

University of Pardubice

Faculty of Arts and Philosophy

Word-Formation in Fiction: Compounds in The Lord of the Rings

Bachelor Thesis

2024

Barbora Slavíková

Univerzita Pardubice
Fakulta filozofická
Akademický rok: 2022/2023

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Barbora Slavíková**
Osobní číslo: **H21213**
Studijní program: **B0231A090018 Anglický jazyk**
Specializace: **Anglický jazyk pro vzdělávání**
Téma práce: **Word-Formation in Fiction: Compounds in The Lord of the Rings**
Téma práce anglicky: **Word-Formation in Fiction: Compounds in The Lord of the Rings**
Zadávací katedra: **Katedra anglistiky a amerikanistiky**

Zásady pro vypracování

Cílem práce je ve zvoleném textu uměleckého stylu analyzovat projevy kompozice (skládání) jako slovo-tvorného procesu. V teoretické části diplomandka vymezí produktivní slovo-tvorné procesy v současné angličtině a podrobně se zaměří na složeniny, jejich typologii, morfologickou strukturu a sémantiku slovních základů. Zároveň charakterizuje typické jazykové prvky zkoumaného žánru s přihlédnutím k autorskému stylu. V praktické části klasifikuje nalezené výskyty složenin podle stanovených kritérií (např. grafické podoby, slovnědruhové platnosti, morfologie základových slov, významových vztahů mezi komponenty) a zhodnotí, jakým způsobem se tento typ pojmenování podílí na funkci stylistické.

Rozsah pracovní zprávy:
Rozsah grafických prací:
Forma zpracování bakalářské práce: **tištěná/elektronická**
Jazyk zpracování: **Angličtina**

Seznam doporučené literatury:

- Bauer, Laurie. 1983. English Word-formation. Cambridge: Cambridge University Press.
Carstairs-McCarthy, Andrew. 2002. An Introduction to English Morphology: Words and Their Structure. Edinburgh: Edinburgh University Press.
Huddleston, Rodney and Geoffrey K. Pullum. 2002. The Cambridge Grammar of the English Language. Cambridge: Cambridge University Press.
Hurford, James R., Brendan Heasley, and Michael B. Smith. 2007. Semantics: A Coursebook. 2nd edition. New York: Cambridge University Press.
Dressler, Wolfgang U., Dieter Kastovsky, Oskar E. Pfeiffer, and Franz Rainer (eds.). 2004. Morphology and Its Demarcations: Selected Papers from the 11th Morphology Meeting. Amsterdam: John Benjamins Publishing Company.
Katamba, Francis. 1994. English Words. London: Routledge.
Leech, Geoffrey and Mick Short. 2007. Style in Fiction: A Linguistic introduction to English Fictional Prose. 2nd ed. Harlow: Pearson Education Limited.
Lieber, Rochelle and Pavol Štekauer. 2011. The Oxford Handbook of Compounding. Oxford: Oxford University Press.
Mandala, Susan. 2010. Language in Science Fiction and Fantasy: The Question of Style. London: Continuum International Publishing Group.
Swales, John M. 1990. Genre Analysis: English in Academic and Research Settings. New York: Cambridge University Press.
Tolkien J. R. R. 2001. The Lord of the Rings. Seven-volume ed. London: HarperCollins Publishers.

Vedoucí bakalářské práce: **Mgr. Eva Nováková, PhD.**
Katedra anglistiky a amerikanistiky

Datum zadání bakalářské práce: **3. dubna 2023**
Termín odevzdání bakalářské práce: **1. května 2024**

doc. Mgr. Jiří Kubeš, Ph.D. v.r.
děkan

Mgr. Olga Roebuck, Ph.D. v.r.
vedoucí katedry

V Pardubicích dne 30. listopadu 2023

Prohlašuji:

Práci s názvem Word-Formation in Fiction: Compounds in The Lord of the Rings jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Beru na vědomí, že v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a směrnicí Univerzity Pardubice č. 7/2019 Pravidla pro odevzdávání, zveřejňování a formální úpravu závěrečných prací, ve znění pozdějších dodatků, bude práce zveřejněna prostřednictvím Digitální knihovny Univerzity Pardubice.

V Pardubicích dne 26. 3. 2024

Barbora Slavíková v. r.

ACKNOWLEDGEMENTS

I would like to express deepest gratitude to my supervisor, Mgr. Eva Nováková, Ph.D., for invaluable advice, constructive criticism, and encouraging support throughout the entire process of creating this Bachelor Thesis.

ANNOTATION

The aim of this paper is to analyse compound lexemes appearing in J. R. R. Tolkien's *The Lord of the Rings* from the viewpoint of word forms and word classes. The theoretical part of this thesis discusses and critically judges approaches towards compounding of various linguists in order to establish the most suitable approach to use in the analysis of the texts. In the analytical part, quantitative and qualitative research is performed and individual findings from the corpus, included in the appendix, are analysed according to the established approaches from the theoretical part.

KEYWORDS

compound, element, open, closed, hyphenated

ANOTACE

Cílem této bakalářské práce je analyzovat slovní složeniny, které se objevují v sérii knih *The Lord of the Rings* od J. R. R. Tolkiena, z pohledu slovních forem a slovních druhů. Teoretická část této práce pojednává o přístupech různých lingvistů k tématu slovních složenin a kriticky je analyzuje za cílem vybrat adekvátní a vhodný přístup k následující analýze textu.

V analytické části je proveden kvantitativní a kvalitativní průzkum a individuální nálezy z korpusu, uvedeného v dodatku, jsou analyzovány s ohledem na ustanovené přístupy definované v teoretické části.

KLÍČOVÁ SLOVA¹

složenina, skladebný prvek, open², closed³, hyphenated⁴

¹ Kategorizace forem, které jsou předmětem analýzy textu, se v českém jazyce nenachází, a proto neexistují vhodné české ekvivalenty. Z toho důvodu se bude na slovní formy složenin odkazovat v češtině anglickou terminologií.

² Open form je taková forma složeniny, ve které jsou skladebné prvky rozdělené mezerou.

³ Closed form je taková forma složeniny, ve které jsou skladebné prvky navázané bezprostředně na sebe.

⁴ Hyphenated form je taková forma složeniny, ve které jsou skladebné prvky spojené spojovníkem.

TABLE OF CONTENTS

INTRODUCTION	9
1 Fiction as a style	10
1.1 John Ronald Reuel Tolkien and his style	10
2 Word-formation	12
2.1 Compounding	13
3 The difficulties of defining a compound word	15
3.1 Structure of a compound lexeme	19
3.1.1 Closed form	20
3.1.2 Open form	21
3.1.3 Hyphenated form	22
3.2 Word class categorization and the syntactic head	24
3.2.1 Compound noun	25
3.2.2 Compound adjective	25
3.2.3 Compound verb	26
4 Analysis of The Lord of the Rings	27
5 Methodology	27
6 Word class	28
7 Word form	29
7.1 Closed form	30
7.1.1 Surnames	32
7.2 Open form	35
7.3 Hyphenated form	38
7.3.1 Neologisms in compounding	38
7.3.2 Proper nouns	39
7.3.3 Quotational compounds	40
7.3.4 Influence of fictional languages	41
7.4 Combined form	43
7.5 Inconsistencies in form	43
CONCLUSION	45

RESUMÉ	47
BIBLIOGRAPHY	51
APPENDIX 1: CORPUS	55
APPENDIX 2: ABBREVIATIONS	62
APPENDIX 3: FIGURES	63

Introduction

There are truths, that are beyond us, transcendent truths, about beauty, truth, honour, etc. There are truths that man knows exist, but they cannot be seen - they are immaterial, but no less real, to us. It is only through the language of myth that we can speak of these truths.

– J. R. R. Tolkien

The aim of this bachelor thesis is to analyse products of the word-formation process of compounding in J. R. R. Tolkien's *The Lord of the Rings*, namely, to explore the word forms from a morphological and semantic standpoint along with their distribution throughout the texts and to identify features and inconsistencies of borderline compound lexemes used by J. R. R. Tolkien.

Firstly, the theoretical part of this thesis includes a brief outline of the life of J. R. R. Tolkien and Tolkien's writing style. The paper proceeds with the description of word formation processes in the English language. Secondly, different linguistic approaches to defining compounding and the compound lexeme are discussed with the aim of establishing the most suitable and accurate approach under which the analysis is undertaken. Thirdly, the categorization of word forms of compound lexemes is explored in a detailed manner. Finally, an outline of word classes that are considered regarding the following analysis is included, together with commentary regarding the typical features of specific word classes.

In the analytical part, the researcher first initiates the research with quantitative graphs depicting the occurrences of individual compound word forms and word classes in *The Lord of the Rings*. Secondly, the findings concerning word class and word form are summarized and analysed in connection with the presented graphs. Afterwards, each category of a specific type of compound word form is discussed via analysis of findings from the corpus and is confronted with approaches from the theoretical part. Further on, a qualitative research is conducted and individual lexemes are thoroughly analysed accordingly to the established approaches and thereupon categorized into their respective groups. Finally, subcategories of chapters concerning word forms are formed discussing significant phenomena recorded during the analysis, specifically compounded surname creation, neologisms, proper nouns, quotational compounds, and fictional language influence on compounding.

1 Fiction as a style

Fiction in literature can be described as a product of linguistic and literary experimentation, where fictional characters are recorded to react to an imaginary world (Abberley 2015, 5). In fiction, there is great space and opportunity for creativity limited only by linguistic restrictions (Abberley 2015, 5), even though according to Leech and Short, these are not limitations in a literary sense, as the language is used in fiction in a way that it refers to a world described as “beyond knowledge” (Leech, Short 1981, 35). To define a style of fictional pieces of literature on a linguistic level can prove to be problematic, as listing all linguistic features of fictional works would result in a strenuous and long process (Leech, Short 1981, 46). In the end, there is no particular style that fiction adopts. According to Leech and Short, even if the author decides to structure sentences only using the bare minimum of words, it is still considered stylistically valuable on the same level any other style of writing would be (1981, 24). It can therefore be implied that works of fiction may adopt an indefinite number of different styles of writing, and so attempting to define specific features is redundant.

1.1 John Ronald Reuel Tolkien and his style

John Ronald Reuel Tolkien was born on 3rd January 1892 in Bloemfontein to Mabel and Arthur Reuel Tolkien (Carpenter 2000). From an early age, Tolkien was highly interested in languages; more specifically in creating imaginary languages (Robinson 2013, 65). Tolkien continued to pursue his fascination with linguistics and proceeded to become a highly esteemed philologist and a research associate for the Oxford English Dictionary. His interest laid both in imaginary languages as well as Old and Middle English (Robinson 2013, 66), for which he has created a vocabulary among other works (Hammond 1993, 1). His studies of Old English and other languages such as Welsh or Norse have built a strong foundation for the creation of Tolkien’s own languages. Tolkien’s fascination with Middle English led him to publish his and E. V. Gordon’s edition of *Sir Gawain and the Green Knight* in 1925 (Carpenter 2000), a poem from the 14th century written in Middle English (Davis 1967, xi), and gave lectures on Beowulf as well (Carpenter 2000). It was during the 1930s that Tolkien started working on his best-known works, the focus of this essay, *The Lord of the Rings* as well as *The Hobbit* (Carpenter 200), which gave Tolkien a historical and cultural background to his invented languages as well as exposure of his works to readers all across the world.

Tolkien’s writing style has been called archaic by critics (Drout 2004, 137). The usage of archaic expressions is likely an influence of his preoccupation with Old and Middle English. Moreover,

both Old and Middle English are sources for Tolkien's imaginary languages, and so it is expected that such expressions would be accompanied by archaic words of the Old and Middle English as well. Aside from Tolkien's invented languages, the impressive capability of distinguishing the direct speech of different races is another one of Tolkien's strengths. For example, in *The Hobbit*, which is a work tightly connected to *The Lord of the Rings*, Tolkien uses differing registers, verbal styles and choices to demonstrate the speaker's race and position within the social hierarchy (Greene 1996, 197). Whereas Thorin, the King Under the Mountain, a dwarf of high standing, speaks high-flown and formal English, a troll of low standing, for example, speaks using simple expressions as well as usually displaying a strong dialect (Greene 1996, 197).

It is needless to say that the presence of Tolkien's invented language items within the texts makes his works unique and adds to the defining features of Tolkien's writing style. Tolkien paid close attention to the aesthetics of a language rather than solely to convey the communicative purpose. He employed his knowledge of phonetics to create a language fitting a specific community, which did not yet exist, however. Tolkien proceeded then to give the imaginary languages a historical background through writing tales about the various people of a fantasy world called Middle-Earth, doing so by associating soft phonetic sounds with the noble Elves and guttural sounds with the rough Orcs and Trolls (Robinson 2013, 66).

Tolkien has created fourteen different languages, though some were given more attention than others. These languages may be involved in the creation of certain compound words, which will be thoroughly explored in the analytical part of this thesis. Tolkien himself proclaimed Quenya his biggest achievement in terms of imaginary languages. It is a language with its own grammatical rules, vocabulary, and a fictional community for which the language was created; the Elves (Robinson 2013, 66). Quenya was supposedly derived from Finnish and Greek, though adopting the spelling rules of the Latin language, all of which Tolkien displayed fluency in.

Another fundamentally important language of Tolkien's is the Sindarin language. Sindarin is a language derived from Quenya, although similarities and connections have been discovered in Welsh, which is another language that Tolkien has mastered over the years (Robinson 2013, 70). Both Quenya and Sindarin are major sources for most non-English expressions in the works connected with Middle-Earth. Tolkien distinguishes between the typical use of Quenya and Sindarin however, as Sindarin is mostly used for naming places in Middle Earth (Algeo 1985, 82), whereas Quenya is the basis for linguistic evolution (Robinson 2013, 66).

Tolkien is sometimes called the “Father of modern fantasy” (Clendenning 2003, 3), as he managed to define the genre for other fantasy authors of then and now (Trichet 2016, 94). The set patterns and themes of fantasy together with invented languages, as inspired by Tolkien’s works, have given depth to other modern fantasy works and are considered the standard for the genre nowadays (Trichet 2013, 95). Also, according to Trichet, no other fantasy author so far came close to the complexity and attention to detail when it comes to languages as Tolkien did. However, many have tried, as Mark Okrand did with the Klingon language, David J. Peterson with the Dothraki language for the *Game of Thrones* series or even Paul Frommer with his Na’vi language connected to the *Avatar* franchise (Trichet 2013, 95). The style of Tolkien’s writing is therefore highly inspirational and acts as a guideline to what a successful fantasy story should resemble.

In addition to inventing languages, Tolkien also participated in creating neologisms within the English language, such as the new fantasy races like “Ents” (Olsen 2008, 39) or “Hobbits” (Shippey 2008, 214), as well as the intriguing word *mathom*, which most likely stemmed from a word in the Old English *mathum* (Straubhaar 2007, 311). The meaning of *mathum*, “treasure” or “gift”, somewhat resembles what *mathom* meant for the Hobbits in *The Lord of the Rings*: “anything that Hobbits had no immediate use for, but were unwilling to throw away”, as Tolkien himself defines (Tolkien in “The Ring Sets Out” 2001, 7). Creating neologisms tied specifically to *The Lord of the Rings* and *The Hobbit* adds depth both to the literary and the linguistic level, which is what makes Tolkien’s works unique and extraordinary.

2 Word-formation

A language is, according to Robins and Crystal, a system of spoken and written symbols used for communication in the form of words (Robins, Crystal 2024). Complex words may be further analysed in detail and divided into smaller items called “morphemes”, which are the smallest units that carry meaning (Plag 2002, 12). Words that do not lie within this category are called “monomorphemic”, which means that they cannot be further dissected into smaller units and that they only consist of one singular morpheme (Plag 2002, 12). This essay will specifically focus on morphologically complex words.

