Analyzing Compound Structures in English and German

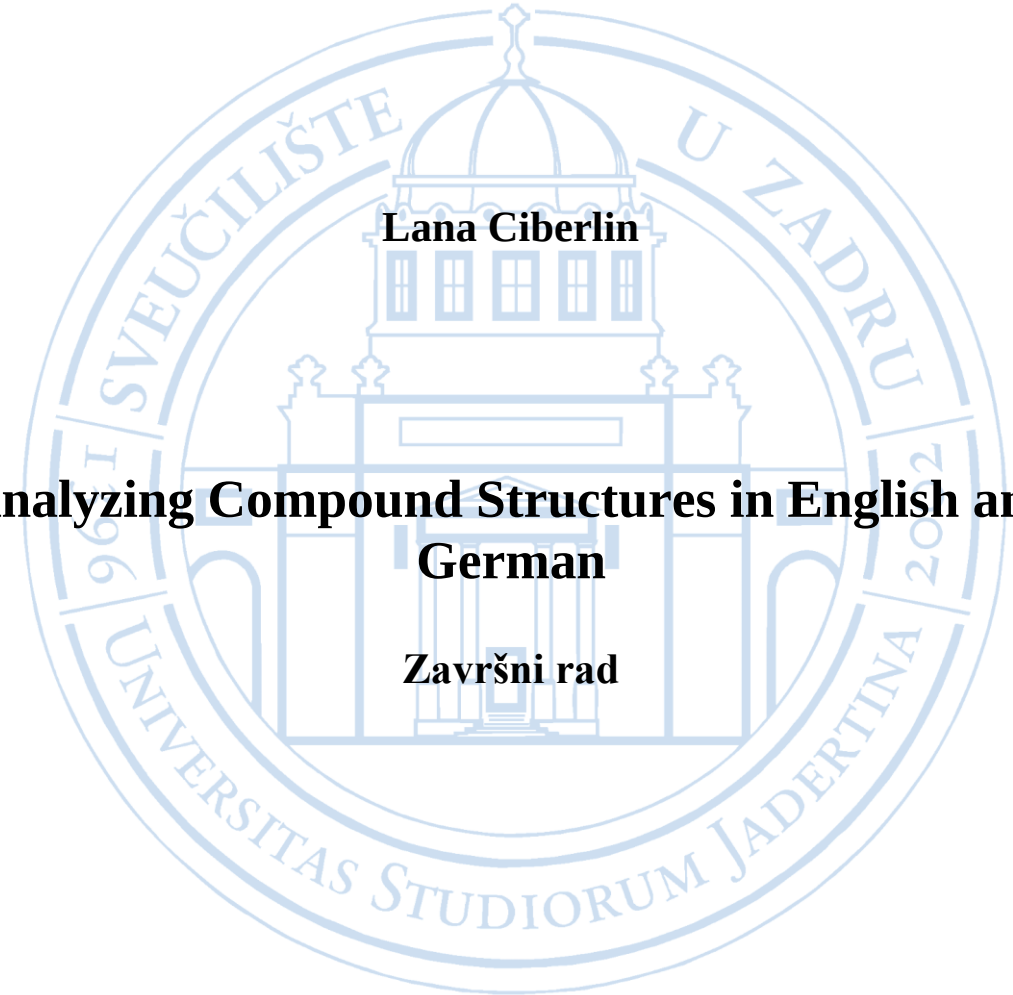
Ciberlin, Lana

Sveučilište u Zadru

Odjel za anglistiku

Sveučilišni prijediplomski studij

Anglistika



Lana Ciberlin

Analyzing Compound Structures in English and German

Završni rad

Zadar, 2024.

Sveučilište u Zadru
Odjel za anglistiku
Sveučilišni prijediplomski studij
Anglistika

Analyzing Compound Structures in English and German

Završni rad

Student/ica:
Lana Ciberlin

Mentor/ica:
Doc. Dr. Sc. Frane Malenica

Zadar, 2024.

