015) sample excluding repeating examples

Based on Table 5, differences are even more noticeable with (2015) sample showing only 6.64% difference (Text 1 – 2015 – 21.72%/Text 3 – 2015 – 28.36%). On the other hand, (1995) sample shows the difference of 18.12% (Text 3- 1995 – 15.21%/Text 1 – 1995 – 33.33%) which is nearly triple compared to the (2015) sample difference. It seems that the exclusion of the repeating compound examples lowers the overall frequency differences across texts in both

samples but it is still in favour of the (2015) sample. Once again, indicating that the (2015) sample contains more evenly distributed compounds across texts, which might be a result of their increasing and more stable percentages in this sample compared to the (1995) one.

When it comes to the length and the complexity of compounds the results show that the 2015 sample contains somewhat more complex compounds compared to the earlier 1995 sample. This supports the proposed hypothesis of the higher complexity of compounds with time. These results are presented in the following Table 6:

Compounds 1995	1281	100.00%
1 WORD	41	3.20%
2 WORDS	958	74.78%
3 WORDS	249	19.44%
4 WORDS	33	2.58%
5 WORDS	0	0.00%

Compounds 2015	2882	100.00%
1 WORD	37	1.28%
2 WORDS	2076	72.03%
3 WORDS	691	23.98%
4 WORDS	55	1.91%
5 WORDS	23	0.80%

Table 6. Compounds according to length – (1995) and (2015) samples

As can be seen, the 2015 sample contains a few percentages less of the 1 word and 2 word compounds compared to the 1995 sample and with the 3 word and 5 word compounds the opposite is true as 2015 sample contains more of the longer compound types. In fact, the 1995 sample contains no examples of 5 word compounds while the 2015 sample contains 23 examples. All of this confirms the hypothesis stating that the compounds get more complex with time. The only evidence to the contrary is a low decrease of the 4 word compounds (less than 1%) in the 2015 sample compared to the 1995 sample. However, the raw compound count is still higher in the 2015 period for all of the compounds and if the repeating examples of compounds are excluded the results are still the same. Table 7 contains the compound length excluding repeating compounds for both (1995) and (2015) samples.

Compounds 1995	420	100.00%
1 WORD	6	1.43%
2 WORDS	275	65.48%
3 WORDS	113	26.90%
4 WORDS	26	6.19%
5 WORDS	0	0.00%

Compounds 2015	663	100.00%
1 WORD	7	1.06%
2 WORDS	425	64.10%
3 WORDS	188	28.35%
4 WORDS	36	5.43%
5 WORDS	7	1.06%

Table 7. Compounds according to length (excluding repeating examples) – 1995 and 2015

The two chosen samples (1995) and (2015) also show further differences across the four texts. Following Tables 8 and 9 contain counts and frequencies based on length of compounds across texts in the two samples.

Compounds 1995	Text 1	Text 1	Text 2	Text 2	Text 3	Text 3	Text 4	Text 4
1-word compound	33	6.43%	3	1.46%	5	2.91%	/	0.00%
2-word compound	371	72.32%	165	80.10%	135	78.49%	287	73.59%
3-word compound	96	18.71%	34	16.50%	26	15.11%	93	23.85%
4-word compound	13	2.53%	4	1.94%	6	3.49%	10	2.56%
5-word compound	/	0.00%	/	0.00%	/	0.00%	/	0.00%
Total	513	100.00%	206	100.00%	172	100.00%	390	100.00%

Table 8. Compound count and frequency across texts according to length – (1995)

Compounds 2015	Text 1	Text 1	Text 2	Text 2	Text 3	Text 3	Text 4	Text 4
1-word compound	12	1.70%	1	0.14%	24	4.07%	/	0.00%
2-word compound	629	88.97%	403	57.41%	395	67.06%	649	73.42%
3-word compound	51	7.21%	292	41.60%	146	24.79%	202	22.85%
4-word compound	15	2.12%	6	0.85%	20	3.40%	14	1.58%
5-word compound	/	0.00%	/	0.00%	4	0.68%	19	2.15%
Total	707	100.00%	702	100.00%	589	100.00%	884	100.00%

Table 9. Compound count and frequency across texts according to length – (2015)