In order for complex words to form, prescribed processes for the creation of them need to exist. Such phenomena are called “word-formation processes” (Plag 2002, 12) or simply “morphological processes” (Bauer 2004, 73). The processes, as categorized by Plag, can either be of concatenative or non-concatenative nature (2002, 15). Concatenative word-formation

processes consist of derivation, back-formation, and compounding, as by definitions the basis of these processes is linking bases and affixes together to form a new lexeme (Plag 2002, 15). On the other hand, non-concatenative processes do not form new words by adding new morphemes, but rather either deleting a part of a word or leaving the original word with no changes of their written form whatsoever (Plag 2002, 15).

Back-formation and derivation are in a way similar, as both are connected to the use of affixes. However, while back-formation deals with the deletion of affixes, or affix-like parts of lexemes (Bauer 2004, 21), through derivation one can create a new word by annexing either a prefix or a suffix (Katamba 2005, 185). Derivation is often referred to as affixation (Bauer 2004, 15) as the term “derivation” may sometimes be used to refer to a broad spectrum of word-formation processes by some linguists (Bauer 2004, 38-39). The process of compounding is the most productive of all word-formation processes, although a highly controversial one (Plag 2002, 169). Compounding will not be discussed in detail yet, as this process will be the sole focus of the rest of the thesis.

A process which does not require any changes of the bases is called “conversion”. It is possible to change the word class of a lexeme through conversion (Plag 2002, 15). To illustrate, the lexeme *water*, as stated by Plag, can be implemented either as a noun or a verb (2002, 15). Additionally, there are multiple processes which consist of deleting parts of the lexemes in order to create new ones. An example of such process is clipping, typically used for shortening of personal names, such as *Elizabeth* being cut to *Liz*, and so on (Plag 2002, 16). However, there are cases of clipping used to reduce long common expressions, such as *telephone* being shortened to *phone* for convenience purposes (Bauer 2004, 30). In addition, it is also possible for clipping and compounding to co-exist, according to Bauer, as he claims the existence of so called “clipping compounds”, where at least one element is a result of the clipping process. A suitable example of this is the lexeme *e-mail*, the first element being a product of clipping of the lexeme *electronic* (Bauer 2004, 30). Other cases of non-concatenative word-formation processes are acronyms and abbreviations (Plag 2002, 17).

2.1 **Compounding**

In the English language, there are multiple word formation processes that people can use to create new words, some being more flexible than others. There are differences in approaches to the categorization of these processes amidst linguists, as word formation processes exist on a

broad spectrum and there is no universal consensus that linguists can seem to agree on. One of the controversial topics of this area of linguistics is how to categorize compounding.

For example, Marchand (1957, 13) claims that there are only two categories of processes: derivation and expansion. Compounding, therefore, is not considered a part of the main categorization, however, it is not completely overlooked either. It is hidden within the expansion process, described as follows: “Expansions of which both the determinatum and the determinant are words (such as steamboat, blackbird, colour-blind, overdo) are called ‘compounds.’” (Marchand 1957, 14)

To clarify the terminology used, determinatum poses as the “head”, the dominant element of a word classified as an expansion (Marchand 1957, I). In addition to this, Katamba specifies the head as being the element in a compound on the right-hand side, the element that determines the syntactic properties of the whole compound word (Katamba 2005, 50). Plag also states that it is a proven theory in English that the head of a compound is always on the right-hand side (Plag 2002, 189).

On the other hand, the determinant resembles the modifier of the head, or as Marchand calls it; a satellite (1957, I). It adds restrictions and modifications to the determinatum (Marchand 1957, 14). If there be a word in which only the determinatum is an independent word but not the determinant, the word would remain a plain expansion. Consequently, words that contain a determinatum which is not an independent word, but an affix, are categorized as derivation.

Similarly to Marchand, Katamba (2005, 55) also supports the idea of only two main word-formation processes, though the categories possess different names: inflection and derivation. In this case, Katamba decides to categorize compounding within derivation, which would seem to go in contrast to what Marchand believes. However, despite the approach of referring to these processes, Katamba holds a very similar belief as Marchand does. What Katamba calls inflection, Marchand calls derivation, and what Marchand refers to as expansion would be the same as derivation for Katamba. For Katamba, derivation means combining lexemes which are already established in the English language to create new lexical items (2005, 41). Therefore, derivation means creating new lexical words, whereas inflection is restricted only for grammatical words. Consequently, compounding logically falls within the category of derivation in this case.

This thesis will, however, choose a different approach: that of compounding being one of the main word-building processes and not a subcategory of another. The approach of Ingo Plag has

been chosen for this study for the reason of dividing derivation and compounding, as derivation has shown to be an ambiguous term between linguists. The position of compounding in word-formation processes can be seen in the following graph.

[1]

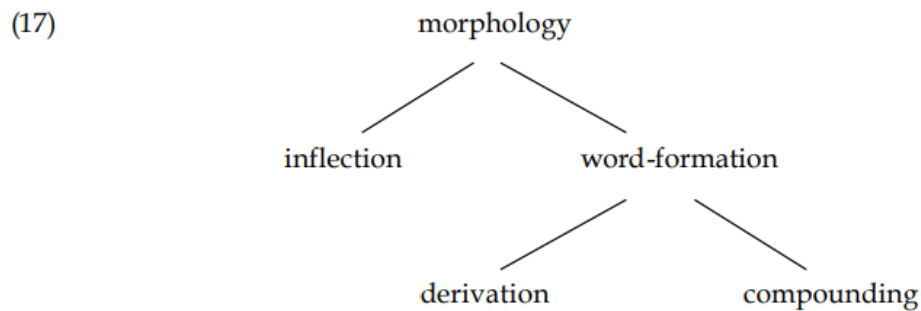


Figure 1 Scheme of the realm of morphology (Plag 2002, 22)

Plag chooses to fully separate inflection from word-formation, which allows derivation and inflection to include a narrower selection of lexical items. Derivation in Plag’s work includes all the other different processes for word-building and compounding here is reserved only for the merging of bound roots of words as Plag describes (2002, 92). As compounding is here categorized as separate from inflection and derivation, this thesis will continue to treat it as such to avoid misunderstanding.

Additionally, since linguists choose to refer to the parts of a compound differently, it is important to establish the terminology for this thesis to avoid confusion. Both Katamba (2005, 49) and Plag (2002, 170) choose to define the individual parts that form a compound not as words but rather as elements. This is because not all lexemes that are considered compounds must necessarily contain only independent words. This problematic regarding compounds containing non-word elements will be dealt with in the following chapter discussing compound words and their definition. For this reason, this essay will continue to use the term “element” to refer to the individual parts of a compound for the sake of clarity.

3. The difficulties of defining a compound word

Creating a universal definition of a compound word that all linguists can agree on has proven to be a difficult task, as no criterion applies to every word that is considered a compound. There are many different criteria and characteristics that can aid in identifying whether something is or is not a compound word. To be able to characterize a compound word, it is necessary to

identify its individual parts, which can prove to be just as difficult as defining the compound word as a whole.

Plag defines a compound as being a combination of two or more root morphemes. On top of that, Plag claims that compound lexemes containing more than two elements and two-element lexemes work accordingly to the same structural and semantic patterns (2002, 182). It is specifically stated by Plag that it is important not to classify a compound lexeme as merging two words into one, as that would not apply to a number of lexical items that cannot be overlooked. For example, as Plag illustrates, the word *biochemistry* is considered a neoclassical compound word even though *bio-* is not an independently functioning word nor an affix either. (Plag 2002, 175) However, *bio-* is categorized as an affix in a relatively recent etymological study by Gerard de Melo (2014 5.2) as well as in the Online Etymology Dictionary, later referred to as OED, which would make the word *biochemistry* a product of derivation.

There seems to be a solution to this ambiguity in defining the prefix *bio* and other similar affixes, such as *geo-*, *anthro-*, etc. Booij has introduced the term “semi-affix” to refer to individual elements of neoclassical compounds (2009, 208). This thesis will not be dealing with neoclassical compounds, however, and for this reason there will be no attention directed towards semi-affixes.

The idea that parts of a compound word do not have to be individual words raises the question of how to distinguish between an affix and an element that is not an independent word. This is a grey area in morphology as there is no absolute rule nor list of affixes that linguists can rely on because of the constant evolution of the English language. De Belder (2011, 282) states that there is currently no clear syntactic distinction between derivation and compounding.

As stated earlier, linguists can often disagree on what is considered an affix or an independently functioning word. To expand on that, there are certain borderline cases of lexemes belonging either to affixation or compounding (Algeo 1991, 7). Algeo illustrates this with the lexeme *happy*. The semantic content of *happy* is “feeling of joy and pleasure”. However, when *happy* appears in compounds such as *trigger-happy*, *power-happy*, or *headline-happy*, the semantic meaning shifts to “tendency of being impulsive or obsessive about something” (Algeo 1991, 7). The semantic shift and the productivity could indicate that *happy* possesses the properties that a product of derivation would have.

Productivity, according to Laurie Bauer, is a morphological feature of certain words to form new words through various morphological processes (2004, 87). If a word has a high

productivity, the lexeme will be involved in the creation of new expressions. Usually, the productive elements tend to appear with certain groups of words. Bauer illustrates this on the suffix *-ation*, which is likely to merge with a verb that ends in *-ise*. On the other hand, it is highly unlikely for it to be connected to a verb ending with the *-en* suffix (Bauer 2004, 88).

To return to the matter of the lexeme *happy*, it works as an independent lexeme within sentences, which an affix is incapable of (Algeo 1991, 7). It can be stated that the lexeme likely experienced the process of grammaticalization. Grammaticalization refers to the circumstance of turning a lexeme into a grammatical expression (Shustova, Osheva, Klochko 2017, 35). The lexeme *happy* has turned into an expression behaving like an affix and affixes are considered to be grammatical elements according to Bauer (2004, 13). A part of the process of grammaticalization is a so called “desemantization”, which is what has been described as a shift in semantic meaning and is what has occurred regarding the meaning of *happy* (Shustova, Osheva, Klochko 2017, 35).

Additionally, there is the issue of distinguishing a compound from a phrase. This affects the open compound words with a gap in their structure, which will be further examined in the following chapters. First, it is necessary to define the term “phrase”. According to Katamba, a phrase is “a syntactic constituent headed by a lexical category” (2005, 187). Rules for distinguishing compound lexemes from phrases have been analysed and tested by Altakhaineh and the most reliable criterion for defining compoundhood according to the study (2016, 82) is adjacency. Lieber and Štekauer have discussed the impossibility of insertion of another element within a compound lexeme in defining the Lexical Integrity Principle (2011, 292), as well as Altakhaineh, referring to it as a rule of insertion in his work (2016, 68). It has been contrasted with inserting elements within phrases which has proven to be possible (Lieber, Štekauer 2011, 13). Lieber and Štekauer illustrate their statement on an example lexeme *blackbird* and the following phrase *a black bird* (2011, 13).

[2]

<i>a black bird</i>	→	<i>a black ugly bird</i>
<i>a blackbird</i>	→	<i>an ugly blackbird</i>

Figure 2 Illustration of the rule of adjacency

Regarding the compound, it is only possible to modify the lexeme as a whole; it is not possible to insert *ugly* in between the individual elements. For relevance, an open compound word from

The Lord of the Rings has been chosen to illustrate this phenomenon further as it is the case of an open compound that resembles the structure of a phrase more clearly. In open compounds, it is especially important to adhere to the rule of adjacency, as in closed compounds it is not as ambiguous to determine the compoundhood. The lexeme *a blackberry tart* has been defined as an open compound. It is possible to modify the compound as one unit, such as “*a sweet blackberry tart*” but it is impossible to insert the modifying adjective within the structure of the compound, such as “*a blackberry sweet tart.*”

Word stress assignment plays a role in defining compound structures as well because of the so called “compound stress-rule”. According to this rule, in compound words it should always be the first element or base which carries the primary stress. (Plag 2002, 175) On the other hand, phrases follow the nuclear stress rule according to which the stress is assigned to the constituent on the right-hand side of the phrase. Lieber, Štekauer (2011, 254) as well as Altakhaineh (2016, 62) identify with the same definition. If this rule were absolute, it would result in a clear distinction between phrases and open compounds.

However, exceptions can be found as a significant amount of compound words follow the nuclear stress rule, although they are still considered to be compounds, such as the words *apple pie* and *summer night* (Plag 2002, 177). This is why distinguishing between a phrase and a compound that does not follow this rule can be difficult or near impossible. In such cases, a more detailed analysis adhering to the rules for compoundhood (Altakhaineh 2016, 61–76) needs to be undertaken in order to retrieve a satisfying answer.

In addition to these rules for compoundhood, there is the case of referentiality that needs to be discussed. According to Altakhaineh, non-referentiality is a common feature of compounds (2016, 69). Despite that, there are instances where the non-head element of a compound is a proper noun, meaning it addresses a specific or a unique referent (Altakhaineh 2016, 69–70). To illustrate such structures, example lexemes *Beatles fans* and *earth science* have been chosen. The first element of the former lexeme is a proper noun used to refer to a specific musical band “Beatles”, so the whole compound is considered referential. The latter does not contain a proper noun; however, the first element refers to a unique entity, thus it shows referentiality (Altakhaineh 2016, 69-70).

In conclusion, the mentioned rules for compoundhood are not entirely absolute, but they are useful tools for linguists to use in identifying phrases and compound lexemes. The analysis of *The Lord of the Rings* will rely heavily on these rules for compoundhood from Altakhaineh.

3.1 Structure of a compound lexeme

To recognize whether a lexeme is a compound or not may be difficult in the case of open compounds. There seems to be ambiguity between the linguists when it comes to defining specific categories of compounds. The three understood categories of forms are closed, hyphenated, and open. However, some linguists define these forms differently by structuring the forms in a hierarchy, such as Constant et al. (2017, 841).

[3]

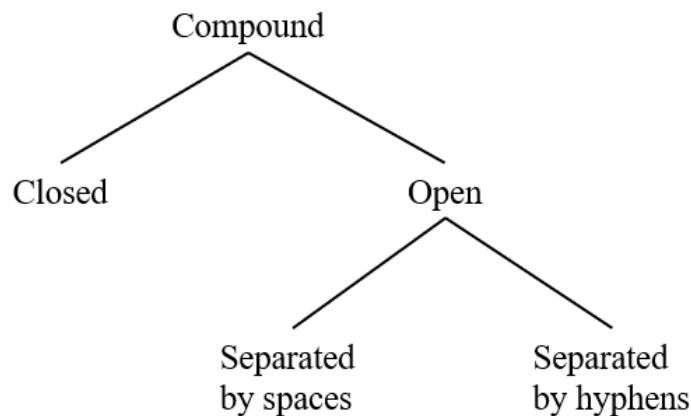


Figure 3 A hierarchy of compound forms created according to the approach of Constant et al.

According to Constant et al. (2017, 841), the compounds in hyphenated form are a subcategory of open compounds and its co-hyponym category are compounds separated by spaces. Once the individual elements of a compound lexeme are not immediately juxtaposed and there is another element between them, either a space or a hyphen, these compounds would be considered open compounds. On the other hand, Lieber and Štekauer do not follow this hierarchy and address all three forms equally. They do not mention the term “open compound” but they do take it into account, though describing the form as a compound with “spaced off constituent elements” (Lieber, Štekauer 2011, 7), the hyphenated compounds being a separate category.

Additionally, there are linguists which briefly acknowledge open compounds but do not go into further detail as to define the relation of open compounds to other forms, as for example Frank A. Smadja and Kathleen R. McKeown define an open compound as an uninterrupted sequence of words (Smadja, McKeown 1990, 253). Even though they do not talk about hyphenated compounds, by the definition it could be implied that hyphenated compounds are excluded

because the usage of hyphen would disrupt the said uninterrupted sequence of words. Therefore, it can be implied that Smadja and McKeown perceive hyphenated compounds to be outside the category of open compounds despite the fact that the hyphenated compounds have not been mentioned.

