A COMPARISON OF COMPOUNDING IN TWO VARIETIES OF ENGLISH: BRITISH AND AMERICAN

Jana Bérešová

Trnava University, Hornopotočná 23, Trnava 918 43, Slovakia

Abstract

Compounding is one of the main word-formation processes alongside derivation. In English, forming new words by combination of two or more words that function as a single word has been recognised since the history of the English language started after the arrival of the Anglo-Saxons in the British Isles. Based on the assumption that compounding in American English has higher productivity than in British English, the aim of our study is to compare the samples of compounds found in contemporary literary prose written in two varieties of English: British and American. The focus will be put not only on the way the compounds are written, but also on their frequency of occurrence, if necessary. An in-depth analysis of particular compounds excerpted from both examined materials will focus on comparing compound types based on the combination of patterns. Specific features will be discussed and commented on from the non-native English-speaking user perspective.

Keywords: *compounding, spelling and hyphenation, nominal compounds, adjectival compounds, verbal compounds*

1. INTRODUCTION

In traditional language teaching, the focus of language education was predominantly based on teaching grammar and vocabulary as discrete-unit items. Nowadays, both words and grammatical structures are used in the environment of context and lexical items operate beyond the level of single words. Due to lexical patterning, vocabulary is viewed in terms of lexical clusters. According to Schmitt (2010), when a string of words acts as a single lexeme with a single meaning, the term commonly used for these lexemes is multiword units. This term covers different kinds of word clusters the behaviour of which are subject to three criteria: institutionalisation, fixedness and non-compositionality (Moon, 1997). The degree to which multiword units are institutionalised, fixed or compositional allows of defining different categories such as fixed expressions, idioms, proverbs, phrasal verbs, etc. However, one of the more common categories are compound words. Compounding as one of the methods of new word formation is still productive and has led to the existence of a large group of compounds in the English language. English belongs to the Germanic family of languages and the most significant difference between noun compounds in German and English is that German ones commonly appear as one word, while the resulting noun chains in English typically feature spaces or hyphens between two or more vocabulary items.

Compounds belong to all major syntactic categories such as nouns (swimming trunks), adjectives (fair-haired), verbs (double-check), adverbs (hereafter) and prepositions (outside) and two main approaches to their naming are recognised. The word *compound* precedes a part of speech, e.g. compound nouns (Huddleston and Pullum, 2016) or a part of speech precedes the word *compound*, e.g. noun compounds (Biber et al., 1999). However, both approaches coincide in the philosophy that a head lexeme determines the syntactic properties of the whole lexeme. Unlike compound nouns, compounds belonging to other parts of speech are commonly written as one word or are hyphenated. The category that needs a thorough analysis are noun compounds that are written separately. Therefore, the criteria emphasised by applied linguists are stress patterns in pronunciation where the main stress is put on the first syllable in the noun compounds, and the modification potential, which is excluded in the case of the first component in noun compounds (Huddleston and Pullum, 2010). The noun compounds that constitute separated words in syntactic constructions may cause some problems as the boundaries between morphological compound and syntactic construction are unclear. The controversial cases of compounds are comprised of those in which the stress is put on the second

syllable (self-esteem) or multiple compounds (tell-tale easyJet tags). These polymorphic words can be viewed as structures involving binary sub-constituents and therefore can be analysed in this sense.

2. MATERIALS AND METHODS

To obtain comparable sets of compounds, the first step was based on selecting a proper source of excerpts. Mentioned in previous articles that deal with the role of contemporary literary prose in English language teaching (Bérešová, 2018), the criteria for our choice have become present-day English, number one bestselling authors, one from the United Kingdom and one from the U.S.A. Due to our previous experience with the quality of the language, these two authors complied with our criteria: Peter James as a UK number one bestselling writer and David Baldacci regarded as # 1 New York Times bestselling author. The second criterion was the year of publishing their crime fiction novels. Both novels were published in 2019, which can justify that the language used in the books can be labelled as contemporary English. The same genre (crime fiction) was expected to enable us to compare the use of the similar inventory of words occurring in the speeches of narrators as well as the characters. However, it is necessary to mention that our familiarity with both writers' styles had a significant influence on choosing their works for an analysis of samples excerpted systematically, consulting dictionaries, the British National Corpus and the Corpus of Contemporary American English.