Raw count numbers are constantly higher in the 2015 sample but even across texts, except for the Text 1, compounds show similar patterns of distribution where 3 word compounds are higher in the 2015 sample compared to 1995 while the opposite is indicated for the 1 word and 2 word compounds. Interestingly, there was not a single example of 5 word compound across 1995 sample texts. Furthermore, there is a possible reason for the lower numbers of more complex compounds. Wasak (2016) mentions that complex compounds are usually constructed from 'strongly lexicalized noun-noun compounds' and concludes that the lack of such strongly lexicalized nominal compounds could be behind lower frequency of more complex compounds. This could especially be true in the case of economics texts as it is expected that in such texts there would be less of the lexicalized nominal compounds compared to other text types. For example, in literary texts or texts within the field of social sciences due to the nature of these texts and wide range of use of many expressions even outside of the specialized, expert circles, it would not be surprising to come across many lexicalized nominal compounds which could lead to further, more numerous and complex cases of compounding. For example, as *textbook* is highly lexicalized it is also often found in other examples such as *morphology textbook* and *grammar textbook.* Another example is that of, *second language* leading to examples such as s*econd language acquisition, second language learning, second language research* etc.

The analysis also referred to the types of compounds according to Plag's model and the results are presented in Table 10.

Type	(1995) Total		(2015) Total	
Nominal	95.32%	1221	98.78%	2847
Adjectiv al	0.31%	4	0.42%	12
Verbal	0.00%	/	0.00%	/
Neoclassical	4.37%	56	0.80%	23
TOTAL	100.00%	1281	100.00%	2882

Table 10. Frequency of compounds according to type in the (1995) and (2015)

The examples of the most frequent compounds in 1995 and 2015 texts are as follows:

Compounds 1995		***Compounds 2015***	
Autocorrelation	32	Arbitrage opportunities	23
Air pollution	8	Batch auctions	73
Contract prices	10	Batch interval	64
Contract wage	10	Discount rates	29
Corruption index	10	Heterogeneity	12
Country effect	22	Index points	10
Cross-section regression	17	Labor fluctuations	19
Excess demand	25	Labor supply	36
Inflation rate	28	Lease length	106
Investment rate	24	Liquidity provider	86
Oxygen regime	8	Market design	54
Panel estimation	15	Price differences	16
Wage contracts	12	Price discounts	25
Price index	21	Product market tightness	84
Production function	24	Property age	20
Red tape index	14	Property size	22
River basins	8	School choice program	11
Vector autoregression	16	Treatment villages	18
Water quality	10	Voucher winners	11

As can be seen from the results in both samples nominal compounds take up over 95% of the total compound count and there are no examples of verbal compounds. Interestingly, the frequency of nominal compounds is even slightly higher in the (2015) period with 98.78% compared to the (1995) period with 95.32%. This supports the overall idea that, on one hand, nominal compounding is the most productive type in the English language. Another interesting result is the presence of otherwise rare coordinative copulative compounds or, to be even more specific, dvandva compounds containing proper names. The following are sample examples of coordinative copulative compounds and dvandva subtype respectively:

Coordinative copulative compounds	**Dvandva subtype**
seller-buyer pair	*Dickey-Fuller* tests
lag-lead distribution	*Phelps-Taylor* model
impulse-response function	*Solow-Cass-Koopmans* model
output-gap coefficient	*Hadrick-Prescott* filter

Table 11. Example of compound classification from the analyzed sample – (1995) and (2015)

On the other hand, verbal compounding is often questionable and, for example, Marchand (1969) does not even agree with the idea that verbal compounding exists in the English language. While others, such as Bauer and Reunof (2001) disagree with Marchand even they stress the low frequency of verbal compounding as the examples in question are often almost limited to cases of back-formation and conversion.

Furthermore, compared to the (1995) period when adjectival compounds were represented with 0.31%, in the (2015) period that number is 0.42% but they are limited to Text 2. On the other hand, neoclassical compounds show the opposite pattern with the frequency of only 0.80% in the (2015) sample compared to the 4.37% frequency in the (1995) sample. Therefore, while nominal and adjectival compounds contain small percentage increase in the (2015) sample, neoclassical compounds are more present in the (1995) sample. This is why it is also important to include results across texts to see if the examples are contained to a single text. The frequency of compounds according to type across texts in the 1995 and 2015 samples is presented in Tables 12 and 13 respectively.