To comment on the last remaining form, closed compounds are vastly different from open and hyphenated compounds. A closed compound can be described as two words that are essentially written as a singular lexeme with no interjection (Aditiawarman, Kartika, Rahmat 2022, 6281) which is what distinguishes them from other forms. The use of closed forms is usually associated with words which are already well-established within the English language, although the level of establishment of a lexeme is not an easily determined criterion (Katamba 2005, 8). Yet, this notion is supported by Schmid and their idea of a degree of entrenchment of lexemes. Entrenchment, as defined by Schmid, is a claim by which speakers almost automatically use specific units for utterances because they are deeply “entrenched” in their memory (2007, 118). This can be accomplished by encountering such units often and repeatedly (Langacker 1987, 59). It can be assumed that for the reason of entrenchment, the compound form of the lexeme *to-morrow* is no longer used (Katamba 2005, 7). The English language has since evolved, and the closed form *tomorrow* now has a higher degree of entrenchment compared to, the now old-fashioned lexeme, *to-morrow*.

It has been chosen for this essay to acknowledge the differences between compounds where elements are separated either by spaces or by a hyphen and the compounds of a closed form. The term “open compound” will refer specifically to those separated by a space and the latter will refer to hyphen-connected elements. Additionally, all three forms of compounds will be understood as being on the same level, none being superior to the other.

3.1.1 Closed form

A compound in a closed form is a lexeme in which two elements are juxtaposed without any interjecting elements between them (Aditiawarman, Kartika, Rahmat 2022, 6281). These compounds may also be referred to as “solid compounds” (Fernández-Domínguez 2010, 206). Lexemes that are compounded in the closed form tend to be well-established within the language. It is assumed that an average language user understands their semantic meaning on account of high frequency of use as well as the overall semantic transparency (Katamba 2005, 8).

The closed form is restricted in a way that the hyphenated and open form is not. There seems to be a limit on how many elements there can be in a compound lexeme. In closed compounds, there are usually only two elements connected, occasionally three, whereas in lexemes like *forget-me-not* or *leave-me-alone look* there are more than two elements joined together by a hyphen (Constant et al. 2017, 841). This restriction is in place possibly for the reason of transparency of meaning, as Plag states that juxtaposing more than two elements could result in confusion and the lexeme might not be easily comprehensible as well as producible (Plag 2002, 172). As for open and hyphenated compounds, there is no limit on the recursivity, although it is discouraged to use exceedingly long constructions for the reason of intelligibility as well (Plag 2002, 172).

3.1.2 Open form

The open form is a highly ambiguous compound form, the reason being that not all compounds of this form adhere to the rules discussed by Altakhaineh (2016, 61). It has been mentioned in Chapter 3 that not all compounds follow the compound stress rule, adhering to the nuclear stress rule instead, which is typical for phrases.

Compound words can often co-exist in multiple forms without solely one correct written form, which implies a freedom of choice of word forms. For example, the results of analysis from Fernández-Domínguez show that the lexeme *wheelchair* can appear in every form. For the purposes of this thesis, consider lexemes 3a, 3c, 3d.

[4]

(3a)	wheelchair	602
(3b)	wheelchairs	178
(3c)	wheel-chair	10
(3d)	wheel chair	4
(3e)	wheel chairs	3
(3f)	WHEELCHAIR	797

Figure 4 Examples of the lexeme *wheelchair* in various compound forms (2010, 206)

Though the findings of the hyphenated and open forms of the lexeme are significantly lower than of the closed form, the lexeme *wheelchair* can and does appear in all forms, similarly to many other compound lexemes, such as *street lamp* and *street-lamp* (Katamba 2005, 7).

Regarding the findings of open compound forms in texts, it can be claimed that the open form is the least prominent. However, this claim is directly contradictory to what Algeo says about word forms. According to Algeo, the open form is the most prominent compared to the less frequent closed and hyphenated form (1991, 7). This statement however does not reflect the reality of the modern English, as the number of closed compounds has significantly increased throughout the years, possibly because of lexemes becoming progressively more established within the language (Katamba 2005, 8). The change of preferred compound forms could be ascribed to the change of language throughout time, as Algeo claimed the open compound form the most frequent in the years 1941–1991, whereas Katamba has claimed the closed compound group the most prominent later in the year 2005.

The issue of defining compound lexemes of open form has been already briefly discussed in Chapter 3, especially concerning the problematic of distinguishing between a phrase and an open compound.

3.1.3 Hyphenated form

A hyphenated compound is a lexeme in which individual elements are joined together with a hyphen (Constant et al. 2017, 841)⁵. The usage of a hyphen in compound lexemes seems to be tied to how established is the lexeme within the English language. This form is usually used in fairly new or even created for one specific situation to convey a particular meaning (Katamba 2005, 8). On the other hand, lexemes that are used frequently often appear in the closed form.

On top of that, lexemes that used to contain a hyphen sometimes experience a loss of hyphenation throughout time for the reason of compactness and simple spelling; for example, the usage of *to-morrow* is no longer common, as stated in Chapter 3.1, and has been exchanged for a more convenient spelling *tomorrow* (Katamba 2005, 8).

Hyphenated compound lexemes are often used as modifiers in English. According to Magat (2014, 153), compound modifiers that precede a noun in a sentence typically appear in the hyphenated form. Although it is a straightforward and simple rule according to Magat, its application is not perfectly consistent in English and exceptions are found nonetheless (Magat 2014, 153). However, the exceptions seem to be restricted to well-established structures in

⁵ Although the definition of a hyphenated compound has been taken from Constant et al., it is important to note that this essay will not adopt the hierarchy of hyphenated compounds being an inferior category of the class of open compounds that is presented in the source.

English, for which Magat illustrates with several open compound structures, such as *high school student* or *criminal defense lawyer* (2014, 153).

The hyphenated compound form allows for many elements to be joined by hyphens into extraordinarily long lexemes. These lexemes tend to be manufactured for a very specific purpose and most of the time, they function as disposable items and are not used repeatedly (Katamba 2005, 105). These compounds usually internally follow syntactic rules, as they are created to include information that would normally be a content conveyed by a longer phrase or a clause. The structure of the compounded phrase remains identical to the original and the spaces between individual lexemes are exchanged with hyphens, which creates a single lexical unit out of a phrase (Katamba 2005, 8).

This is called a “creative use of hyphenation” and Katamba sets as an example the lexemes *simple-to-serve* or *fresh-from-the-farm*. (2005, 8). It is worth noting that both the example lexemes are used in sentences as modifiers.

[5]

<i>simple to serve</i>	→	<u><i>simple-to-serve</i></u> dishes
<i>fresh from the farm</i>	→	<u><i>fresh-from-the-farm</i></u> eggs

Figure 5 Illustrations of the hyphenation process from the original phrase

It can be assumed that if one decides to manufacture such a compound structure from a phrase, there is a significant chance that the usage of such a compound lexeme will acquire the role of an attribute within the utterance. However, this is to be understood as a tendency rather than a rule. Additionally, Plag mentions the lexeme *jack-in-the-box*, which follows the rules of syntax internally and did arise from a phrase, but it does not necessarily function as a modifier (2002, 174).

Plag’s approach to creative compounding is that such lexemes should not be considered compounds for their internal syntactic properties and should therefore be considered “lexicalized phrases” instead (2002, 175). For this essay, the approach of Katamba has been chosen and such creative structures will be regarded as compound lexemes. However, the term used for referring to these structures has been adopted from Ryšavá (2012, 2) and Rybínová (2021, 7). Instead of “creative use of hyphenation”, the term “quotational compound” has been

chosen for this paper, as most of the research devoted to these structures has operated with the term.

3.2 Word class categorization and the syntactic head

The topic of word class of compound lexemes must be preceded by a relevant and brief summary of the nature of a syntactic head. In a compound lexeme, the relations between the whole compound and its elements are described by Bauer as hyponymous, where the whole compound is a hyponym of its syntactic head (2004, 51). The syntactic head is found on the right-hand side of the compound (Lieber, Štekauer 2011, 6) and its function is to determine morphological features of the whole compound (Plag 2002, 174), including the gender, inflection, and the word class of the compound (Bauer 2004, 52). The function of the syntactic head regarding the transferring of features on the compound is called “feature percolation” (Plag 2002, 174).

It has been clarified that the syntactic head is responsible for the compound lexeme’s feature percolation (Plag 2002, 174). The head can be either a noun, an adjective or a verb (Bauer 2004, 52), which already categorizes compound lexemes into three groups according to the word class of the head. Within each category, there are further arrangements to be made regarding the word class of the modifying lexeme since the word class of the non-head element does not influence the word class of the whole compound. This is supported by Plag and illustrated by a table containing all possible compound types, including subtypes based on the word class of the non-head element of a compound (2002, 185).

[6]

(18) Inventory of compound types, revised

	noun (N)	verb (V)	adjective (A)
noun	<i>film society</i>	<i>brainwash</i>	<i>stone-deaf</i>
verb	<i>pickpocket</i>	<i>stir-fry</i>	-
adjective	<i>greenhouse</i>	<i>blindfold</i>	<i>light-green</i>
preposition	<i>afterbirth</i>	-	-

Figure 6 A table of compound types based on word class from Plag

According to Delahunty and Garvey (2010, 134), there are more categories to recognize regarding the word-class categorization of compounds than Plag suggests. Apart from

neoclassical compounds, which are not expected to appear in *The Lord of the Rings*, Delahunty and Garvey also mention compound adverbs, though comprising only a minor category (2010, 134). On top of that, Bauer claims that there is a fifth category: that of prepositions (Bauer 2004, 33). This essay will continue to address word class categories of compounds as a group of nouns, verbs, adjectives, and adverbs. Prepositions will not be subject to analysis as they lie within grammatical words and the focus of this thesis is aimed at lexical words.

3.2.1 Compound noun

Compound nouns, or “nominal compounds” (Plag 2002, 185), pose as the largest represented group of compound lexemes according to Algeo, who claims that around 90% of all new compounds are nouns, verbs and adjectives forming a minority amongst this type of word formation process (Algeo 1991, 7). The reasoning for nouns being the most prominent group of compounds is the favouritism of nouns in the English language (Algeo 1991, 7). The meaning of a sentence or a clause usually results from nouns being the carriers of the semantic content. Nouns usually pair with “dummy verbs” such as *have*, *do* or *take*, the meaning of which is highly flexible and dependent on the surrounding lexemes (Algeo 1991, 7). Such combinations often supersede the usage of lexemes such as *dine* by using the phrase *have dinner* instead (Algeo 1991, 7). Algeo’s argument for the semantic superiority of nouns is that “there are more new things to name than there are new events or qualities” (1991, 7), claiming that for nouns there is a much more room for expansion and creativity than there is for adjectives or verbs.

To categorize compound nouns according to Plag’s table in Figure 6, it is the most flexible category when considering the range of possible word classes of the non-head elements (2002, 185). A noun as the head element can be combined with another noun, an adjective, a verb, or a preposition. The noun-noun compound structure is the most prominent among compounds (Plag 2002, 185). Compounds of the adjective-noun structure are a numerous group as well. Finally, there are verb-noun constructions and preposition-noun compound lexemes, which are not as common and are mostly produced by implementing two specific prepositions: *after* and *under* (Plag 2002, 194).

3.2.2 Compound adjective

Compound adjectives, also called adjectival compounds (Plag 2002, 194), are compounds that contain an adjective in the position of the syntactic head, meaning on the right-hand side of the lexeme. There are also numerous sub-categories within this classification to be addressed.

Semantics play a role in the understanding of one of the sub-categories, the noun-adjective compound lexemes. The noun can either serve as a modifier or as an argument of the following adjective (Plag 2002, 194–195). For example, the lexeme *dog-lean* is based on a comparison relationship, the semantic understanding being “lean as a dog”, whereas in the lexeme *structure-dependent*, the noun acts as an argument of the adjective, meaning “dependent on structure” (Plag 2002, 195). Semantic differences within compound structures, such as noun-adjective sub-category provides, are not taken into consideration in Delahunty’s and Garvey’s compound classification.

Adjective-adjective compound lexemes are not as common as the former sub-category (Plag 2002, 195), yet there are two types of such structures. In lexemes such as *icy-cold* or *blueish-green*, the first adjectival element of the compound stands in the position of a modifier (Plag 2002, 195). Its function is to further specify the semantic content of the latter adjective. According to Plag, there are also adjective-adjective compounds which seem to function as “copulative compounds”, meaning that both elements of the compound are equal, neither being superior to the other, displacing the syntactic head completely. Examples of this are lexemes *sweet-sour* or *bitter-sweet*, as illustrated by Plag (2002, 195).

3.2.3 Compound verb

Compound verbs, or verbal compounds (Plag 2002, 197), are lexemes with a verb as a syntactic head. Such compounds can take nouns, adjectives, and verbs as the non-head elements. However, according to Plag, many compound verbs are not an immediate result of compounding but rather of conversion or back-formations (2002, 197). Bauer states that conversion is a derivational process that includes changing the word class of the lexeme without no change in its form (2004, 36). To comment on back-formation, Bauer defines back-formation as a type of a word-formation process that consists of removing an affix from a lexeme, and thus creating a new lexeme (2004, 21). Additionally, Plag restricts back-formation only to removing suffixes (2002, 48). Interestingly, Katamba seems to contradict his own statements when it comes to back-formation, since the definition is identical to Bauer’s, however, Katamba then proceeds to call the deleted affixes “a peripheral part of a word which is wrongly analysed as a suffix” (2005, 184). In this thesis, such removed parts will be considered affixes, as Bauer suggests, if they should be further mentioned.

To return to the topic of compound verbs, verb-verb structures are, according to Plag, not subject to any other word-formation process other than compounding (2002, 198). In fact, verb-

verb compound lexemes seem to function as copulative compounds, as both elements of the compound have their own independent semantic meaning and one is not the head of the other (Plag 2002, 198). An example of such a verbal compound is the lexeme *stir-fry*, describing the action of *stirring* and *frying* (Plag 2002, 198).

4. **Analysis of The Lord of the Rings**

The aim of the thesis is to examine the frequency of occurrence of various types of compound structures within Tolkien's *The Lord of the Rings*. The original scope of the analysis was intended to be the complete trilogy apart from the appendix. However, the scope has been reconsidered throughout the research as the number of findings has proven to be so extensive that it would be impractical to include more than the first three of the seven books. For this reason, *The Return of the King: The War of the Ring*, *The Return of the King: The end of the Third Age* and *The Two Towers: The Ring Goes Fast* will be excluded from the analysis and will not be dealt with for the sake of compactness.

The analytical part of this thesis is structured to involve a detailed analysis of compounds in order to categorize the findings into specific groups, each focusing on a different word-form or word-class of the lexemes. Chapters 3.1 to 3.1.3 will cover word-form classes and in Chapters 3.2 to 3.2.3 the word-class categorization will be discussed.

5. **Methodology**

The researcher will extract 200 compound words from the first three books of *The Lord of the Rings* and input the lexemes into a corpus which will be available in the Appendix. As the lexemes surrounding the compound in a clause will not be analysed in detail, the findings will be presented as they appear with no further syntactic context, unless deemed necessary for the proper understanding of the usage of the compound. The identified compound structures will be categorised into specific groups based on the classifications suggested in the theoretical part. The frequency of occurrence of each group of compounds will be discussed and the findings for each type of compound will be commented on separately.

The analysed selection of compound lexemes may be additionally presented in the sentences from which they have been extracted, if the sentential context is deemed influential to the analysis of the compound lexeme in any way. A singular lexeme may appear in more than one category if it shares features with multiple assorted groups of compounds. Every lexeme from the full list of findings will be labelled with a number in the range of 1–200.