Izjava o akademskoj čestitosti

Ja, **Lana Ciberlin,** ovime izjavljujem da je moj **završni** rad pod naslovom **Analyzing Compound Structures in English and German** rezultat mojega vlastitog rada, da se temelji na mojim istraživanjima te da se oslanja na izvore i radove navedene u bilješkama i popisu literature. Ni jedan dio mojega rada nije napisan na nedopušten način, odnosno nije prepisan iz necitiranih radova i ne krši bilo čija autorska prava.

Izjavljujem da ni jedan dio ovoga rada nije iskorišten u kojem drugom radu pri bilo kojoj drugoj visokoškolskoj, znanstvenoj, obrazovnoj ili inoj ustanovi.

Sadržaj mojega rada u potpunosti odgovara sadržaju obranjenoga i nakon obrane uređenoga rada.

Zadar, 18. rujna 2024.

TABLE OF CONTENTS

1. INTRODUCTION

According to Plag (2002), most productive way for creating new words in English language is compounding. It is an essential component in word construction in many languages, but English and German in particular. The history of each language adds to their expressiveness, flexibility, and complexity. Compounds are frequently employed to form new vocabulary that expresses meanings more precisely than basic words. Gaining knowledge about the compounds and processing their meaning, it can help one better understand their linguistic structures. In both English and German compounds are coined using various word classes, such as nouns, adjectives, verbs and prepositions. What is more, noun and adjective compounds are greatly common and vital for communication. English compounds can be single words, separated words, or words with hyphens, whereas German compounds tend to be lengthier and more complex. Despite their differences, both languages function on the right-hand head rule, according to which the right-hand element is the head which determines grammatical category and meaning of a compound. In this work, English and German compounds will be compared, starting with a general overview of their creation and how they are used in both languages. Afterwards the emphasis will be on noun and adjective compounds, which are two most used in these two languages. Lastly, the idea of semantic transparency will be examined, along with how it affects how compounds are understood and processed. This will show the distinctions and parallels between approaches to compound word construction and interpretation in English and German language.

2. COMPOUNDING

Compounding is a significant word-formation process in Germanic languages. Through compounds new words such as *woodcutter* or *writing-pen*, verbs like *side-step* or adjectives as *apple-green* are coined (Clark et al 1986: 7). As Lieber and Štekauer (2011) stated, it is difficult to define compounds because roots and stems are the components of compounding structure in certain languages whereas in some they are simply stems and roots and nothing more. Some researchers claim that when two or more words are joined to form a unit, we can refer to it as a compound. However, this definition of a compound comes from the absence of inflectional morphemes. What is more, the question how to distinguish compounds from phrasal forms arises since there are not enough widely acknowledged criteria for defining a compound. Lohde (2006) argues that compound structures are formed by joining two or more words and those are then called immediate constituents. According to Bauer (2020), the internal organization of compounds distinguishes them from other lexemes since they contain two or more lexemic bases that could function as the head of phrases on their own and could be inflected when they occur alone. The last component of a compound is usually where inflection is shown whether it be regular or irregular. This element could be referred to as the head of this compound. There are many compounds that appear in different dictionaries under various spellings.

According to Altakhaineh (2016), compounds in the German language are usually written as a single word. Only few instances can be found, where they are hyphenated as in *rot-grün* or *schwarz-rot-gold*. In the English language, on the other hand, this cannot be studied because some compound structures are written as single word like overflow, some with hyphen, as in ice-cream, others without a hyphen. Dixon (2014) claims that the roots that make compound structure can either be part of the same word-class like in instance *hen-house*, both constituents being nouns or they can be of different word-classes as in colour-blind with right-hand constituent being a noun and left-hand being an adjective. There are several types of compounds, one of which is the basic categorization which classifies compounds as they act as a distinct word-classes, those are compound nouns, adjectives, verbs and prepositions. Scalise and Bisetto (2011) distinguish root and synthetic compounds as well as endocentric and exocentric. They defined endocentricity and exocentricity as the presence, that is absence of the head and according to Benczes (2006), English language distinguishes two main types of compounds: endocentric and exocentric.