1995	Text 1		Text 2		Text 3		Text 4		Total	
Nominal	90.06%	462	99.51%	205	96.51%	166	99.49%	1221	95.32%	1221
Adjectival	0.39%	2	0.00%	/	0.00%	/	0.51%	4	0.31%	4
Verbal	0.00%	/	0.00%	/	0.00%	/	0.00%	/	0.00%	/
Neoclassical	9.55%	49	0.49%	1	3.49%	6	0.00%	56	4.37%	56
TOTAL	100.00%	513	100.00%	206	100.00%	172	100.00%	1281	100,00%	1281

Table 12. Total frequency of compounds according to type across (1995) sample

2015	Text 1		Text 2		Text 3		Text 4		Total	
Nominal	98.30%	695	98.29%	690	98.13%	578	100.00%	884	98.78%	2847
Adjectival	0.00%	/	1.71%	12	0.00%	/	0.00%	/	0.42%	12
Verbal	0.00%	/	0.00%	/	0.00%	/	0.00%	/	0.00%	/
Neoclassical	1.70%	12	0.00%	/	1.87%	11	0.00%	/	0.80%	23
TOTAL	100.00%	707	100.00%	702	100.00%	589	100.00%	884	100,00%	2882

Table 13. Total frequency of compounds according to type across text (2015) sample

The following are sample examples of the most frequent adjectival and neoclassical compounds:

Adjectival compounds		**Neoclassical compounds**	
data-consistent	2	Autocorrelation	32
time-invariant	2	vector autoregression	16
product-slack	3	reduced-form vector autoregression	1
Labor-tight	5	biodiversity	1
Product-tight	4	Endogeneity	5
		long-run exogeneity	1
		Heterogeneity	12

Table 14. Adjectival and neoclassical compounds across texts - (1995) and (2015)

When it comes to specific types of compounds, frequencies differ across texts so, for example the adjectival type is contained to a single text in 2015 and two texts in 1995 and the neoclassical type is present in two 2015 texts (*heterogeneity*) compared to three in 1995 texts (*autocorrelation*, *biodiversity*, *endogeneity*). Interestingly, adjectival compounds were often incorporated into larger units:

risk-free discount rate
asset-specific quantity
economy-wide price

However, if we focus on the neoclassical compound type that showed higher frequency in the 1995 sample compared to the 2015 sample we can see that most of the examples (49 out of 56) are contained to a single Text 1 which, as a result, impacted the final frequency percentage. Once again repetitions accounted for most of the examples, for example *autocorrelation* – neoclassical – repeated 32 times in Text 1. To further check this possible influence of a specific text type the total frequency based on compound types across texts excluding the repeating examples was determined and the results can be found in Tables 15 and 16.

1995	Text 1		Text 2		Text 3		Text 4		Total	
Nominal	97.14%	136	99.06%	105	96.87%	62	99.09%	109	98.09%	412
Adjectival	0.71%	1	0.00%	/	0.00%	/	0.91%	1	0.48%	2
Verbal	0.00%	/	0.00%	/	0.00%	/	0.00%	/	0.00%	/
Neoclassical	2.14%	3	0.94%	1	3.13%	2	0.00%	/	1.43%	6
TOTAL	100,00%	140	100,00%	106	100,00%	64	100,00%	110	100,00%	420

Table 15. Compound types frequency across texts excluding repetitions – 1995

2015	Text 1		Text 2		Text 3		Text 4		Total	
Nominal	99.31%	143	98.11%	156	99.47%	187	100.00%	172	99.25%	658
Adjectival	0.00%	/	1.89%	3	0.00%	/	0.00%	/	0.45%	3
Verbal	0.00%	/	0.00%	/	0.00%	/	0.00%	/	0.00%	/
Neoclassical	0.69%	1	0.00%	/	0.53%	1	0.00%	/	0.30%	2
TOTAL	100.00%	144	100.00%	159	100.00%	188	172.00%	172	100.00%	663