It is important to mention that the graph created and used for the purposes of this thesis has been constructed solely from the lexemes contained in the corpus of this thesis. The percentages of each category of compounds may change significantly if the complete trilogy be the object of analysis, and so the overall order from the most to the least prominent could show differing results. For this reason, it is important to acknowledge the limits of the corpus and be perceptive to the fact that the results of this analysis do not apply to the whole trilogy as such.

6. **Word class**

The topic of word class will not be analysed in detail, as the aim of the thesis is to focus primarily of the compound word form of the findings. However, a graph has been created to illustrate the frequency of the word classes of compound lexemes in order to reflect the claims about word classes from the theoretical part of this paper.

[7]

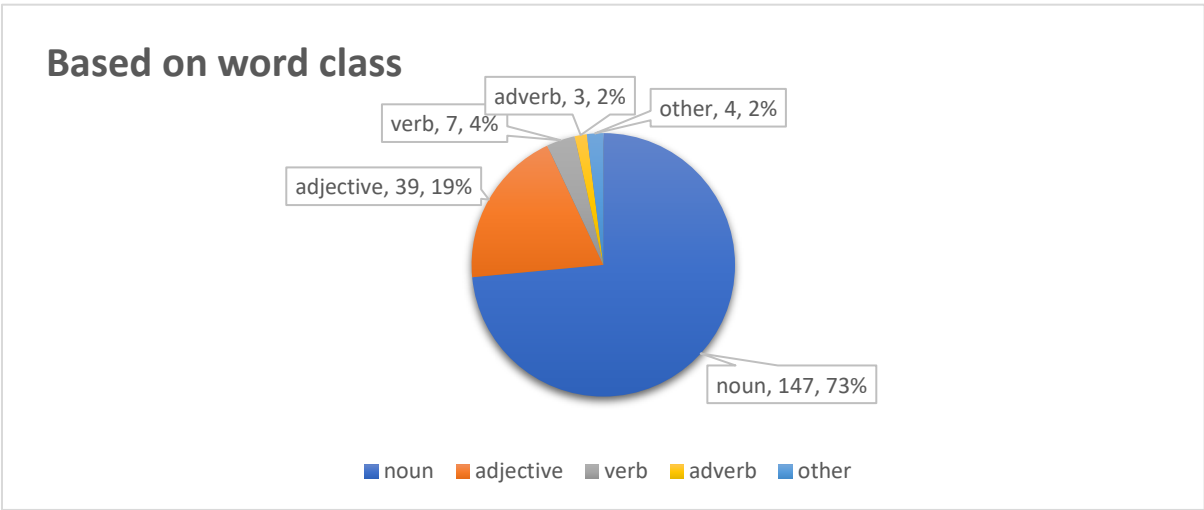


Figure 7 A graph presenting the distribution of word classes within 200 occurrences

As can be seen from the percentages presented in Figure 7, nouns represent the majority of the analysed compounded lexemes. This is greatly supported by Algeo and partially by Plag, who specifically highlights the noun-noun structure of compounded nouns as the largest group of all compound lexemes. As there have been no statements made about the frequency of other word classes, such as adjectives, verbs and adverbs, there can be no claims as to why the adjectives form the second largest group of the lexemes and why verbs and adverbs represent the minority.

The classification named “other” represents compounded lexemes, which do not fall within either of the four other categories. Although it has been said that the thesis will focus on lexical words, which only consist of nouns, adjectives, adverbs, and verbs, a small number of interjections have been deemed highly interesting from the linguistic standpoint. However, the scope of this paper does not allow for a detailed analysis of these lexemes, as that would reduce the space available for analysing other compound lexemes. Therefore, the interjections included in the corpus will not be discussed in this thesis.

7 **Word form**

Tolkien has used a plethora of compound lexemes in his works, which fortunately gave a substantive basis for the analysis. Before proceeding to the analysis of word forms of the findings, a brief summary of the topic is in order. A compound word form concerns the orthographic structure of a compound lexeme. Three distinct categories have been established in Chapter 3.1: the closed, the hyphenated, and the open form. The closed compounds contain two or more elements merged together with no other morphological changes or additional content. The hyphenated compound resembles two or more elements that are connected via a hyphen. Finally, the open compounds appear as though the elements function as two separate morphological units, as they are divided by a space, which is typical for free combinations. However, the open compounds adhere to various rules of compoundhood, which have been described and established in Chapter 3. These types of compound word forms have been identified in *The Lord of the Rings* and are the topic of the following chapters.

[8]

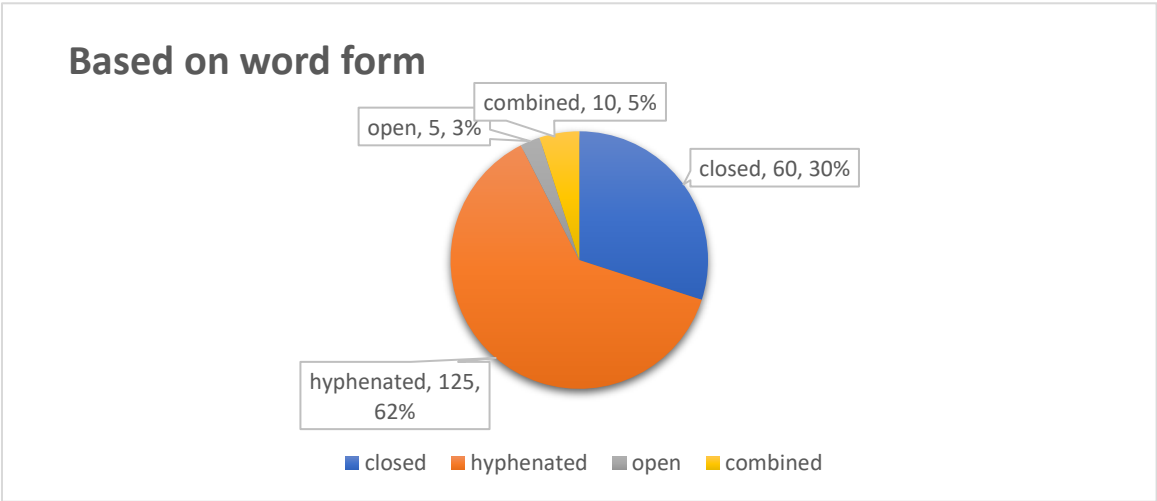


Figure 8 A graph presenting the distribution of word forms within 200 occurrences

The first graph shows the distribution of individual compound word forms, clearly showing that the most prominent group is the hyphenated form at 62% occurrence rate. This result is not surprising when Tolkien's creativity in manufacturing words is considered, because it has been already stated in Chapter 3.1.3 that the hyphenated form tends to be used for new unique compounds denoting new entities. Tolkien was a highly esteemed philologist and linguist, and so the creation of unique compound structures is not unusual in *The Lord of the Rings*. Therefore, the hyphenated form becomes a useful tool for Tolkien's writing.

The closed form being the second most prominent group with 30% occurrence rate is expected, as most well-established compound lexemes tend to appear in the closed form as a result of great semantic transparency and their frequent use.

Additionally, it is not uncommon for open compounds to assert the position of a minority, only 3% of the findings, as they are considered to be generally the least prominent group (Fernández-Domínguez 2010, 206) and, on top of that, a highly ambiguous word form of compounds (Altakhaineh 2016, 61). As a result of the open form being an ambiguous form of compounds, it is possible that other researchers and linguists might disagree on the categorization of specific lexemes that has been established for this thesis. For example, the findings directly contradict the statement of Algeo John, since according to Algeo, open compounds should be the most prominent group (Algeo 1991, 7). This goes in contrast with the 3% occurrence rate in *The Lord of the Rings*, which can be clearly observed in the presented graph in Figure 8.

Throughout the analysis of the first of the three Tolkien's works, it has been observed that another unexpected group of compounds appeared in the texts. The compound group consisting of 10% of the analysed compound lexemes, chosen to be called "combined form" for the purposes of this thesis, has appeared multiple times and has not been identified in any linguistic papers during the analysis. This group of lexemes will be individually discussed in Chapter 7.4.

7.1 Closed form

Compounds of the closed form are of the second highest occurrence rate amongst the lexemes in the corpus. Closed-form lexemes have already been established as a frequent compound form, used especially for lexemes which are well-established within the language and semantically transparent (Katamba 2005, 8). It so follows that most of the closed-form

compounds are not indeed products of Tolkien's creative writing, but rather expressions commonly used in English language, usually to be found in dictionaries for their firm establishment within the English vocabulary despite their compounded nature. The exceptions to this are compounds containing proper names of places, surnames, or names of unique fantasy races specific to the genre, for example (193) *Entwood*, or (13) *Sandyman*, functioning as a surname.

A number of lexemes classified as closed compounds are included in the dictionaries of English, such as the nouns (91) *eyelid* and (85) *firewood*, or the adverb (21) *outside*. These compound lexemes refer to entities that are a part of the established English vocabulary. Closed compounds appear in every word class in the text. It is important to remind again that this thesis is determined to analyse the lexical word classes only; that means nouns, adjectives, adverbs, and verbs. To set the rest of the prototypical examples of closed compounds, adjectives such as (173) *humpbacked* or (116) *underfed* and many other lexemes, especially of nouns, have been found, though most of the closed compound lexemes are not to be analysed further for their prototypical nature.

A unique example of a closed compound lexeme included in the corpus is (22) *handicraft*. The uniqueness stems from the elements being connected via an additional third element in the form of *i*. The definition of the compounding process has been extensively covered throughout the theoretical part of this thesis. The etymology of the lexeme *handicraft* is of crucial importance in this case, as the elements from which the lexeme originate were, according to OED, *hand* + *craft*. The presumed underlying factor for this exceptional case of compounding according to OED is its Old English origin. In the Old English language, the lexeme appeared in the form *hændecraft*, the semantic content remaining identical with the contemporary form. And although the etymological analysis of the lexeme leads no further, *handicraft* is directly linked to the lexeme *handiwork*, from which it is possible to trace the etymological meaning of the *i* in OED. The Old English form is *handgeweorc*, linking the *i* element with the *ge-* in the Old English expression. In the language's shift towards Middle English, the Old English prefix *ge-* translates as *i*. This lexeme morphologically influenced the former and this gave rise to the creation of the lexeme (22) *handicraft*. The usage of such a lexeme is coherent with Tolkien's linguistic background and expertise in Old English.

If the lexeme were to be compounded prototypically according to contemporary English, eliminating the influence of the Old English prefix, the product would be the lexeme *handcraft*.

According to OED, the lexeme *handcraft* is considered an established word in English. This fact has been further consulted with MWD and OLD as well. Further information has been collected regarding *handcraft*, which has led to the usage of the lexeme being located specifically in North America, according to OLD. Additionally, *handcraft* and *handicraft* are stated to convey the identical semantic meaning, the only difference being that (22) *handicraft* has been recorded to have a tendency of appearing in both British English and in North American English.

6.1.1 Surnames

The closed form is often used by Tolkien for the creation of personal names. First, it is important to state that the prototypical first name and surname in English does not contain a hyphen. Though there are undoubtedly cases of names in hyphenated form as well, it is not the standard. In *The Lord of the Rings*. Tolkien has placed an intense focus on compounded closed-form surnames. Most of Tolkien's characters do not possess a surname, however, the significant exception to this are Hobbits. Hobbits are the only race that are openly referred to by their names and surnames, and what is more, their surnames are in most cases results of compounding.

The possible reasoning for solely Hobbit surnames may be of a literary standpoint rather than linguistic. The reader is likely supposed to emotionally connect to the Hobbits, them being the most crucial members of the fellowship and since a significant portion of the story is focused on the fate of Frodo and Sam, the two main protagonists of Hobbit race. Therefore, the Hobbit characters were likely assigned surnames to appeal to contemporary human culture.

There have been mentions of residents of the settlement of Bree possessing surnames. The reason for the expansion of surnames to Bree may be geographical and social, as there are Hobbits, who live in Bree together with men, and it is situated in a close proximity from the Shire; the country occupied by Hobbits, which implies that the surname culture might have been simply adopted from the neighbouring community of Hobbits. The following table contains examples of compounded Hobbit surnames.

[9]

(49) Proudfoot (13) Sandyman (50) Bracegirdle (51) Brockhouse

Figure 9 Examples of Hobbit surnames that are products of compounding

It is necessary to note that the attempt at uncovering the understanding of surnames is outside the scope of this essay, the topic being subject to patronymics. However, the surnames included in Figure 9, which were created for *The Lord of the Rings*, have been deemed important to include due to their compounded nature, thus making them relevant to the topic of the thesis.

Although all the surnames presented in Figure 9 are all products of compounding, the word class of the individual elements sometimes differs. The first two surnames, lexemes (49) and (13), form a distinct group that can be defined as adjective-noun compound, which has been already discussed and defined in Chapter 3.2.1. In the compound surname (49) *Proudfoot*, the lexeme *proud* stands in the role of the modifier, whereas *foot* is the modified element. The same analysis can be applied to the lexeme (13) *Sandyman*.

The defining feature of the other identifiable group of compound surnames also concerns the structure and the word class of the compounded elements: the lexemes (50) *Bracegirdle* and (51) *Brockhouse* are structured as noun-noun compounds. OLD does not provide a definition of the lexeme *brock*, as it is not a part of the British English vocabulary but is rather an American expression. Therefore, the Merriam-Webster Dictionary, referred to as MWD, has been consulted and the meaning of *brock* is “badger”. It might be no coincidence that the lexeme *brock* is of Celtic origin, specifically from the Welsh expression *broch*, as Tolkien has indulged in studying the Welsh language, which might have been a direct motivation for this specific surname.

It can be argued that (50) *Bracegirdle* belongs in the verb-noun category, as the lexeme *brace*, according to OLD, can either be understood as a noun or a verb. In this specific case, the classification of the lexeme is ambiguous from both the morphological and semantic standpoint. According to Chapter 3.2.1, the verb-noun structure is a recognized compound form as well as the noun-noun structure, and so with the identical written form, both options are valid. Including the semantics point of view does not offer any further answers, because the meaning of the noun and the verb is almost identical. The meaning of the noun *brace* in OLD defines it as a tool for holding certain things supported in position. The verb conveys a very similar meaning, *to brace something* meaning supporting something by making it more stable and solid. Although the noun-noun structure is more prominent in English than the verb-noun structure, the intended

form cannot be claimed for certain, making this example of a compound surname in a way ambiguous.

The relationship between the compounded elements of a surname is not always completely transparent, as creativity and particularly the personal intentions and motivations of Tolkien play a crucial part in understanding any hidden meaning of the surnames. For example, the surname *Sandyman* could either indicate someone with a hair colour reminiscent of the colour of sand, or a person who works with sand. However, the motivation of the surname creation is abstract and impossible to be uncovered without a direct statement from J. R. R. Tolkien regarding this matter, that it is not likely that this essay would be able to discover the semantics and the motivation with absolute clarity.

Additionally, a problematic, and seemingly ambiguous case regarding the nature of compounded surnames have been found. One of the main characters of the fellowship, Legolas, has been referred to as Legolas Greenleaf in *The Lord of the Rings* as well as in a related work from Tolkien called *The Fall of Gondolin*. At first, it would seem that (61) *Greenleaf* resembles a surname, as it is situated in the immediate vicinity of the first name and is in form similar to the surnames of Hobbits included in Figure 9. However, (61) *Greenleaf* cannot be considered a surname, as it is not hereditary. This can be illustrated by the fact that the full name of the father of Legolas, Thranduil, does not contain (61) *Greenleaf*, therefore the name could not have been inherited.