2.1. Materials

While contemplating the materials that will be analysed to provide us with samples necessary for our analysis and comparison of data gained from two texts written in two varieties of English, two crime fiction novels were taken into consideration. However, the last decision resulted from the fact that books chosen for excerpting were recently published and had to be read without being aware of the plot of the story. The book written by Peter James was analysed as the first one from page 1 to page 10 to gain 68 excerpts of different English compounds. This range of pages contained 11,444 characters, which represented 2,450 words. The second analysed book was written by David Baldacci whose English is an American variety and to obtain 79 excerpts of English compounds required an analysis of 10 pages (2,807 words, 12,864 characters). However, once the number of words was balanced for both texts at the total of 2,450 words used in each text, the actual number of compounds taken from the American text was 68. The second book offered us more complex compounds, which can be broken down into binary constituents. However, we decided to ignore them due to the purposes of our research. Subtracting the words that were used in the examined texts multiple times or problematic, the total number of compounds (56) in each variety of English was analysed in depth.

The first reading of both books provided us with samples that complied with the criteria. However, after analysing them more thoroughly, some of them were excluded as they were classified as idioms or derived words in different sources of consultation. In the initial stage of excerpting samples, the focus was put on their parts of speech as the former idea was to find out the proportion of both open-ended classes such as nouns, adjectives, verbs, adverbs and closed-ended classes (pronouns, prepositions, etc.). However, the examined pages contained compounds combined with other lexemes, for example, numeral compounds such as *twenty-four seven*. Analysing such samples, it was possible to conclude that this numeral *twenty-four* belongs to the category of compounds and is extended by the expression *seven* having one meaning 'all the time' (24 hours a day, 7 days a week), functioning as an adjective or an adverb in informal English (24/7).

The second reading was aimed at justifying the initial choice of excerpts, narrowing our attention to classifying compounds according to the nature of their heads. While the first more general approach to the classification of the parts of speech was strictly based on the official identification in the dictionary and taking the context, in which the compound was used in the text, into account, the second approach was strictly oriented on the constituents and their combination forming a pattern.

Since we decided to exclude specific cases of compounds and the same compounds repeated several times, by coincidence, in both varieties we gained 65 samples that were suitable for statistical

analyses. From the formerly read ten pages of an American text it was necessary to reduce the word count that resulted in decreasing the second text by one page and a half to maintain quality. The results of a quantitative analysis and comparison arose from the data achieved under the same conditions and having common features. Our research was planned to describe a phenomenon called the word-formation process of compounding accurately and factually.

2.2. Methods

Stating the research problem and the purpose of this study, the methods applicable in the field of linguistics were planned to be used. The first step was data collection – excerpting all the expressions that might have complied with the previously mentioned characteristic features of compounds, followed by quantitative data analyses. A qualitative analysis of data was necessary to achieve a better understanding of the research problem. A collection of excerpts was analysed in detail, using the dictionary entries and the British National Corpus or the Corpus of Contemporary American English. In this stage, we could recognise inconsistency in different sources that provided us with a variety of information. During our learning and teaching processes, we have learnt that regularities in English grammar are accompanied by a large number of irregularities, and exceptions to the rule are commonplace.

As mentioned above, it was easier to distinguish compound adjectives and adverbs from phrases than noun compounds as their spelling based on two words written separately required an in-depth analysis to decide whether a sample complies with the process of compounding and a selected compound can be considered as a representative of this word-formation process. Some compound nouns needed to be checked in different dictionaries, which revealed that these sources of information are not consistent in presenting the same words in the same way. In some cases, it was needed to consult the excerpted expression in the British National Corpus (BNC) or the Corpus of Contemporary American English (CCAE) to get proper information about its frequency of occurrence.