Table 16. Compound types frequency across texts excluding repetitions - 2015

We can see that all four text from both (1995) and (2015) period once again contain the highest frequency of nominal compounds (96+%) compared to other compound types, and there are no examples of verbal compounds at all. There was only a single example of an adjectival compound in Text 1 - 1995 (*data-consistent*) and it was repeated twice and only three examples of adjectival compounds in Text 2 – 2015 *product-slack* (3 repetitions), *labor-tight* (5 repetitions) and *product-tight* (4 repetitions). Similar was true for neoclassical compounds with only two actual examples in (2015) one in Text 1 (12 repetitions) and one in Text 3 (11 repetitions) both of which were the same neoclassical compound – heterogeneity. In 1995 there were several examples such as *autocorrelation, biodiversity, endogeneity and exogeneity*. Once again, as with adjectival compounds, neoclassical compounds were often used to create larger nominal compounds such as in the example *autocorrelation coefficient or* for adjectival compounds: *price-level-adjusted (value).* Therefore, because of the recursive nature of compounds very often adjectival and other compounds end up being incorporated into larger nominal structures which could potentially explain the highest frequency of the nominal compound type.

To conclude, overall, 2015 sample contains far more compounds with and without repeated examples which supports the hypothesis that the number of compounds increases in time. Similarly, while 2015 period did not contain very high frequency of more complex compounds, frequencies were still higher for this period compared to 1995 sample which for example, contained zero 5-word compounds. Furthermore, 2015 sample texts showed more stable distribution across texts. All the differences based on the total frequency, length and complexity and compound types seem to suggest that compounds are gradually becoming more complex and frequent over time which confirms the proposed hypothesis.

Conclusion

The present research focused on the analysis of the total number of compounds in the two chosen Journal Volumes and the analysis of the complexity and type of the identified compound examples. The hypothesis that the frequency of compounds in economic tests is increasing with time along with their complexity has been confirmed. The results revealed that the frequency of compounds is in fact higher in the 2015 Journal Volume, with a quite notable difference compared to the 1995 Journal Volume. The complexity of compounds has also increased with time, as only 2015 Journal Volume contained examples of the five-word compounds and a higher number of three-word compounds. The recursive nature of compounds as well as the notion that, with time,

basic compound elements get more lexicalized and as a result used more often, could explain this noted trend.

Total frequency of compounds for the 2015 sample has more than doubled compared to the 1995 sample with 69.23% compared to 30.77%. If repetitions are excluded, the frequency results are still very close with 61.22% in favour of 2015 sample and 38.78% for the 1995 sample. This brings the increase of compound to 158% with single count and to extremely high 225% including repetitions. This, in turn, shows that compounds are more frequently used in the 2015 period and, if repeated examples are any indicators, they have become more lexicalized.

Based on these results it would seem useful to investigate compound types and their classification and complexity within other areas of economics and to check whether the trend is present in general. A drawback of this research is its focus on a specific branch within economics and there are no other existing results to make comparisons across other disciplines. Similarly, because much of the current research focuses on the semantics of compounds, it would be useful to include frequency and classification analysis and to see whether there are any similarities between certain semantic interpretations and compound types based on form or complexity. Research of compounds including all of their aspects and not focusing only on semantics, forms, types or text types should be done in order to fully analyze their current position and possible trends of their development in the English language. With that in mind, this thesis provides a small contribution to research on compounds, opens further research questions and offers a basis for making more inclusive comparisons or drawing more general conclusions.

References

1. Agathopoulou, E. (2011). "EAP Materials: Insights from Corpus-Based Studies"
2. Bauer, L. (1983). *English Word-Formation.* Cambridge: Cambridge University Press.
3. Brady, C. (1982). *'Warriors' in Beowulf: an analysis of the nominal compounds and an evaluation of the poet's use of them. Anglo-Saxon England*, 11.
4. Brugnano, M. (2015). *Investigating the Language of Economics: A Linguistic Analysis and Translation of "Basic Economics"* by T. Sowell (Chapters 20 and 21).
5. Butnariu, C. and Veale, T. (2008). *A Concept-Centered Approach to Noun-Compound Interpretation*. School of Computer Science and Informatics University College Dublin
6. Cabre, M. (1999). *Terminology: Theory, Methods, and Applications*. John Benjamins Publishing.
7. Fabb, Nigel (1998) "Compounding" in A. Spencer and A. M. Zwicky (eds.) *The Handbook of Morphology*, Blackwell: Oxford, 66-83.
8. Gleitman, L. and Gleitman, H. (1970). *Phrase and paraphrase*. New York: W.W. Norton.
9. Greenbaum, S. (1996). *The Oxford English grammar*. London: Oxford University Press.
10. Huddleston, R. D., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge, UK: Cambridge University Press.
11. Kim, S. and Baldwin, T. (2005). *Automatic interpretation of noun compounds using WordNet similarity*. . In Proceedings of the 2nd International Joint Conference on Natural Language Processing (IJCNLP2005), pages 945–956.