To resolve the issue of what *Greenleaf* poses as, it is necessary to explore the meaning of the first name of this character. As has been mentioned in 1.1, Tolkien has created fictional languages on which he built his literary works. In 1.1, the existence of Sindarin language has been established, further described as a language from which Tolkien built most proper nouns for *The Lord of the Rings*. The name Legolas is of Sindarin origin and can be divided further into individual morphemes of the Sindarin language. The morpheme *laeg* represents the English lexeme *green* and is connected with the morpheme *golas*, referring to a collection of leaves or foliage. The influence of Tolkien's fictional language interferes with personal names on the morphological level in this manner and under this logic, *Greenleaf* poses as an epithet⁶. An epithet can be defined as a word or a phrase that occurs in the role of a name, according to

⁶ The decision for identifying (61) *Greenleaf* as an epithet is additionally supported by an article on the website Tolkien Gateway. It is important to note that the reason for citing the Tolkien Gateway is because it is the only available resource discussing the topic of patronymics specifically in the case of the given lexeme.

MWD. Additionally, the epithet is supposed to characterize the person or thing, to which it is connected. Therefore, it can be argued that the translation of Legolas' Sindarin name to English serves a descriptive and characterizing function, achieving the properties of an epithet together with the placement of following the first name.

The nature of surnames is not widely explored in the context of Tolkien's fantasy; therefore, it can only be assumed, with the aid of reliable linguistic and patronymic resources, as to what the role of patronymics is in *The Lord of the Rings*.

7.2 Open form

The open-form compound lexemes constitute a minority of 3% from the findings listed in the corpus of this essay. The resulting number of open forms is coherent with the claims made about its frequency discussed in Chapter 3.1.2, except for the approach of Algeo, to which the findings are contradictory. The usage of open compounds by Tolkien, although amounting only to a few example lexemes, has shown certain tendencies, especially concerning the "combined form" phenomenon, which is going to be analysed in a separate Chapter 7.4. The following table contains examples of the open compounds taken from the corpus.

[10]

(113) <i>inside shutters</i>	(114) <i>cold hour</i>
(115) <i>death's door</i>	(182) <i>Common Tongue</i>

Figure 10 Examples of open compounds found in *The Lord of the Rings*

The open compounds are the main compound form for which establishing the compoundhood according to the rules devised by Altkhaineh is necessary, as there is great ambiguity in distinguishing between open compounds and free combinations. The ambiguity in identifying a lexeme as an open compound structure is the highest of all categories.

The lexeme (113) *inside shutters* can be proved to be an open compound by applying the rule of insertion by Altkhaineh (2016, 68) or by following the Lexical Integrity Principle (Lieber, Štekauer 2011, 292). The rule of insertion is a claim by which it is impossible to insert a word between the individual elements of a compound lexeme. If the lexeme was potentially modified by an adjective, because of its compounded structure the adjective would be situated in the initial position. For example, if the inside shutters were green, the resulting form of the whole

noun phrase would be *green inside shutters*. This shows the coherency with the rule of insertion, as it would not be grammatically correct to insert the adjective *green* between the elements with the result being the word *inside green shutters*.

Regarding the possibility of the lexeme being simply a free combination or collocation, a counterargument can be applied to this specific case, considering the order of adjectives. The lexeme *inside* can be categorized as an adjective, which allows for the theory of a restrictive sequence of adjectives to be applied (Sopher 1962, 192). If *inside* is categorized as an adjective denoting the purpose and quality of the shutters, the potential further modification cannot interject between *inside* and *shutters*. Quality and purpose denoting adjectives are the category of adjectives that appear the closest to the modified noun, following the grammatical rules for the order of adjectives from the Cambridge Dictionary. This means that even if *inside shutters* was a free combination and not an open compound, it would still be impossible to insert any further adjectives, as every other type of adjective would be positioned before the adjective *inside* regardless. This concludes the confirmation of compoundhood of the lexeme (113) *inside shutters*.

The lexeme (114) *cold hour* refers to a concept of temperature drop, specifically the low temperatures before dawn based on the sentential context presented in Figure 11.

[11]

“The **cold hour** before dawn was passing.” (Tolkien in “The Rings Sets Out” 2001, 232)

Figure 11 Sentential context for the lexeme (114) *cold hour*

The compoundhood of the lexeme can be verified by the rule of modification. Altakhaineh claims that unlike in phrases, the first element of a compound lexeme cannot be modified. In this case, it could be argued that it is possible to modify the lexeme using for instance the amplifier “very”, resulting in a construction *very cold hour*. However, the addition of the lexeme *very* breaks the compoundhood of the structure that it becomes a free combination referring to a different concept than the original compounded lexeme. Therefore, (113) *cold hour* has been classified as an open compound.

The lexeme (115) *death’s door* has been labelled an open compound despite the seemingly opposing rule of inflection devised by Altakhaineh (2016, 73). According to Altakhaineh, the elements of a compound word cannot take inflection. This would normally disqualify (115)

death's door from being recognized as a compound lexeme for the possessive genitive case of the first element. However, as Altakhaineh subsequently admits with reference to Bauer et. al., there are instances in English, where compound lexemes appear in the form of descriptive genitives (2013; quoted in Altakhaineh 2016).

[12]

children's hour girl's club no-man's-land child's play

Figure 12 Examples of descriptive genitives as presented in Altakhaineh (2016, 74)

The example compounds listed reflect the form of (115) *death's door*. Therefore, in the case of this particular lexeme, the compoundhood in the form of a descriptive genitive has been confirmed. Descriptive genitives, according to Rosenbach (2006, 78), can be divided into two types of s-genitives: a determiner genitive and a descriptive genitive. Rosenbach quotes Quirk et al. in the claim that descriptive genitives usually possess an idiomatic relationship with the head element, which can be observed in the case of the lexeme (115) *death's door*. To illustrate this theory, the sentential context is provided in Figure 13.

[13]

“The chief danger is that the poor beast is probably at **death's door**.” (Tolkien in “The Rings Sets Out” 2001, 236)

Figure 13 Sentential context for the lexeme (115) *death's door*

The meaning of the lexeme is not understood literally, as a real existing door belonging to the entity of death, but rather as an idiomatic expression of someone being on the verge of dying, according to MWD. This feature of idiomaticity is coherent with the classification of the lexeme as a descriptive genitive.

Finally, the lexeme (182) *Common Tongue* has been chosen to represent the class of open compounds. The compoundhood of the lexeme can be likewise confirmed by the principle of modification by Altakhaineh, which claims that it is impossible to modify the first element of a compound lexeme (2016, 62). Altakhaineh additionally discusses the rule of referentiality, based on the tendency of compound lexemes being non-referential in nature (2016, 69). However, Altakhaineh subsequently adopts the perspective of Bauer et al. and acknowledges the existence of compounds, where either the first element or the whole compound contains a

proper noun (Bauer et al. 2013: quoted in Altakhaineh 2016). Proper nouns are inherently referential, and so even though (182) *Common Tongue* is objectively a proper noun, it does not necessarily follow that it is impossible to classify the lexeme as a compound based off its referentiality. For this analysis, the lexeme is considered to be an open-form proper noun constructed by the process of compounding.

In the chapter devoted to discussion about open compounds, it has been concluded that the findings are coherent with the approaches discussed in the theoretical part regarding the high ambiguity of open-form compound lexemes.

7.3 Hyphenated form

The hyphenated compounds represent the largest category among the findings, the whole corpus consisting of 62% of hyphenated form. The fact that they are the most prominent group of all compounds considered in the research for this paper is supported by the approach of Katamba, as he states that the hyphenated form is often used specifically for fairly new expressions. Tolkien was engaged in word and language creation from an early age (Robinson 2013, 65) and that is reflected in the style of *The Lord of the Rings* in the way of hyphenated compounds, which can be manufactured for one or more instances and function more as disposable items rather than expressions with the potential to be used repeatedly (Katamba 2005, 150).

The usage of hyphenated compounds covered every word class, though some classes appeared more frequently than others. For example, the most prominent word class of the hyphenated compounds were nouns. This is consistent with the statement of Algeo where he claims that around 90% of compound lexemes belong in the category of nouns, the adjectives, adverbs, and verbs being only secondary in frequency.

7.3.1 Neologisms in compounding

Although compound lexemes are said to usually be of non-referential nature, it has been established in the theoretical part of this thesis that referential compounds exist. The majority of proper nouns in *The Lord of the Rings* concerns names of fictional races, such as Hobbits and Ents. Hobbits are a race of short human-looking creatures with large feet who live in underground dwellings, according to Merriam-Webster, and Ents represent a race of living humanoid tree-like creatures burdened with protecting the Fangorn Forest of Middle Earth, as

defined on Tolkien Gateway⁷. Tolkien has used mainly hyphenation and closed forms in creating new compounds containing the names of the fictional races, as can be seen in examples such as (105) *hobbit-talk* and (180) *Entmaiden*. The differences in capitalization and in compound word form will be discussed in the following chapter.

The neologism *mathom* created by Tolkien solely for *The Lord of the Rings* works has already been mentioned in Chapter 1.1 discussing Tolkien's style and his ambitions for creating fictional languages. Tolkien has additionally used the word *mathom* to create a hyphenated compound lexeme (23) *Mathom-house*. It is worth noting that Tolkien decided to capitalize the lexeme, which indicates that it is considered a proper noun. This could be due to the fact that *Mathom-house* was described as an establishment exclusively of the Hobbits, or possibly only in the settlement of Hobbiton. If it assumed that it is a Hobbiton-exclusive building, (23) *Mathom-house* is most likely capitalized because it is a unique entity and there is only one *Mathom-house*.

7.3.2 Proper nouns

Amongst all names for places introduced in *The Lord of the Rings*, some places carry names that are results of compounding. With regards to the hyphenated form, (1) *Barrow-downs* can be taken as an example. The lexeme has shown a level of productivity, since from it a new lexeme (158) *Barrow-blade* has been derived. The compounding process is not entirely prototypical, as the first element of the resulting compound does not contain the full original lexeme but merely its part in the form of the first element. Since in *The Lord of the Rings*, no entity that can be referred to as "Barrow" exists, it is implied that the lexeme (158) *Barrow-blade* must be derived specifically from the lexeme (1) *Barrow-downs* and could be further defined as "a blade from Barrow-downs". If the compounding process was prototypical, the compound would result in being constructed from three individual elements. That is most likely the reason for the elimination of the *-downs* element in (158), as it has been claimed by Plag in Chapter 3.2.1, that it is discouraged to use long constructions to avoid unintelligibility (2002. 172).

The lexeme (158) *Barrow-blade* could be a case of a combination of word-formation processes, combining compounding and clipping. Clipping is a process in which a certain part of a longer word is removed while keeping the same semantic content (Bauer 2004, 30). As an example,

⁷ Tolkien Gateway was chosen as a source as a result of the absence of a definition of the term "Ent" in any major dictionaries.

Bauer illustrates clipping with the lexeme *telephone*, which through the process of clipping results in a lexeme *phone*, by removing the *tele-* part from the beginning of the word. The explanation of clipping would account for the shortening of the element (1) *Barrow-downs*. The lexeme (158) *Barrow-blade* cannot be associated with any one word-formation process, and so it follows that the lexeme is a result of a combination of word-formation processes, which have been identified to be compounding and clipping.

Additionally, Tolkien has uniquely created a compound lexeme using a first name of a character in (62) *Gollum-creature*. No other compound resembling this structure has been found. It is not clear why Tolkien decided to construct the compound, since referring to the creature by its first name *Gollum* would suffice. In *The Lord of the Rings*, it has been made clear that “Gollum” is a first name that Sméagol adopted, and so the structure cannot be attributed to being a reference to a type of a creature or a race, making the lexeme a unique occurrence in this analysis.

7.3.3 Quotational compounds

Quotational compounds have been briefly described in Chapter 3.1.3. A few examples of these compound lexemes have been identified, including both well-established compound lexemes in the English language and compounds which convey specific situational information. Three examples of quotational compounds in question are listed in the following figure, each illustrating varying degrees of lexicalization of a compound.

[14]

(88) *forget-me-not* (135) *Sword-that-was-Broken* (196) *you-can't-have-it-so-go-to-sleep*

Figure 14 Examples of quotational compounds found in *The Lord of the Rings*

The quotational compound (88) *forget-me-not* is a well-established compound referring to a type of flower. This lexeme is extraordinarily lexicalized, which is not typical for quotational compounds, as they are mostly formed as nonce formations (Ménová 2012, 18–19) quoted in Rybínová, 2021), the usage and meaning of them being aimed at a specific situational context.

The lexeme (135) *Sword-that-was-Broken* can be assumed to have been created exclusively by Tolkien, as the compound is a proper noun, on account of the capitalization and its referential nature to a unique entity in the story of *The Lord of the Rings*. The capitalization of the lexeme appears in two elements, both of which are lexical words; the grammatical words contained within the compound lexeme are not capitalized. This quotational compound has been used

throughout the analysed texts multiple times consistently. Therefore, it can be claimed that (135) *Sword-that-was-Broken* has become to some degree lexicalized, though only within the context of *The Lord of the Rings*.

Finally, the lexeme (196) *you-can't-have-it-so-go-to-sleep* has been identified as a lexeme containing an uninterrupted sequence of finite clauses transformed into a quotational compound. The compound lexeme consists of nine individual elements, which makes this lexeme a suitable example showcasing the freedom of recursivity in hyphenated compounds. The compound was not used by Tolkien in any more instances. For this reason, the quotational compound (196) has been deemed not lexicalized but rather an expression used to convey a particular message for one specific situation.

7.3.4 Influence of fictional languages

Tolkien's fictional languages have already been established to bear a great influence on the creation of names of places and first names of characters. The Old Entish language has not been thoroughly explored, as Tolkien focused on Quenya and Sindarin to a higher degree. Nevertheless, Tolkien has offered two direct translations of Entish words in *The Lord of the Rings*, which have led to intriguing and extraordinary compound lexemes to materialize.

It is important to note that these lexemes have been taken from books outside the scope of this essay. This thesis was initially meant to cover the whole trilogy, however, throughout the process of the analysis, the number of compounds has far exceeded the expectations as well as the capacity of the corpus. Despite that, the two lexemes have been chosen to be included in this paper for their extraordinary compounded nature as well as to pay respect to Tolkien's linguistic inventions.

First, it is important to specify a crucial detail considering the grammatical rules of Old Entish. Individual words in Old Entish take form of long strings of expressions that tell the story of the given entity that is being referred to (Tolkien in "Appendices" 2001, 137). Therefore, words translated from Old Entish to English appear in sentences or longer paragraphs of text.

The first translation of an Old Entish word appears in the sixth book *The End of the Third Age* and the lexeme is supposed to denote the word "orc". The main character of the race of Ents, Treebeard, talks about orcs as (6) *evileyed-blackhanded-bowlegged-flinthearted-clawfingered-foulbellied-bloodthirsty*. The extraordinarily long and complex translated lexeme is an adjective. However, its proper English equivalent is "orc", which is a noun. The structure of the

Old Entish word seems to be strikingly different from the English language regarding the word class incoherence.

However, from the sentential context it is implied that the translation of the Old Entish word is not complete, as the character Treebeard suffices to referring to the creatures as “those vermin of orcs” in the midst of the direct translation. The sentential context is provided in the following figure.

[15]

“For there was a great inrush of those, *burárum*, those **evileyed-blackhanded-bowlegged-flinthearted-clawfingered-foulbellied-bloodthirsty**, *morimaite-sincahonda*, *hoom*, well, since you are hasty folk and their full name is as long as years of torment, those vermin of orcs...” (Tolkien in “The End of the Third Age” 2001, 114)

Figure 15 Sentential context for the lexeme (6)

Therefore, as Tolkien did not offer the complete translated expression, there is a possibility that if the lexeme was in its full form, the syntactic head could pose as a noun and the word classes would be cross-linguistically coherent through translation.