2.1 NOUN COMPOUNDS

Noun-noun compounds are the most widespread type of compound in the German language as well as in English (Olagunju 2010: 32). Noun compounds, just as it is claimed by Adams (2013), consist of a final nominal element modified by one or more additional parts that can be verbal, adjectival, or nominal when used independently. The author distinguishes five different categories; compound made of a noun and deverbal noun, two nouns, genitive -s and a noun, adjective and noun and exocentric. As for a compound made of noun and deverbal noun like bicycle-repairing, deverbal head cannot stand alone unless it is clear from the context is the modifier of the compound, which functions as the object in a similar sentence. Due to the fact that the verb-object link in these compounds only functions when the verb is transformed into a noun, they are frequently referred to as synthetic. Noun-noun compounds are frequently stressed on the modifying noun like advice centre or success rate. It should be noted that there is a difference between the use of genitive s in compounds and phrases. Although *Valentine's Day* or *Newton's Law* seem to name a specific person, the modifier in these instances actually identifies certain principle and date. Furthermore, certain adjectives function like nouns when they modify other nouns, particularly those that are derived from nouns like military, rural or herbal. Those are then not gradable as they do not alter their intensity. For instance, *arms sales* and *military sales* are identical in their structure even though *arms* is a noun and military an adjective. These combinations are referred to as compounds, i.e. they function as a single unit rather than individual words. On the other hand, German noun compounds are able to transmit a lot of information in a little space since they can be significantly lengthier than any other word-formation product. The length of German noun compound can be best seen in the example of *Rindfleischetikettierungsüberwachungsaufgabenübertragungsgesetz.* The interpretation of nominal compounds seems to be very flexible. For example, *Fischfrau* can have following meanings: woman who buys fish, wife of a fishman, woman who eats fish, woman who is as cold as a fish or woman who stands next to fish (Donalies 2005: 61-62). As most of the German compounds are noun compounds, depending on the context they can be composed of two, three or four different nouns, claims Schlücker (2012). For example, *Haupt* and *Stadt* are combined to form *Hauptstadt*. Nonetheless, German is recognised for its ability to form lengthier compounds as it could be seen in the example mentioned above. Due to the fact that German is inflected language, these compounds can be found in variety of forms according to how they are used in sentences, including cases and numbers as German has four cases and two numbers. According to Duden (2009), German noun compounds are in almost all instances written as one word although there are some examples where hyphen is used in order to make both pronunciation and understanding easier for speakers. Moreover, if the left-hand constituent is a

letter, formula or abbreviation as it is the case with *y-Achse, CO_2-Ausstoß* and *Kfz-Steuer*. Another example when hyphen is used is when there is listing like *Garten-* und *Campingmöbel* in order to shorten the sentence. To make the pronunciation easier, German noun compounds also have infix -e-, -s-, -es-, -n-, -en-, -er- or -ens-.

2.2 ADJECTIVE COMPOUNDS

As stated by Adams (2013), most adjective compounds end in -ing, as in *life-threatening*, or -ed, as it is the case with *security-coded*. Non-head element of a compound is usually noun, adjective and in some instances adverbs. The author divided adjective compounds in four categories: the ones ending in -ing, -ed ending, other suffixes and non-derived adjectives. Whereas certain noun and -ing ending compounds are acceptable and often employed, others are constrained by the meaning and word structures of the constituents involved. The semantic and syntactic meaning of the words determines the rules governing these formation process. Although most adjective compounds have -ing ending this type of compound has rather limited formation since some combinations may seem inappropriate or just do not work in the English language. That being said *library-working student* sounds unusual. On the other hand, compounds that relate to place or time like *ocean-going ship* do work. Combining two adjectives is a typical form of German adjective compound. For instance, *bitterkalt*. This compound makes the idea of coldness stronger. Another example is *hellblau* which refers to a particular shade of blue. Both adjective compounds provide a wider range of descriptive options in the language since they are more specific than their constituent words. When joined with a verb or a noun root they form compounds helpful, not only in everyday usage, but also in technical and scientific contexts like *eisfrei* or *wasserdicht*. Therefore, it could be said that adjective compounds in German language serve as some kind of intensifiers. Adjective compounds are constrained by the standard, as opposed to the almost unlimited noun compounds (Donalies 2005: 76-80).