As far as verb compounds are concerned, there were only two samples found in each variety of English and written as one word or hyphenated. On the other hand, an American text contained more compound adverbs than a British one. Prepositions were not represented in the text of British-English variety. Since some words can belong to different parts of speech, for example, *outside*, which can be an adverb as well as a preposition, in order to indicate their part of speech, the context in which such a word was used was born in mind.

From the beginning there was a tendency to exclude the compounds such as s*omeone* or *somehow* from our statistical data as they were repeated more times in both texts, commonly more in the American one. They can be classified as either an independent pronoun formed from the quantifier *some* and the numeral *one* or an adverb consisting of a quantifier and an adverb. They are part of a closed-item group and their more frequent use was not significant for our research aims. Therefore, they were calculated only once.

3. RESULTS

Compounds are commonly found in all word classes, presented clearly in a variety of theoretical grammar-reference books. Focusing on a variety of compounds in our research, the largest number of compounds occurred in the category of nouns, in which the number of compound nouns exceeded 60% of the total number of excerpted samples, followed by compound adjectives (23.21%) found in the British text. While the number of nouns in American English was lower (51.78%), the number of adjectives was higher (28.57%). Both texts contained equally represented compound verbs – in a ratio two samples in British English to two samples in American English. The representation of compound adverbs appeared a little higher in an American text in contrast with a British one. The compound numerals (3.57%) were found only in the American text. Compound pronouns (7.15%) and prepositions (1.79%) were represented equally in both varieties.

This first distribution of samples is based on their parts of speech classification arising from the text (context). The idea behind this distribution was to find out an inventory of compound types. However,

analysed words could have had more than one part of speech and therefore the part of speech in the text as well as its official label in the dictionary were used in the above-mentioned statistical analysis.

As regards spelling, the compound nouns excerpted from the British-English text written as a single word accounted for 42.86% while those written as two words were represented by 34.28% of the excerpted samples. However, in American English, the compounds written as two separate words made up 62.07% in comparison with single-word compounds that constituted 31.03%. Our assumption regarding the spelling of compounds in American English was proved.

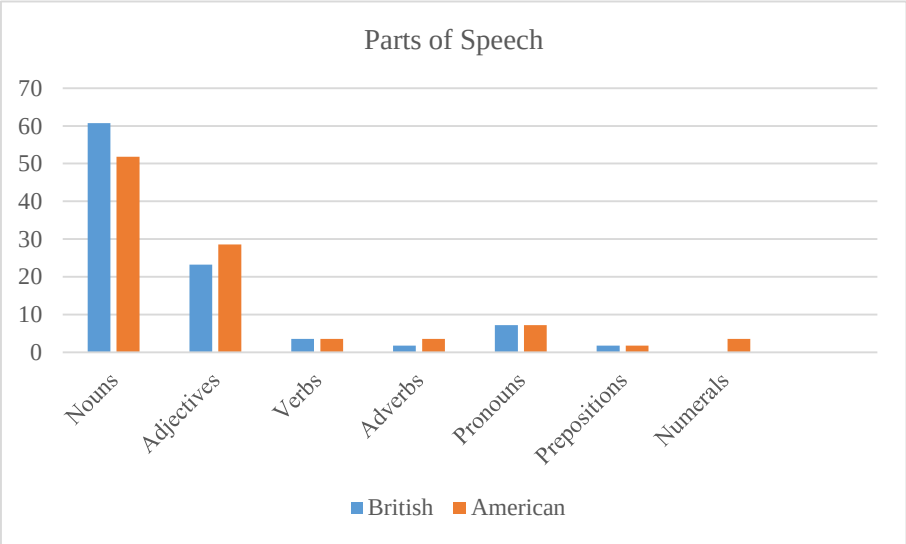

Fig. 1. The distribution of parts of speech

To distinguish the first distribution of compounds based on all parts of speech from the second one focusing on the main compound types, we decided to use different terms.