12. Lapata, M. (2002). *The disambiguation of nominalizations*. Computational Linguistics, 28(3), pp. 357–388.
13. Lieber, R., & Štekauer, P. (2011). *The Oxford handbook of compounding*. Oxford: Oxford University Press.
14. Matthews, Peter H. (1991) *Morphology*, Cambridge University Press: Cambridge.
15. Mattiello, Elisa (2003) *Slang compounds as one case of morphological complexity*. In *Complexity in Language and Text*, Lavinia Merlini Barbaresi (ed.), 343-377. Pisa: Edizioni Plus - Pisa University Press.
16. Moskowich, I., & Garcia, B. C. (2006). *Lop-webbe and henne cresse: Morphological Aspects of the Scientific Register in Late Middle English*
17. Nakov, P. and Hearst, M. (2006). *Using Verbs to Characterize Noun-Noun Relations.* EECS and School of Information University of California at Berkeley
18. O S´eaghdha, D. (2008). *Learning compound noun semantics.* Ph.D. thesis, Computer. Laboratory, Universityof Cambridge
19. O S'eaghdha, D.,Copestake, A. 1998. *Interpreting compound nouns with kernel methods*. Natural Language Engineering 1: 1-27. Cambridge University Press.
20. Pisoschi, C. (2012). *"Compounding in Early Modern English: Shakespeare's sonnets"*. University of Craiova.
21. Plag, I. (2003). *Word-formation in English*. Cambridge: Cambridge University Press.
22. Rosario, B. and Hearst, M. (2001). *Classifying the Semantic Relations in Noun Compounds via a Domain-Specific Lexical Hierarchy.* Proceedings of the 2001 Conference on Empirical Methods in Natural Language Processing.
23. Scarpa, F. (2001). *La traduzione specializzata*. Milano: Hoepli.
24. Tratz, S. and Hovy, E. (2010). *A Taxonomy, Dataset, and Classifier for Automatic Noun Compound Interpretation.* Information Sciences Institute University of Southern California
25. Vanderwende, L. (1994). A*lgorithm for Automatic Interpretation of Noun Sequences.* In Proceedings of the 15th Conference on Computational linguistics, pages 782–788
26. Ziering, P. (2018). Indirect supervision for the determination and structural analysis of nominal compounds.

VRSTE I FREKVENTNOST UPOTREBE SLOŽENICA U TEKSTOVIMA IZ EKONOMSKOG REGISTRA

Sažetak

Tvorba složenica u engleskom jeziku je veoma kompleksan lingvistički fenomen. Analiziranje složenica je često vrlo zahtijevno jer je potrebno analizirati vrste riječi koje ulaze u tvorbu složenica kako bi se napravila razlika između složenih nominala i imeničkih složenica. Zbog složenosti tvorbe ponekad je veoma teško ispravno prevesti značenje složenice, posebno kada nema šireg konteksta. U ovom radu se istražuju frekvencija i vrste složenica u odabranim tekstovima iz ekonomskog registra prikupljenih iz časopisa iz oblasti ekonomije objavljenih 1995. i 2015. godine. Analiza obuhvata podatke o frekvenciji i poređenja različitih tipova složenica u

rasponu od dvadeset godina. Rezultati su pokazali da se frekvencija upotrebe složenica u ekonomskim tekstovima povećala, jednako kao i njihova sloenost. Najfrekventnije su bile imeničke složenice, a najmanje zastupljene su bile neoklasične složenice.

***Ključne riječi**: složenice, tvorba, ekonomski tekstovi, nominal, neoklasični*