3. CREATION OF NEW WORDS THROUGH COMPOUNDS IN ENGLISH AND GERMAN LANGUAGE

Compounding requires a distinction between the process that led to the creation of a particular form and the final product of word-formation. One of the more appropriate criteria for defining compounds is stress. English compounds are said to have stress on the left-hand constituent. Same rule can be applicable to German as well. However, there can be some inconveniences

with this. Specific groups of people as well as individual native speakers may pronounce some forms differently. What is more, the pragmatic circumstances and context in which a word is spoken might affect the way a particular compound is pronounced. Another way for defining and distinguishing a compound from phrase is according to syntactic criteria. If there is not a single additional element that can be placed in between the two elements, this form is considered to be a compound rather than a phrase. However, different languages have different syntactic criteria for defining a compound structure (Lieber, Štekauer 2011: 8-12). As it is claimed by Berg (2017), German phrases translate English compounds more often than English phrases translate German compounds. On the other hand, English phrases tend to be translated by German compounds more often than German compounds are used as a translation for English phrases. That is so because German is especially prone to compounding in comparison to English and other languages.

3.1 COMPOUNDS IN THE ENGLISH LANGUAGE

It is evident that people used the process of compounding since the period of Old English. During that period there were several types of compounds, the most numerous were nouns. They would contain two nouns, an adjective and a noun or simply verbal noun and a noun. In the Middle English period infix -s- was added in between two nouns as it is the case with doomsday. Another feature that was not existing in Old English is that pronouns were used along with a compound to determine the gender of it. During this period compounding was heavily influenced by the French language. Another way of creating a compound in Middle English was by adding words to determine the place like *bak* "back" and *doun* "down" in *back-side* and *downfall* (Podgorski 2008: 3-5). Plag (2002) argues that in the concept of modifier-head structure, the head, or the right-hand member, is modified by the left-hand constituent. The head of a compound is its most significant component as it determines the overall meaning of a compound. English compounds contain a crucial systematic feature regarding their head, better known as the right-hand head rule. Not every word from every word class can be joined with another to form a compound. That is why the majority of the compound's syntactic and semantic features are defined by its head. For example, if the compound contains noun as its right component, the whole compound will be classified as noun *(beer bottle).* Compound's head also has the nature to take the plural marking for the compound *(park commissioners).* Nominal compounds involve nouns, verbs and adjectives as the constituent that is not the head. In English, noun-noun compounds are the most widespread types of compounds. Certain noun-noun compounds are ambiguous and can be interpreted on a variety of ways based on some external factors like context in which it appears. Benczes (2006) claims that endocentric compounds are the hyponyms of the head element which means that in the structure apple tree