The second lexeme (7) *Forestmanyshadowed-deepvalleyblack Deepvalleyforested Gloomyland* differs in that the direct translation poses as a compound, however, the semantic content is understood as a sentence, which is illustrated in the sentential context in the following figure.

[16]

“It may be rendered ‘**Forestmanyshadowed-deepvalleyblack Deepvalleyforested Gloomyland**’, and by which Treebeard meant, more or less: ‘there is a black shadow in the deep dales of the forest’.” (Tolkien in “Appendices” 2001, 138)

Figure 16 Sentential context for the lexeme (7)

It cannot be identified with absolute certainty, whether the lexeme is truly a compound or whether this entry in the corpus consists rather of three individual compound lexemes, as the rules for compoundhood are not applicable in this case. The compound itself is highly influenced by the fictional language and functions purely as a direct translation of the Old Entish utterance not meant to be used in this form, but rather in the form of the following English sentence denoting the meaning from Figure 16.

7.4 Combined form

During the analysis of compound word forms, an issue in classification has been encountered in several cases. The lexemes such as (169) *limestone-boulder*, (96) *blackberry tart* or (58) *fireside-story* are results of multiple compounding processes, thus forming a more complex word form. In this thesis, the form has been given the name “combined” for the nature of two different compound word forms present. There is no mention of such compound word form existing in any reliable linguistic sources. This implies that these lexemes are not examples of a type of a fourth word form, but rather a compound lexeme constructed from two elements, the first of which is in itself a compound. For example, the lexeme (58) *fireside-story* was constructed from the words *fireside* and *story* in the hyphenated compound word form. The element *fireside* was previously constructed from the elements *fire* and *side* and are of the closed form. This concludes that no fourth compound word form exists.

7.5 Inconsistencies in form

It is necessary to mention that Tolkien was not always consistent with the way in which he created compound lexemes. There have been numerous instances, where a lexeme appeared in multiple word forms throughout the books. The lexemes being referred to regarding this issue are listed in the following figure.

[17]

(155) <i>door-way</i>	(156) <i>doorway</i>	(183) <i>watchman</i>	(184) <i>watch-man</i>
(185) <i>foot-hill</i>	(186) <i>foothill</i>	(187) <i>herd-folk</i>	(188) <i>herdfolk</i>
(145) <i>night-fall</i>	(146) <i>nightfall</i>	(147) <i>highroad</i>	(148) <i>high-road</i>

Figure 17 Examples of inconsistencies in word form

The reason for this incongruence of compound forms is not transparent, as no link between the word-form used and its semantic content has been found. The hyphenated modifier rule mentioned by Magat in the Chapter 3.1.3 cannot be applied in this case, as the lexemes included in Figure 17 do not function as modifiers in their separate sentential contexts. The perspective of shifting from hyphenated towards closed form in terms of development of language throughout time also cannot be considered a factor in this case, the reason being that the compound lexemes analysed appear within one to two books, which were written by Tolkien in close succession, all being published within 2 years (Carpenter 2000). Therefore, it is unlikely

that the development of language and the level of establishment of a compound word could change during the time that Tolkien was writing. Other editions of *The Lord of the Rings* have not been taken into account, and so it is possible that the compound word form at least partially depends on the editor of a particular edition. This has not been neither confirmed nor denied.

Additionally, irregularities in Tolkien's works are present even regarding the region of Tolkien's choice of words. The lexemes (72) *wooly-footed* and (111) *woolly-footed* have both been found in Tolkien's works. According to the OLD, the lexeme (111) *woolly-footed* is an expression commonly found in the British English, whereas (72) *wooly-footed* is associated with American English. The reasons for this incongruence can be various. The lexeme (72) appears in the first book *The Rings Sets Out*, whereas the lexeme (111) is recorded to appear in the third book *The Treason of Isengard*, meaning that the lexemes were not used in the context of one book. Therefore, Tolkien may have come into contact with American English in the midst of completing the third book, which may have influenced the choice of preferred expressions. However, as mentioned in the previous paragraph, the books have been written in close succession, and therefore the influence of American English, if there was any, would not be likely to significantly influence Tolkien's active vocabulary. It is also possible that the inconsistency was caused by an error of either Tolkien or the editors of Tolkien's works. This could be further explored if other editions of *The Lord of the Rings* were part of the analysis.

Finally, a brief commentary on the already discussed lexemes from Chapter 7.3.1 is in order, as the lexemes containing the word "Hobbit" share inconsistent capitalization when compounded. The lexemes such as (26) *hobbit-architecture* or (25) *hobbit-building* come in stark contrast to lexemes (30) *Hobbitry-in-arms* and (14) *Hobbit-lore* in terms of the capital first letter. The same applies to lexemes (177) *ent-stride*, (178) *ent-house*, (179) *Entwife*, and (180) *Entmaiden*, which among other differences contain capitalization inconsistencies as well. As Tally confirms (2019, 2), Tolkien's choice of capitalization of all races besides the mentioned is inconsistent and no link between the semantic content and the capitalization was found. However, Tolkien was a highly esteemed linguist, and so it can be assumed that Tolkien had reasons underlying the choice in capitalization, even though Tolkien's motivation and basis for this was not uncovered in this paper.

CONCLUSION

To summarize, this bachelor thesis aimed to analyse the distribution of different word forms of compound lexemes in the first three books of *The Lord of the Rings*. 200 findings have been collected during the research out of which a corpus was created, which served as a basis to the analytical part of this paper. Subsequently, examples from the corpus were interpreted accordingly to approaches defined in the theoretical part.

As a result of the thorough analysis, the scope of the essay did not allow for all significant phenomena to be mentioned and to be given its respective chapters. The intention was to include a chapter discussing ablaut motivated compounds, for which several examples have been identified in the texts, which is a unique group of compound lexemes created with regards to phonetic motivation. Moreover, recurring productive affix-like elements, such as *-wood*, *-land*, or *-forest* were intended to be commented on in the analysis as well. Additionally, all lexemes were intended to be analysed from the semantic standpoint, categorized as either endocentric or endocentric. Restrictions on the scope of the paper did not allow for a detailed analysis of these morphologically significant lexical items, however, it may present as a topic for future research into the matter of compounding.

It is shown that Tolkien relied on compounding in *The Lord of the Rings* by the number of findings in merely the first three books exceeding the expectations and the scope of the essay. To conclude the specific results of the analysis regarding the form, the hyphenated form of compound lexemes is the most prominent that Tolkien preferred to use. The hyphenated form dominated the other forms with 62% occurrence as opposed to the 30% of closed-form compounds and the remaining 3% of open and 10% of “combined” form being applied by Tolkien. As has been stated in Chapter 3.1.3, this is connected to the tendency of the hyphenated form being used primarily in new compound structures and subsequently so to Tolkien’s motivation for creating new languages and words discussed in Chapter 1.1 concerning Tolkien’s life and style.

The rates of occurrences of word classes are coherent with the statement of Algeo that nouns pose as the majority of all compound lexemes, as the findings revealed a 73% occurrence of compounded nouns, as well as the approach towards adjectives being the second most prominent group of compound word classes with 19% occurrence within the scope of the corpus.

The semantic properties of compound lexemes were analysed and illustrated on a number of findings, although the scope of this essay could not include an equal amount of attention to each corpus entry, focusing on lexemes that required a detailed commentary and research rather than prototypical compounds.

The conclusion is that Tolkien has in the majority of cases used the hyphenated compound word form and that the majority of compounds used by Tolkien were identified as nouns.

RESUMÉ

Tato bakalářská práce se zaměřuje na výzkum slovních složenin v prvních třech knihách sedmidílného vydání knižní série *The Lord of the Rings* od autora J. R. R. Tolkiena. Práce je rozčleněna do dvou hlavních částí. Teoretická část je pokrytá v práci od Kapitoly 1 do Kapitoly 3.2.3 a zabývá se teoretickými přístupy k tématu slovního skládání a kontrastování daných přístupů různých lingvistů. Obsahem prvních dvou kapitol je stručné shrnutí stylu beletrie a následně život a styl autora J. R. R. Tolkiena. Dílčím cílem teoretické části je vytvořit adekvátní podklad složený z vhodných morfologických přístupů, který bude sloužit jako základ pro část analytickou. Následně analytická část se věnuje jednotlivým poznatkům a nálezům z knih *The Lord of the Rings*, jejichž analýza se opírá o přístupy definované v části teoretické. V praktické části se výzkumník pokouší o odhalení případných tendencí v produktech slovního skládání a o jejich morfologicky výstižnou klasifikaci. V určitých případech se výzkumník také snaží na základě teoretických znalostí tématu slovních složenin a oblasti morfologie a sémantiky anglického jazyka určit potenciální motivaci J. R. R. Tolkiena za vytvářením specifických složených struktur. Cílem práce je prostřednictvím kvantitativní a kvalitativní analýzy identifikovat Tolkienem nejvíce používaný slovní druh a slovní formu složenin a detailně analyzovat nálezy dle adekvátně zvolených postupů a postojů definovaných v teoretické části této bakalářské práce.

Teoretická část této práce se zprvu zabývá definicí beletrií, přičemž na základě odborných zdrojů je lingvisticky definovat žánr beletrie téměř nemožné pro rozsáhlost škály možných stylů psaní. Proto se v Kapitole 1.1 práce zaměřuje na život J. R. R. Tolkiena a jeho autorský styl. Velký důraz je kladen Tolkienově motivaci pro tvoření fiktivních jazyků a jeho odborné lingvistické expertíze v oblasti především Staroangličtiny, Velštiny, Latiny, Severštiny, Řečtiny a Finštiny, která výrazně inspirovala a ovlivnila tvorbu Tolkienem vytvořených jazyků.

Dále se teoretická práce Kapitolou 2 začíná zabývat morfologickými slovotvornými procesy, včetně derivace, se speciálním zaměřením na proces skládání. Výzkumník porovnává odlišné přístupy ke klasifikaci i pojmenování činitelů slovních složenin a procesu skládání a následně definuje přístup zvolený pro tuto bakalářskou práci. Následuje Kapitola 3, jenž se zaměřuje především na úskalí a problematiku definování slovní složeniny, jako například dvojznačnosti open složenin a frází, či rozdílům v klasifikaci složenin v rámci slovotvorných procesů. Důraz je kladen na pravidla pro identifikaci tzv. “compoundhood”, což se dá v češtině definovat jako “stav bytí složeninou”. Tato pravidla jsou následně hojně používána v praktické části k analýze

open složenin. Tato kapitola se následně rozděluje na tři jednotlivé podkapitoly, každá zabývající se jinou slovní formou složenin. Kapitola 3.1.1 se zaměřuje na closed formu slovních složenin, která se vyznačuje bezprostředním spojením jednotlivých skladebných prvků, a na vlastnosti a tendence spojené s touto formou složenin. Stejným způsobem je následně zpracována Kapitola 3.1.2 o open formě a Kapitola 3.1.3 o hyphenated formě složenin. V Kapitole 3.1.2 se taktéž ustanovuje schopnost slovních složenin se objevovat v jazyce v několika různých formách.

Kapitola 3.2 krátce definuje termín “syntaktická hlava”, která je nepostradatelnou a nezbytnou součástí prototypické slovní složeniny, a pojednává o tendencích a frekvencích složenin o různých slovních druzích. Na závěr teoretické části této bakalářské práce kapitoly 3.2.1 až 3.2.3 referují o jednotlivých slovních druzích včetně podkategorií slovních druhů jednotlivých skladebných prvků.

Analytická část je zahájena stanovením rozsahu analýzy *The Lord of the Rings* a její metodologie. Kvalitativní metody zastupují názorné grafy čerpající z vytvořeného korpusu o 200 nálezech, znázorňující poměry slovních forem a druhů složenin. Kvantitativní metoda je prováděna prostřednictvím detailní analýzy jednotlivých nálezů, diskusí o jejich případných dvojznačnostech a úskalích, identifikací compoundhood dle daných pravidel, a jejich následnou teorií oprávněnou kategorizací.

V kapitole zaměřující se na closed složeniny se výzkumník pokouší detailně analyzovat jednotlivé nálezy z hlediska etymologického, cizojazyčného a patronymického. Unikátní výraz *handicraft* byl analyzován z hlediska cizojazyčného s ohledem na Tolkienovu zkušenost se Staroangličtinou, která byla určena motivací pro použití daného lexému. Předmětem výzkumu se stala také ojedinělá příjmení Hobitů, jenž jsou produktem procesu skládání. Diskutována byla sémantická motivace za tvorbou daných příjmení a sémantický vztah mezi jednotlivými skladebnými prvky. Následně byl analyzován lexém *Greenleaf* pro jeho unikátní roli epitetu v kombinaci s procesem skládání a vlivem fiktivního jazyka na strukturu lexému a jeho vznik.

Obsah kapitoly zabývající se open formou složenin se skládá primárně z aplikování pravidel pro zjišťování compoundhood u konkrétních nálezů z důvodu vysoké úrovně dvojznačnosti mezi složeninami a frázemi. Součástí této kapitoly je i definice termínu “deskriptivní genitiv” či “popisný genitiv”, který byl ojediněle identifikován v *The Lord of the Rings*. U každého ze čtyř prezentovaných nálezů se compoundhood potvrzuje jiným z pravidel pro compoundhood

pro rozmanitost užití a pro individuální přístup adekvátní pro konkrétní nálezy. Větný kontext je uveden u složenin, u kterých má vliv na správné porozumění lexému.

Nejkomplexnější formou složenin, hyphenated formou, se zabývá Kapitola 7.3, která je dále rozdělena na čtyři tematické podkapitoly. Na začátku kapitoly je uveden výsledek kvantitativního výzkumu výsledků průzkumu *The Lord of the Rings*, jehož výsledky jasně určují, že hyphenated forma je nejčastější používanou formou složenin v rámci korpusu. V Kapitole 7.3.1 se výzkumník zaměřuje na užití Tolkienových specifických neologismů, například slova *mathom*, s kterým se pojí etymologický průzkum tohoto lexému, a dvou fiktivních humanoidních ras *Hobbit* a *Ent* ve formě hyphenated složenin. Následuje kapitola 7.3.2 diskutující vlastní jména, která jsou produkty procesu skládání. V rámci tohoto tématu se výzkumník snaží identifikovat slovtvorný proces použitý při tvorbě lexému *Barrow-blade*, výsledkem čehož je diskuse o kombinaci procesu skládání a zkracování. Následující Kapitola 7.3.3 se věnuje takzvaným “quotational” složeninám, jejichž forma se podobá větám více než slovům. Tři příklady tohoto typu složenin byly prezentovány, analyzovány a následně seřazeny podle stupně lexikalizace. Poslední podkapitola hyphenated složenin je předmětem průzkumu vlivu Tolkienových fiktivních jazyků na proces skládání v jazyce anglickém. Výsledkem bylo potvrzení značeného působení fiktivních jazyků na proces skládání s ohledem na morfologickou strukturu složenin.

Kapitola 7.4 stručně pojednává o neočekávané a neuznávané skupině “kombinovaných” složenin, které byly několikrát nalezeny v textu. Jedná se o složeniny, které jsou produktem několika stupňů procesu skládání, a to i různých forem. Žádný odborný zdroj tuto formu neuznává, a proto nebyla považována za stěžejní bod analýzy.

V poslední kapitole analytické části se výzkumník snaží dostatečně okomentovat veškeré nekonzistence slovních složenin, zahrnující lexémy objevující se ve dvou různých formách v rámci jedné až tří knih, lexémy z různých variant anglického jazyka, nejednotné používání jedné formy v souvislosti se skladebnými prvky a kapitalizaci lexémů.