3.1. Compound Types Based on the Combination of Patterns

Although there are numerous theoretical approaches to compounding, in general, it is possible to conclude that the heads of compounds occur on the right-hand side. These heads can be modified or complemented (Carter and McCarthy, 2013). Most applied linguists distinguish three main categories: nominal compounds, adjectival compounds and verbal compounds that contain different compounding patterns.

3.1.1. Nominal compounds

Nominal compounds contain the largest variety of compounding patterns. According to Huddleston and Pullum (2016), compound nouns are divided into two main groups: noun-centred and verb-centred. The reason for avoiding the distinction between noun-centred and verb-centred nominal compounds resulted from a low number of samples we decided to comment on. Our focus is on the first constituent of nominal compounds and therefore our discussion will be aimed at three sub-groups: a noun-noun pattern, an adjective-noun pattern and a verb-noun pattern. Nominal compounds were the most numerous in the total number of all samples and the noun-noun pattern (e.g. *suitcase, ponytail, airport, apartment building*, etc. in British English and *catcall, radar detector, life sentence, law enforcement*, etc. in American English) was the most frequent sub-group within this compound type.

Figure 2 displays the distribution of the patterns. A noun + noun pattern dominates in both varieties. Although there were a number of samples that deserve commenting the focus of our article enabled us to discuss only some of them. The first one is the compound *arrivals hall*, in which the plural interpretation is restricted to the non-head and not inherited by the whole compound as it expresses the meaning that the hall is only one where incoming passengers from many different flights arrive. The

second sample found in the British text was the compound *luggage tag* that was discovered to be replaced by a more typical UK expression *luggage label* (McIntosh, 2013).

The British-English text contained more noun and adjective first constituents than the American-English one did. As regards research procedures, it was necessary to use them systematically. The American text contained expressions in which different nouns were preceded by an adjective *visiting.* One of them was given a label of a compound – *visiting hours*, the expressions that was found in the dictionaries and in the Corpus of Contemporary American English (361 hits, out of which 79 occurred in fiction). On the other hand, the expression *visiting room*, which was not an entry in any dictionary, but occurred in 129 hits with the highest number of them (39) used in fiction.

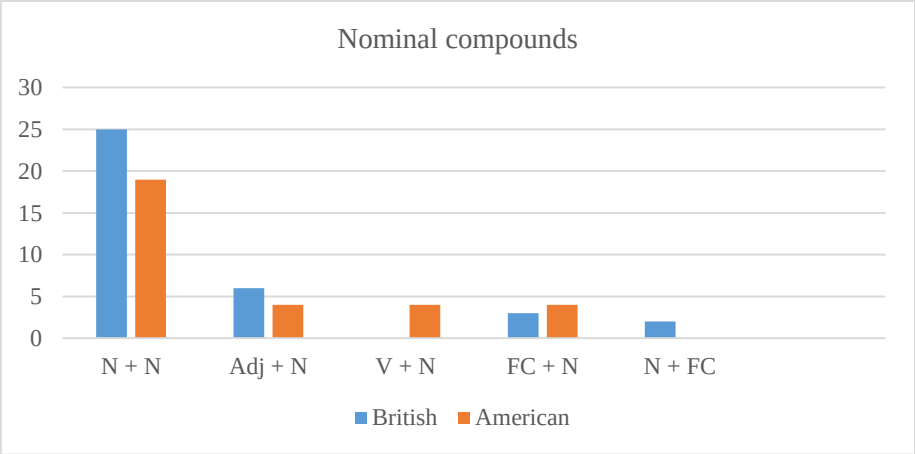

Fig. 2. Nominal compounds – the combination of patterns

On the other hand, a verb + noun pattern was represented only in American English, for example, in the compounds such as *runway* or *jumpsuit*. The compound *runway* is used in the meaning of 'a long, narrow stage used by models in a fashion show' in American English, which has its British-English equivalent *catwalk.* Figure 2 displays two categories that should be explained. Other categories of first components were comprised of compounds in which the first constituent did not belong to the main categories and the first constituent was, for example, a preposition (inmate). However, there were some samples, in which the second constituent belonged to other parts of speech such as *break-up*, which can be characterised as verb-centred nominal compound as the central element is verbal and the non-verbal constituent is a preposition.