it is a kind of tree. On the other hand, exocentric compounds are rather metaphors or metonymy. Plag (2002) claims that forms such as *laser printer, book cover* or *letter head* can be classified as endocentric compounds since they are all noun-noun compounds that indicate a subclass of the referents of the head. One way to explain these compounds would be that their semantic head is located inside the compound. Compounds can be classified according to the IS A test. To explain this in other words, laser printer is a kind of printer and book cover is a kind of cover. Compound structures such as *greybeard, loudmouth* and *pickpocket* fall into the classification of exocentric compounds. They all refer to people and their semantic heads are considered to be outside of the compounds. In the English language, as well as in German, compounds are classified into specific word classes according to the right-hand head rule. The head is the central constituent of a compound and is expected to be on the right if we are to observe the right-hand head rule. A compound *dustbin* is a kind of bin and is endocentric, meaning that the head, which is right constituent, determines the word class of a whole compound. Expressions where no component appears to be central to the whole or of the same word class are referred to as exocentric as *paperback* or *lowlife*. It can be argued that life from the previous example is not the head of the compound as its plural form has *lowlifes* instead of *lowlives* (Adams 2013: 3-5). As it was mentioned earlier, stress in compounds can help determine whether a certain form is a compound or a phrase. Compounds are often stressed on the first element, whereas in phrases the stress is on the last word. This is best seen in the example of a compound *blackboard*, meaning "a board to write on" and phrase *a black board*, "a board that is black". Although this rule is applicable to most nominal compounds, there are a number of exceptions to the rule. Some of the exceptions are *geologist-astronomer, apple pie* and *summer night*. With noun-noun compounds it is crucial to remember that new words can be added to existing compounds to create new ones, for example *university teaching award committee member training* (Plag 2002: 171-181). Another example where the first element of a compound is stressed is *bathroom towel designer*. Compounds are represented by tree diagrams as it is Figure 1.

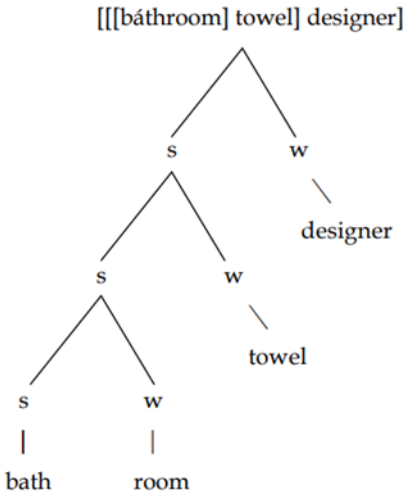


Figure 1: tree diagram bathrom towel designer

Other than endocentric and exocentric compounds Plag (2002) states that there is another kind of compound, those are copulative compounds. Neither of the two components of the compound appears to be more significant than the other. It is possible to describe them as having two conceptual heads, neither of which is inferior to the other, both members equally contribute to the meaning of the compound. The structure *poet-translator* can be interpreted as a person who is both a poet and a translator.

3.2 COMPOUNDS IN THE GERMAN LNGUAGE

Similar to English compounds, German compounds existed already in the 8th century. Though nominal, verbal, adjectival and prepositional compounds have all been documented, only nominal and adjectival compounds have met the most development up to the present days (Schlücker 2012: 4). In the German language, as stated by Donalies (2005), is compounding one of the main processes for forming new words. However, it should be noted that this is only applies to nominal and adjectival compounds. German compounds are made up of a minimum of two words *(Königsmantel)* and it is not uncommon for them to consist of more than two words *(Königsmantelfabrikant).* There are two main types of compounds, Determinativkompositum and Kopulativkompositum. Most widely used and most productive type of compounds is without a doubt determinative compound. They are made up of a root and determinative word or in German Grundwort and Bestimmungswort. In the compound, the two elements are not essentially interchangeable but are distinguished by subordination (Lohde 2006: 36). Donalies (2005) suggested that Determinativkompositum is the so-called normal case of compounds, meaning that it can be divided into two units as it is in the example of *Gartenhaus* or *Gartenhausidylle.* In the first example the two constituents are *Garten* and *Haus*, whereas in the second example first constituent is *Gartenhaus* and second one is *Idylle.*