Výsledkem analytické části jsou Tolkienem objektivně nejvíce preferované formy a druhy slovních složenin, které jsou založené na sběru složenin do korpusu bakalářské práce a jejich následném vyhodnocení prostřednictvím analýzy. Nejčastěji užívaným slovním druhem složenin byla identifikována podstatná jména, což souhlasí a také navazuje na předpoklady uvedenými v teoretické části. Co se týče slovní formy složenin, nejčastější formou byly určeny hyphenated složeniny. Výsledky jsou koherentní s předpoklady spojené s Tolkienovým

lingvistickým pozadím a s možností kreativního tvoření neomezeně komplexních slovních složenin.

BIBLIOGRAPHY

Primary sources

Tolkien, John R. R. 2001. *The Lord of the Rings*. Seven-volume ed. London: HarperCollins Publishers.

Secondary sources

Abberley, Will. 2015. *English Fiction and the Evolution of Language, 1850–1914*. Cambridge: Cambridge University Press.

Aditiawarman, Mac. Diana Kartika, and Wahyudi Rahmat. 2022. *The Use Of Digital Newspapers As A Learning Media Close Compounds Words (Words Forms) English Sentences In The Daily Use Of Students in West Sumatra*. Vol. 19.

Algeo, John. ed. 1991. *Fifty Years Among the New Words: A Dictionary of Neologisms, 1941–1991*. Cambridge: Cambridge University Press.

Algeo, John. 1985. “The Toponymy of Middle-Earth.” *A Journal of Onomastics*: 80–95.

Altakhaineh, Abdel R. M. 2016. *What is a compound? The main criteria for compoundhood*. Al Ain: Al Ain University of Science and Technology.

Bauer, Laurie. 2004. *A Glossary of Morphology*. Washington, D.C.: Georgetown University Press.

Bauer, Laurie, Rochelle Lieber, and Ingo Plag. 2013. “The Oxford Reference Guide to English Morphology.” in *What is a compound? The main criteria for compoundhood*. Al Ain: Al Ain University of Science and Technology.

Booij, Geert E. 2009. “Compounding and Construction Morphology.” in *The Oxford Handbook of Compounding*, edited by Rochelle Lieber and Pavol Štekauer, 322–347. Oxford: Oxford University Press

Carpenter, Humphrey. 2000. *J. R. R. Tolkien: A Biography*. Boston: Houghton Mifflin.

Clendenning, George. 2003. *Fantasy: Its Definitions, Its Critics, and J.R.R. Tolkien*. Youngstown: Youngstown State University.

- Constant, Mathieu, Gülşen Eryiğit, Johanna Monti, Lonneke van der Plas, Carlos Ramisch, Michael Rosner, and Amalia Todirascu. 2017. *Multiword Expression Processing: A Survey*. Association for Computational Linguistics.
- Davis, Norman. 1967. *Sir Gawain and the Green Knight*. Oxford: Oxford University Press.
- De Belder, Marijke. 2011. *Roots and affixes: Eliminating Lexical Categories from Syntax*. Netherlands: LOT.
- De Melo, Gerard. 2014. *Etymological Wordnet: Tracing The History of Words*. Reykjavik: European Language Resources Association (ELRA).
- Delahunty, Gerald P. and James J. Garvey. 2010. *The English Language: From Sound to Sense*. West Lafayette: Parlor Press.
- Drout, Michael D. C. 2004. "Tolkien's Prose Style and its Literary and Rhetorical Effects." *Tolkien Studies* 1: 137–163
- Fernández-Domínguez, Jesús. 2010. *Productivity vs. Lexicalization: Frequency-Based Hypotheses on Word-Formation*. Poznań: School of English, Adam Mickiewicz University.
- Frey, Albert R. 1933. *Sobriquets and Nicknames*. Cambridge: The Riverside Press.
- Greene, Dierdre. 1996. "Tolkien's Dictionary Poetics: The Influence of the *OED*'s Defining Style on Tolkien's Fiction." *Mythlore: A Journal of J.R.R. Tolkien, C.S. Lewis, Charles Williams, and Mythopoeic Literature* 21 (2): 195–199.
- Hammond, Wayne G. 1993. *J. R. R. Tolkien: A Descriptive Bibliography*. New Castle, Delaware: Oak Knoll Press.
- Katamba, Francis. 1994. *English Words*. London: Routledge.
- Langacker, Ronald. 1987. *Foundations of Cognitive Grammar*, vol. 1. Stanford: Stanford University Press.
- Leech, Geoffrey N. and Michael H. Short. 1981. *Style in Fiction: A Linguistic Introduction to English Fictional Prose*. New York: Longman Inc.
- Lieber, Rochelle and Pavol Štekauer. 2011. *The Oxford Handbook of Compounding*. Oxford: Oxford University Press.
- Magat, Joan A. 2014. "Hawking Hyphens in Compound Modifiers." *Legal Communication & Rhetoric: JALWD* 11: 153–175.

- Marchand, Hans. 1967. "Expansion, Transposition, and Derivation." *La linguistique* 3. Fasc. 1. 13–26
- Ménová, Martina. 2012. *Phrasal compounds in contemporary British newspapers*. Praha: Univerzita Karlova
- Olsen, Corey. 2008. *The Myth of the Ent and the Entwife*. Tolkien Studies. West Virginia University Press.
- Plag, Ingo. 2002. *English Word-formation*. Cambridge: Cambridge University Press.
- Robins, Robert H. and David Crystal. "language." *Britannica*, March 12, 2024. <https://www.britannica.com/topic/language>
- Robinson, Christopher L. 2013. *What Makes the Names of Middle-earth So Fitting? Elements of Style in the Namecraft of J. R. R. Tolkien*. Vol. 61 (2) Paris: HEC-Paris.
- Rosenbach, Anette. 2006. "Descriptive Genitives in English: A Case Study on Constructional Gradience." *English Language and Linguistics* 10. Cambridge: Cambridge University Press.
- Rybínová, Veronika. 2021. *Phrasal Compounds in English Fiction*. Praha: Univerzita Karlova.
- Ryšavá, Andrea. 2012. *Noun Phrase in English: Its Form, Function and Distribution in Text*. Olomouc: Filozofická fakulta Univerzita Palackého v Olomouci.
- Schmid, Hans-Jörg. 2007. "Entrenchment, Salience and Basic Levels, " in *Handbook of Cognitive Linguistics*. Oxford: Oxford University Press.
- Shippey, Tom. 2008. *Review of The History of The Hobbit*. Tolkien Studies. West Virginia University Press.
- Shustova, Svetlana V., Elena A. Osheva, and Konstantin A. Klochko. 2017. "Desemantization of Functional Grammatical Causatives in the Aspect of Grammaticalization," *XLinguae Journal*: 34–41
- Smadja, Frank. And Kathleen McKeown. 1990. *Automatically Extracting and Representing Collocations for Language Generation*. New York: Columbia University.
- Sopher, Hadas. 1962. *Sequence of Adjectives*. *ELT Journal*, Volume XVI, Issue 4. Florida Atlantic University.

Straubhaar, Sandra B. 2007. *Review of A Tolkien Mathomium: A Collection of Articles about J. R. R. Tolkien and His Legendarium*. Tolkien Studies. West Virginia University Press.

Tally, Robert T. Jr. 2019. *Demonizing the Enemy, Literally: Tolkien Orcs, and the Sense of the World Wars*. San Marcos: Texas State University.

Ten Hacken, Pius. ed. 2016. *The Semantics of Compounding*. Cambridge: Cambridge University Press.

Trichet, Juliette. 2016. *Philology in Tolkien's Work: The Creation of a Myth through the Power of languages*. Lyon: Université Jean Moulin Lyon 3.

APPENDIX 1: CORPUS

For the sake of clarity, the classification form of the first recorded lexeme in the corpus is analysed, so that the information can be understood comprehensively.

[1] **Barrow-downs** (noun, hyphenated, 1: contents)

- *noun* = word class
- *hyphenated* = compound word form
- *1* = the order number of the book within the 7-volume edition
- *contents* = the page number

Additionally, lexemes of a combined form contain a further specification of the compound forms present, which immediately follows the specification of form. The order of the individual forms resembles the order in which they appear within the compound. As an example, the lexeme (85) *fireside-story* is listed below to illustrate the corpus entries for combined form compounds.

[85] **fireside-story** (noun, combined: closed+hyphenated, 1: 55)

Lexemes (6) and (7) are an exception to the listing of compounding processes. Their non-prototypical and extraordinarily long form is not listed fully. Each compounding process that is identified is mentioned once, instead of listing all forms continuously. In all other lexemes of combined form, the order and number of the compound forms correspond to the forms of the lexeme in question.

[1] **Barrow-downs** (noun, hyphenated, 1: contents)

[2] **well-intentioned** (adjective, hyphenated, 1: note on the text)

[3] **long-expected** (adjective, hyphenated, 1: contents)

[4] **paperback** (adjective, closed, 1: x)

[5] **non-royalty-paying** (adjective, hyphenated, 1: x)

[6] **evileyed-blackhanded-bowlegged-flinthearted-clawfingered-foulbellied-bloodthirsty** (adjective, combined: closed+hyphenated, 6: 114)

[7] **Forestmanyshadowed-deepvalleyblack Deepvalleyforested Gloomyland** (other, combined: closed+hyphenated+open, 7: 138)

[8] **ten-fingered** (adjective, hyphenated, 1: xvi)

- [9] **outbreak** (noun, closed, 1: xvi)
- [10] **story-germ** (noun, hyphenated, 1: xvii)
- [11] **over-lap** (verb, hyphenated, 1: xvii)
- [12] **motor-car** (noun, hyphenated, 1: xix)
- [13] **Sandyman** (noun, closed, 1: xix)
- [14] **Hobbit-lore** (noun, hyphenated, 1: 1)
- [15] **Mirkwood** (noun, closed, 1: 4)
- [16] **highland** (noun, closed, 1: 4)
- [17] **good-natured** (adjective, hyphenated, 1: 2)
- [18] **water-mill** (noun, hyphenated, 1: 1)
- [19] **well-ordered** (adjective, hyphenated, 1: 1)
- [20] **well-farmed** (adjective, hyphenated, 1: 1)
- [21] **outside** (adverb, closed, 1: 6)
- [22] **handicraft** (noun, hyphenated, 1: 4)
- [23] **Mathom-house** (noun, hyphenated, 1: 7)
- [24] **well-to-do** (other, hyphenated, 1: 8)
- [25] **hobbit-building** (noun, hyphenated, 1: 9)
- [26] **hobbit-architecture** (noun, hyphenated, 1: 9)
- [27] **many-tunnelled** (adjective, hyphenated, 1: 10)
- [28] **Riddle-game** (noun, hyphenated, 1: 15)
- [29] **Longbottom** (noun, closed, 1: 11)
- [30] **Hobbitry-in-arms** (noun, hyphenated, 1: 13)
- [31] **dwarf-hoard** (noun, hyphenated, 1: 14)
- [32] **orc-mine** (noun, hyphenated, 1: 15)
- [33] **Undertowers** (noun, closed, 1: 19)
- [34] **odd-looking** (adjective, hyphenated, 1: 32)
- [35] **well-spoken** (adjective, hyphenated, 1: 28)
- [36] **tunnel-making** (verb, hyphenated, 1: 30)
- [37] **shoe-making** (verb, hyphenated, 1: 2)
- [38] **widespread** (adjective, closed, 1: 19)
- [39] **Bywater** (noun, closed, 1: 28)
- [40] **post-office** (noun, hyphenated, 1: 34)
- [41] **thunder-clap** (noun, hyphenated, 1: 35)
- [42] **eating-house** (noun, hyphenated, 1: 34)

- [43] **overlooked** (adjective, closed, 1: 34)
- [44] **goblin-barker** (noun, hyphenated, 1: 35)
- [45] **elf-fountain** (noun, hyphenated, 1: 35)
- [46] **dwarf-candle** (noun, hyphenated, 1: 35)
- [47] **life-size** (adjective, hyphenated, 1: 36)
- [48] **red-golden** (adjective, hyphenated, 1: 36)
- [49] **Proudfoot** (noun, closed, 1: 37)
- [50] **Bracegirdle** (noun, closed, 1: 37)
- [51] **Brockhouse** (noun, closed, 1: 37)
- [52] **merrymaking** (noun, closed, 1: 41)
- [53] **black-leather** (adjective, hyphenated, 1: 41)
- [54] **strong-box** (noun, hyphenated, 1: 41)
- [55] **front-door** (noun, hyphenated, 1: 46)
- [56] **waste-paper basket** (noun, combined: hyphenated+open, 1: 49)
- [57] **birthday-present** (noun, combined: closed+hyphenated, 1: 49)
- [58] **fireside-story** (noun, combined: closed+hyphenated, 1: 55)
- [59] **Hundred-weight feast** (noun, combined: hyphenated+open, 1: 55)
- [60] **East-West Road** (noun, combined: hyphenated+open, 1: 57)
- [61] **Greenleaf** (noun, closed, 3: 124)
- [62] **Gollum-creature** (noun, hyphenated, 1: 71)
- [63] **sharper-eyed** (adjective, hyphenated, 1: 70)
- [64] **Fire-mountain** (noun, hyphenated, 1: 81)
- [65] **hobbit-kind** (noun, hyphenated, 1: 69)
- [66] **overthrow** (verb, closed, 1: 69)
- [67] **eavesdropping** (verb, closed, 1: 83)
- [68] **grass-clipping** (noun, hyphenated, 1: 84)
- [69] **keen-eared** (adjective, hyphenated, 1: 70)
- [70] **river-bed** (noun, hyphenated, 1: 70)
- [71] **grass-border** (noun, hyphenated, 1: 83)
- [72] **wooly-footed** (adjective, hyphenated, 1: 202)
- [73] **Halfelven** (noun, closed, 1: 87)
- [74] **half-empty** (adjective, hyphenated, 1: 91)
- [75] **sandy-haired** (adjective, hyphenated, 1: 91)
- [76] **zig-zag** (verb, hyphenated, 1: 97)

- [77] **hemlock** (noun, closed, 1: 148)
- [78] **cobweb** (noun, closed, 1: 144)
- [79] **nightmare** (noun, closed, 1: 142)
- [80] **grey-green** (adjective, hyphenated, 1: 150)
- [81] **willow-leaf** (noun, hyphenated, 1: 152)
- [82] **willow-bought** (noun, hyphenated, 1: 152)
- [83] **over-power** (verb, hyphenated, 1: 154)
- [84] **nonsense-word** (noun, hyphenated, 1: 157)
- [85] **firewood** (noun, closed, 1: 156)
- [86] **willow-tree** (noun, hyphenated, 1: 154)
- [87] **willow-trunk** (noun, hyphenated, 1: 154)
- [88] **forget-me-not** (noun, hyphenated, 1: 162)
- [89] **standstill** (noun, closed, 1: 160)
- [90] **twig-finger** (noun, hyphenated, 1: 168)
- [91] **eyelid** (noun, closed, 1: 165)
- [92] **dew-drop** (noun, hyphenated, 1: 169)
- [93] **rain-water** (noun, hyphenated, 1: 169)
- [94] **raindrop** (noun, closed, 1: 173)
- [95] **rain-washed** (adjective, hyphenated, 1: 174)
- [96] **blackberry tart** (noun, combined: closed+open, 1: 204)
- [97] **lantern-light** (noun, hyphenated, 1: 200)
- [98] **morning-light** (noun, hyphenated, 1: 167)
- [99] **candlelight** (noun, closed, 1: 175)
- [100] **overhead** (adverb, closed, 1: 184)
- [101] **innkeeper** (noun, closed, 1: 199)
- [102] **slowcoach** (noun, closed, 1: 202)
- [103] **dark-green** (adjective, hyphenated, 1: 206)
- [104] **weather-beaten** (adjective, hyphenated, 1: 206)
- [105] **hobbit-talk** (noun, hyphenated, 1: 206)
- [106] **long-stemmed** (adjective, hyphenated, 1: 206)
- [107] **travel-stained** (adjective, hyphenated, 1: 206)
- [108] **hey-diddle-diddle** (other, hyphenated, 1: 210)
- [109] **deedle-dum-diddle** (other, hyphenated, 1: 211)
- [110] **Midyear's Day** (noun, open, 1: 224)