3.1.2. Adjectival compounds

Analysing adjectival compounds, we could discover again different approaches to their determination. According to Huddleston and Pullum (2016), adjectival compounds can be divided into two main groups: adjective-centred and verb-centred. The low numbers of samples did not allow us to distinguish different specific groups. However, we will discuss some samples. Our collection of samples enabled us to distinguish the subgroups of adjectival compounds such as noun + adjective, adjective + adjective, participles as heads and other forms.

Varieties of English	**N + Adj**	**Adj + Adj**	**-ing participle as head**	**-ed participle as head**	**other forms**
British		15.36	23.1	23.1	38.44
American	6.25	12.5	18.75	37.5	25.0

Table 1. Adjectival compounds and their combinations of patterns

Comparing the data presented in Table 1, it is possible to conclude that adjectival compounds based on -ing and -ed participles as heads are the highest in numbers. While in British-English samples (e.g. *expensive-looking, serious-looking, dark-coloured, shaven-headed,* etc.) they account for 46.2%, in American English (e.g. *faint-hearted, rubber-soled, face-planting,* etc.) their representation was 56.25%. On the other hand, the differences are seen in the number of the -ed participle as head (higher in American English) and in the combinations in which the first constituents were also nouns preceding the participles in American English. Some adjectival compounds with the -ed participle can receive a passive interpretation.

The British text did not contain any noun + adjective patterns, while American English was comprised of this type of adjectival compounds, for example, *stick-thin.* On the other hand, other forms in the British text were more numerous than in the American text, for example, numerals + adjectives, e.g. *two-storey* (building). This category involved measure expressions which combined with adjectives to larger compounds such as *fifty-seven-year-old*.

3.1.3. Verbal compounds

This group was small and only four samples were found, two in the American text and two in the British one. The former one provided us with the verb *become*, which can be characterised as a combination of verb + verb and the verb *understand* (preposition + verb). The British text contained the verb *double-check* (adjective + verb) and the verb *update* (preposition + verb). Without the context, the head constituents in two examples taken from the British text could have been classified as nouns. Therefore, two-step procedures were adopted in analysing the excerpts.

3.2. Special Groups of Compounds

A special group of compounds commonly mentioned in lexicology- or grammar-reference books are compounds where at least one of the constituents is a combining form. They are usually called neo-classical compounds. Our collection of samples contained one of these compounds (photograph). Although these compounds are prominent in scientific and technical texts, some of them are commonly used.

Another group that might attract non-native speakers of English are phonologically motivated compounds. In our case, the collection of samples provided us only with one sample – *flipflops*, excerpted from the British text. While in this text, the meaning of this compound was related to a special kind of shoes, in American English and Australian English, these shoes are called *thongs.* Both constituents of the mentioned compound have their meanings and the compound is motivated by vowel contrast.

4. DISCUSSION

In consistency with the theoretical framework presented in the introductory chapter of this paper, the first reading provided us with numerous samples that had been considered relevant for the studied topic, but only the consecutive verification of them enabled us to conclude that the expressions should be classified differently, for example, as idioms (*Dutch courage* or *give birth*). The next problematic expressions were those that had stress on the second constituent. After using other criteria, these expressions were either included into a collection or excluded from the list.

In the following lines, it seems to be reasonable to mention some specific compounds belonging to our list of excerpts. One of them was the adjective *handsome*, known only from the 15th century, formed from two words *hand + some* (German handsam), being used in the meaning of ‘easy to handle or manipulate, or to deal with, or use in any way’. In the 16th century, a new meaning was recognised – ‘having a fine form/figure or beautiful with dignity’, which is the prevailing current use (Simpson and Weiner, 1991, p. 1080).

Although numerous nominal compounds have a stress put on the first constituent, there are exceptions based on meaning relationships that are temporal, locative or causative. The American text contained the word *Boston Marathon* stressed rightward, which can be considered an exception to the prevailing

compound stress rule that is commonly put on the left-hand constituent of the compound. However, this sample was not included into our collection as it was used in a multiple-word expression *a Boston Marathon bomber*.