German-speaking readers can break down compounds into coherent components. Therefore, segmentation is not a communication issue. However, linguists still deal with it since there are several ways in which they can be segmented. This is the best seen in the example of *Straßenbahnfahrer*. However, the boundaries of segmentation cannot always be clearly defined like it is the case in the example of *Kindergeburtstagsfeier*. This compound can be parted into *Kinder*, meaning "children" and *Geburtstagsfeier*, "birthday party". Then the meaning of the compound would be the birthday party for or with children. Another possibility of separation of the compound into constituents would be *Kindergeburtstagsfeier*, "children's birthday party", and *Feier*, "party" where the meaning of the whole compound is "the celebration on the occasion of children's birthday". In this type of compound, the order of the elements within a compound cannot be changed as it would change the meaning of that newly created word (Olagunju 2010: 24). With the help of the diagram three types of Determinativkomposita can be distinguished; left-handed, right-handed and bilaterally branched determinative compounds. An example of left-handed compound is *Tüllgardinenstange* as it is divided into *Tüllgardinen* and *Stange* and then first constituent consists of *Tüll* and *Gardinen* (Figure 2). Compound *Metallgardinenstange* is an instance of a right-handed compound (Figure 3) and *Edelmazipankonditortorte* is bilateral (Figure 4).

Figure 2: tree diagram Tüllgardinenstange

Figure 3: tree diagram Metallgardinenstange

Figure 4: tree diagram Edelmarzipankonditortorte

Joint elements in German compounds do not belong to the structure, they are semantically empty and serve as connection between two or more elements within a particular compound like -s- in *Arbeitsamt*. Donalies (2005) follows the example of English compounds and claims that there is also a Righthand Head Rule in German, which states that the right-hand element decides the syntactic use of a compound, what kind of word it belongs to, how it is inflected and so on. To understand this better it will be explained on the example of a compound *Kurtisanenschuh*. The whole compound is masculine since the last, that is right-hand element Schuh is masculine. The author distinguishes syntactic and semantic head of a compound. In determinative German compounds, the syntactic head is always the semantic head. Main accent is on the first element. In the German language, all nouns are written with a capital letter. In some compounds, only the first element is written this way however, there are instances when both elements are capitalised as in *KopfHörer*, but in recent years it can be seen that only the left element is capitalised. It is more common for German compounds to be written as one word even though there are some examples of them being hyphenated. As it is argued in Duden (2009) English borrowings can be written either as one word or two words like *Softdrink* or *Soft Drink*. Lohde (2006) claims that in spoken language, copulative compound words are far less frequent. The components of a compound are equal rather that subordinate to one another which is in contrast to endocentric compounds. Moreover, they are also known as additive or conjunctive compounds since they share the same status and of the same word class. Paraphrasing demonstrates equality because in the example of *Dichter-Diplomat* that person is both poet and a diplomat. These types of compounds are most commonly found in professional and technical expressions such as terminology related to fashion, newspaper language or science. What is more, they can be found in names with or without a hyphen such as *Hosenrock, Maler-Ortnithologe, Schwefellwasserstoff* or *Hans-Peter*. Olagunju (2010) stated that in most cases, the first element is stressed like in the examples *Hutschachtel.* However, both units are equally stressed in copulative compounds like *armenisch-deutsche-Beziehung*. Most of the examples used above can be classified as noun compounds as well as adjective compounds.

4. SEMANTIC TRANSPARENCY

Both English and German have a large number of compounds, which add to the complexity and adaptability of both languages. The degree of semantic transparency of these language units, which are created by combining two or more words, can vary. Semantic transparency has been defined as the extent to which each constituent's meaning is preserved in the compound's overall meaning while indicating the extent to which the meaning of a compound can be

predicted based on the meaning of its constituents at the same time. A compound's meaning might be understood immediately from its constituent parts, which would then be transparent meaning, like *bedroom*, or it can be more idiomatic and less clear, opaque, as in *hogwash* "rubbish". Acquiring the understanding of the balance between semantically transparent and opaque compounds is essential to appreciating the construction and processing of meaning in language (Gagné, Spalding 2016: 2). The objective is to investigate the ratio of semantically transparent to opaque compounds in English and German, with a particular emphasis on the impact of this difference on the frequency of usage. Given the structural differences in compounding processes between English and German, the comparison between two languages is especially interesting as English compounds tend to favour either hyphenation or open forms, whereas German compounds are more often merged as single words. There were several research conducted in the past with the aim of examining semantic transparency in the two languages.