- [111] **woolly-footed** (adjective, hyphenated, 3: 195)
- [112] **lamplight** (noun, closed, 1: 228)
- [113] **inside shutters** (noun, open, 1: 230)
- [114] **cold hour** (noun, open, 1: 232)
- [115] **death's door** (noun, open, 1: 236)
- [116] **underfed** (adjective, closed, 1: 236)
- [117] **road-side** (noun, hyphenated, 1: 278)
- [118] **roadside** (noun, closed, 1: 229)
- [119] **graveyard** (noun, closed, 1: 272)
- [120] **caveyard** (noun, closed, 1: 272)
- [121] **paveyard** (noun, closed, 1: 272)
- [122] **shinbone** (noun, closed, 1: 272)
- [123] **thinbone** (noun, closed, 1: 272)
- [124] **heavy-hearted** (adjective, hyphenated, 1: 263)
- [125] **cloud-veiled** (adjective, hyphenated, 1: 263)
- [126] **half-asleep** (adverb, hyphenated, 3: 86)
- [127] **half-veiled** (adjective, hyphenated, 1: 263)
- [128] **half-stripped** (adjective, hyphenated, 1: 270)
- [129] **saddle-skirt** (noun, hyphenated, 1: 277)
- [130] **drinking-cup** (noun, hyphenated, 2: 20)
- [131] **looking-glass** (noun, hyphenated, 2: 13)
- [132] **summer-light** (noun, hyphenated, 2: 29)
- [133] **Windlord** (noun, closed, 2: 61)
- [134] **Horse-lord** (noun, hyphenated, 2: 61)
- [135] **Sword-that-was-Broken** (noun, hyphenated, 2: 69)
- [136] **Redhorn** (noun, closed, 2: 89)
- [137] **Cloudyhead** (noun, closed, 2: 89)
- [138] **Redhorn Gate** (noun, combined: closed+open, 2: 90)
- [139] **mountain-pass** (noun, hyphenated, 2: 93)
- [140] **ankle-deep** (adjective, hyphenated, 2: 96)
- [141] **mountain-side** (noun, hyphenated, 2: 97)
- [142] **cliff-wall** (noun, hyphenated, 2: 98)
- [143] **snowstorm** (noun, closed, 2: 98)
- [144] **snow-clad** (adjective, hyphenated, 2: 99)

- [145] **night-fall** (noun, hyphenated, 2: 104)
- [146] **nightfall** (noun, closed, 1: 90)
- [147] **highroad** (noun, closed, 2: 113)
- [148] **high-road** (noun, hyphenated, 1: 204)
- [149] **baggage-pony** (noun, hyphenated, 2: 114)
- [150] **dwarf-gate** (noun, hyphenated, 2: 119)
- [151] **Dwarf-door** (noun, hyphenated, 2: 116)
- [152] **dwarf-axe** (noun, hyphenated, 2: 110)
- [153] **will-o'-the-wisp** (noun, hyphenated, 2: 131)
- [154] **mithril-ring** (noun, hyphenated, 2: 135)
- [155] **door-way** (noun, hyphenated, 2: 144)
- [156] **doorway** (noun, closed, 2: 201)
- [157] **doom-boom** (other, hyphenated, 2: 146)
- [158] **Barrow-blade** (noun, hyphenated, 2: 144)
- [159] **orc-chieftain** (noun, hyphenated, 2: 144)
- [160] **man-shape** (noun, hyphenated, 2: 148)
- [161] **fir-wood** (noun, hyphenated, 2: 157)
- [162] **orc-voice** (noun, hyphenated, 2: 147)
- [163] **counter-spell** (noun, hyphenated, 2: 147)
- [164] **spear-thrust** (noun, hyphenated, 2: 148)
- [165] **mithril-coat** (noun, hyphenated, 2: 158)
- [166] **Moria-gate** (noun, hyphenated, 2: 161)
- [167] **beast-shape** (noun, hyphenated, 2: 176)
- [168] **waybread** (noun, closed, 2: 202)
- [169] **limestone-boulder** (noun, combined: closed+hyphenated, 2: 232)
- [170] **wizard-lord** (noun, hyphenated, 2: 241)
- [171] **S-rune** (noun, hyphenated, 3: 8)
- [172] **goblin-soldier** (noun, hyphenated, 3: 8)
- [173] **humpbacked** (adjective, closed, 3: 23)
- [174] **mountain-maggot** (noun, hyphenated, 3: 53)
- [175] **tree-herd** (noun, hyphenated, 3: 77)
- [176] **tree-talk** (noun, hyphenated, 3: 78)
- [177] **ent-stride** (noun, hyphenated, 3: 80)
- [178] **ent-house** (noun, hyphenated, 3: 81)

- [179] **Entwife** (noun, closed, 3: 82)
- [180] **Entmaiden** (noun, closed, 3: 87)
- [181] **beggar-man** (noun, hyphenated, 3: 110)
- [182] **Common Tongue** (noun, open, 3: 131)
- [183] **watchman** (noun, closed, 3: 133)
- [184] **watch-man** (noun, hyphenated, 2: 224)
- [185] **foot-hill** (noun, hyphenated, 3: 154)
- [186] **foothill** (noun, closed, 3: 4)
- [187] **herd-folk** (noun, hyphenated, 3: 156)
- [188] **herdfolk** (noun, closed, 3: 35)
- [189] **stouthearted** (adjective, closed, 3: 158)
- [190] **Westfold-man** (noun, combined: closed+hyphenated, 3: 161)
- [191] **thunderbolt** (noun, closed, 3: 173)
- [192] **half-orc** (noun, hyphenated, 3: 168)
- [193] **Entwood** (noun, closed, 3: 185)
- [194] **clang-bang** (noun, hyphenated, 3: 207)
- [195] **crash-crack** (noun, hyphenated, 3: 207)
- [196] **you-can't-have-it-so-go-to-sleep** (adjective, hyphenated, 3: 239)
- [197] **Wilderland** (noun, closed, 1: 4)
- [198] **Dunland** (noun, closed, 1: 4)
- [199] **Greenwood** (noun, closed, 1: 3)
- [200] **Buckland** (noun, closed, 1: 8)

APPENDIX 2: ABBREVIATIONS

OLD – Oxford Learner’s Dictionary

MWD – Merriam-Webster Dictionary

OED – Online Etymological Dictionary

APPENDIX 3: FIGURES, TABLES AND GRAPHS

[1]

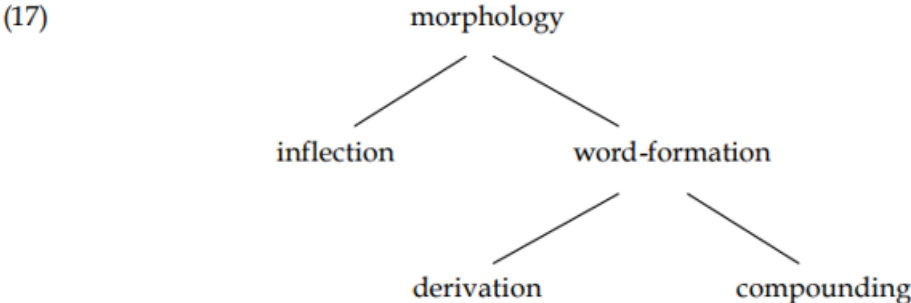


Figure 1 Scheme of the realm of morphology (Plag 2002, 22)

[2]

a black bird → *a black ugly bird*
a blackbird → *an ugly blackbird*

Figure 2 Illustration of the rule of adjacency

[3]

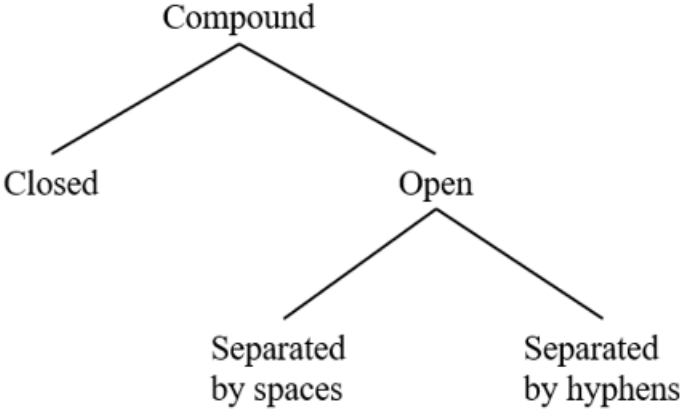


Figure 3 A hierarchy of compound forms from Multiword Expression Processing Survey

[4]

(3a)	wheelchair	602
(3b)	wheelchairs	178
(3c)	wheel-chair	10
(3d)	wheel chair	4
(3e)	wheel chairs	3
(3f)	WHEELCHAIR	797

Figure 4 Examples of the lexeme *wheelchair* in various compound forms (2010, 206)

[5]

simple to serve → *simple-to-serve* dishes
fresh from the farm → *fresh-from-the-farm* eggs

Figure 5 Illustrations of the hyphenation process from the original phrase

[6]

(18) Inventory of compound types, revised

	noun (N)	verb (V)	adjective (A)
noun	<i>film society</i>	<i>brainwash</i>	<i>stone-deaf</i>
verb	<i>pickpocket</i>	<i>stir-fry</i>	-
adjective	<i>greenhouse</i>	<i>blindfold</i>	<i>light-green</i>
preposition	<i>afterbirth</i>	-	-

Figure 6 A table of compound types based on word class from Plag

[7]

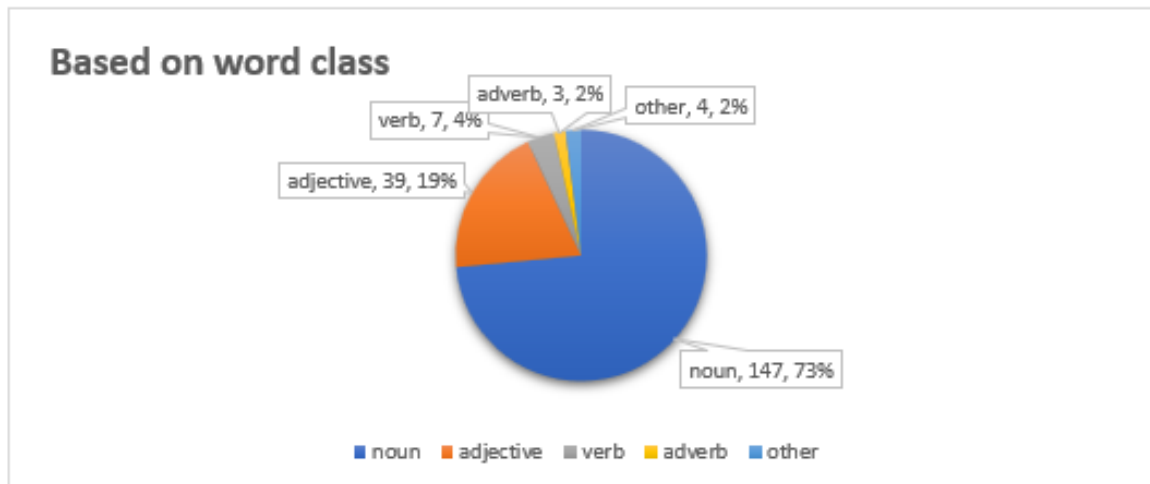


Figure 7 A graph presenting the distribution of word classes within 200 occurrences

[8]

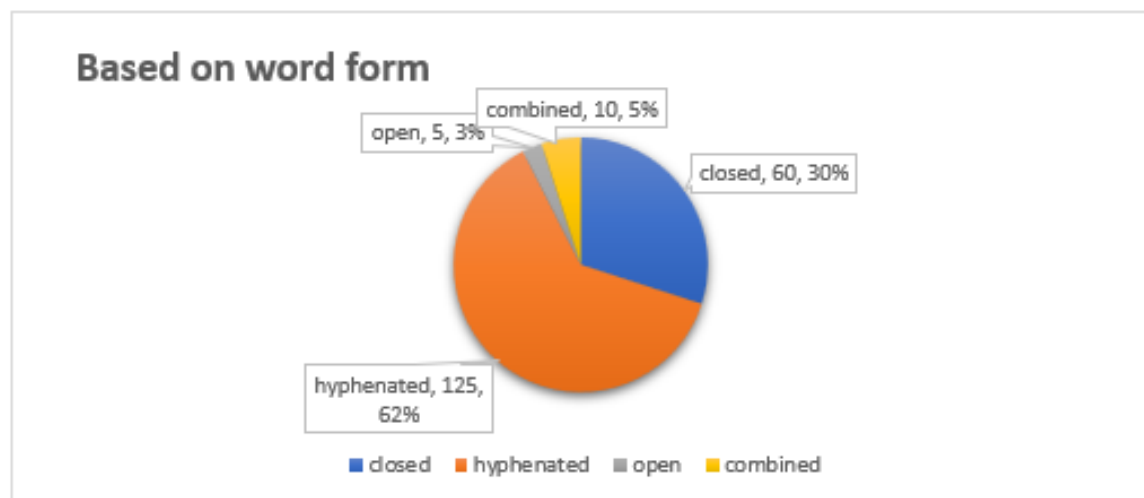


Figure 8 A graph presenting the distribution of word forms within 200 occurrences

[9]

(49) Proudfoot (13) Sandyman (50) Bracegirdle (51) Brockhouse

Figure 9 Examples of Hobbit surnames that are products of compounding

[10]

(113) *inside shutters* (114) *cold hour*
(115) *death's door* (182) *Common Tongue*

Figure 10 Examples of open compounds found in *The Lord of the Rings*

[11]

“The **cold hour** before dawn was passing.” (Tolkien in “The Rings Sets Out”, 2001, 232)

Figure 11 Sentential context for the lexeme (114) *cold hour*

[12]

children's hour girl's club no-man's-land child's play

Figure 12 Examples of descriptive genitives as presented in Altakhaineh (2016, 74)

[13]

“The chief danger is that the poor beast is probably at **death's door**.” (Tolkien in “The Rings Sets Out”, 2001, 236)

Figure 13 Sentential context for the lexeme (115) *death's door*

[14]

(88) *forget-me-not* (135) *Sword-that-was-Broken* (196) *you-can't-have-it-so-go-to-sleep*

Figure 14 Examples of quotational compounds found in *The Lord of the Rings*

[15]

“For there was a great inrush of those, *burárum*, those **evileyed-blackhanded-bowlegged-flinthearted-clawfingered-foulbellied-bloodthirsty**, *morimaite-sincahonda*, *hoom*, well, since you are hasty folk and their full name is as long as years of torment, those vermin of orcs...” (Tolkien in “The End of the Third Age”, 2001, 114)

Figure 15 Sentential context for the lexeme (6)

[16]

“It may be rendered ‘**Forestmanyshadowed-deepvalleyblack Deepvalleyforested Gloomyland**’, and by which Treebeard meant, more or less: ‘there is a black shadow in the deep dales of the forest’.” (Tolkien in “Appendices”, 2001, 138)

Figure 16 Sentential context for the lexeme (7)

[17]

(155) <i>door-way</i>	(156) <i>doorway</i>	(183) <i>watchman</i>	(184) <i>watch-man</i>
(185) <i>foot-hill</i>	(186) <i>foothill</i>	(187) <i>herd-folk</i>	(188) <i>herdfolk</i>
(145) <i>night-fall</i>	(146) <i>nightfall</i>	(147) <i>highroad</i>	(148) <i>high-road</i>

Figure 17 Examples of inconsistencies in word form