The last issue for discussion is that compounds can be composed of more words than two. However, these multi-constituent compounds can be divided in binary compounds as its constituents. In the American text, the expression *an Oklahoma City bombing conspirator* contains four words. This multiple-word expression is comprised of a binary structure *city bombing* meaning 'a terrorist bombing in the city during which people were killed and buildings were destroyed', being more specified by the name of the city in which it happened (*Oklahoma City bombing*) and completed by another constituent *conspirator* 'the plotter who conspired with other people to do something harmful' (a truck packed with explosives was detonated outside the federal building in Oklahoma City) forming the above-mentioned expression. The process of stacking new words on an existing compound to form a new compound is common in both varieties of English.

The expressions such as *Mr Too-Good-To-Be-True* are unique and enable the language users to express the idea in a condensed, but still clearly stated message. These expressions are difficult to be translated into other languages and once it is done through free translation, the expression loses its charm.

5. CONCLUSIONS

Using numerous books related to the topic enabled us to reveal that compounding is viewed as one of the most productive word-formation processes, presented similarly, however applied linguists use different terms to cover this complex topic. Variation in the styling of compound words in English is frequent and widespread and the compound entries in dictionaries differ from one other. Bilingual dictionaries were more consistent in presenting compounds than monolingual ones. However, it was necessary to consult several ones to get a satisfactory answer to the question whether an excerpted expression can be labelled as a compound.

Reading contemporary literary prose in both varieties of English for many years, we have obtained the impression that American English uses compounding as a word-formation process more frequently and therefore we tended to find out a real situation in both varieties. However, answering such an open-ended question needs more samples taken from different books written by British and American writers to draw a valid conclusion. During our research, we discovered that compounding is a complex word-formation process and the basic definitions presented by numerous linguists are not easily applicable once working with authentic samples. Although we used all the criteria that are recommended in theoretical books, it was necessary to apply different approaches to identify the compounds.

While adjective compounds were mostly distinguishable due to their spelling based on the use of hyphens to link words, it was quite demanding to distinguish between phrases and compounds dealing with nominal compounds in American English. Each time we found out the combination of two separate words, we checked them against the criteria mentioned above. During this experience, we could discover that some criteria that seemed to be clearly stated and applicable, did not work properly in numerous cases.

All the statistical data have proved that there are not significant differences between compounds used in British English or American English. Compounding is still a productive process and compounds are occurring frequently in contemporary English as they can express the complexity of societal life.

REFERENCES

1. Baldacci, D 2019, *Long Road to Mercy*, Grand Central Publishing, New York.
2. Bérešová, J 2018, 'Stylistically Marked Words in Contemporary Literary Prose', *language, Individual & Society,* Volume 12, pp. 38-45.
3. Biber, D, Johansson, S, Leech, G, Conrad, S & Finegan, E 1999, *Grammar of Spoken and Written English,* 1st edn, Longman, Harlow.
4. McCarthy, M 2013, *Cambridge Grammar of English,* 7th edn, Cambridge University Press, Cambridge.
5. Huddleston, R & Pullum, GK 2016, *The Cambridge Grammar of the English Language,* 9th edn, Cambridge University Press.
6. James, P 2019, *Dead at First Sight*, Macmillan, London.
7. McIntosh, C (ed.) 2013, *Cambridge Advanced Learner's Dictionary*, 4th edn, Cambridge University Press, Cambridge.
8. Moon, R 1997, 'Vocabulary connections: Multi-word items in English', in N Schmitt & M McCarthy (eds.), *Vocabulary: Description, acquisition, and pedagogy,* Cambridge University Press, Cambridge, pp. 40-63.
9. Schmitt, N 2010, *Vocabulary in Language Teaching,* 11th edn, Cambridge University Press, Cambridge.
10. Simpson, JA & Weiner, ESC 1998, *The Oxford English Dictionary – Volume VI,* 2nd edn, Clarendon Press, Oxford.