The first research conducted by Frisson et al. (2008) is divided into two experiments. The aim of the first experiment was to determine whether reading time is impacted by un unspaced compound word's transparency, meaning a compound was written as a single word without spaces. Three different kinds of semantically opaque compounds were compared with matching sets of transparent compounds. First one is the one where second part of a compound word is related to the overall meaning and first part unrelated. Second is the compound whose meaning is related to the first part but not to the second. Last one is where neither of the constituents is related to the overall meaning of a compound. Authors came to conclusion that there was little variation in reading times between opaque and transparent compounds across all kinds. This shows that the overall meaning of a compound word is not created from its individual parts, even though the elements contribute to how it is processed. The purpose of the second experiment was to determine how much compounds written as two words affected the reading time and to identify any effects of transparency. They came to conclusion that transparency does affect processing when readers are forced to put together the meaning from the components. Both experiments showed that transparency has no effect on reading time for unspaced compounds since readers can recognize the words as entire words. Processing speed is impacted when the compound is spaced because readers must rely more on the transparency of the constituents to understand the word.

Similarly to the first research, Gagné and Spalding (2016) compare different sets to examine the semantic transparency. The difference is that they classified compounds into four categories: transparent-transparent, opaque-transparent, transparent-opaque and opaque-opaque, whereas Frisson et al. (2008) in only three categories. Participants were divided into

two groups to collect ratings on semantic transparency. On a computer screen, words were displayed in a random order. Using a slide bar that went from 0 (extremely unpredictable) to 100 (highly predictable), they judged how much the constituents contribute to the overall meaning first. Afterwards they evaluated each constituent separately based on how much they maintain their meaning in relation to the whole word. They concluded that the semantic transparency of a compound is not solely determined by how similar its constituents are to one another. This implies that transparency is not well perceived by the presence or absence of links between constituent and compound representation in one's mental lexicon (Gagné, Spalding 2016: 2-16). As for the German language, the first research on semantic transparency was conducted on native and non-native speakers, as Borgwaldt and Lüttenberg (2010) claim in their research. The focus is on how native and non-native speakers see German compounds and does improving second language's proficiency affect how and why semantic transparency is perceived. Forty native Russians who later learnt German as well as forty native German speakers rated the semantic transparency of forty German compounds. On a scale from 1 to 7, 1 being the weakest connection and 7 the strongest one, participants judged how strongly the meaning of each compound connected to its parts. Semantic transparency of the compounds was perceived similarly by non-native as it was by native speakers overall. However, the findings imply that a compound's head may be the most noticeable component to speakers who are less experienced, who will consequently interpret it as having a greater meaning. Furthermore, because of their more varied vocabulary, advanced speakers may begin to focus more on the specifying modifier and evaluate its relationship to the compound's meaning as the stronger one. Sagel et al. wanted to test the potential effects of lexical frequency and transparency on the time course of compound typing on 45 German participants who were instructed to write compounds. Semantic transparency has an influence on compound typing but there was a strong connection between a compound frequency and semantic transparency. It was discovered that frequency effects are seen in both opaque and transparent compounds, indicating the role of morphological component in compound formation. This implies that the morphological structure of compounds, regardless of their semantic transparency or opacity, has an impact on the processing system. What is more, assuming that transparent compounds have a meaning representation that is more strongly connected to their constituents and opaque compounds have a meaning representation that is more strongly connected to the word form as a whole, one can explain the effect of semantic transparency, which is found with low-frequency compounds.

5. SEMANTIC TRANSPARENCY RESEARCH

The idea of semantic transparency is believed to be essential to an explanation of how compounds are processed and represented in the mind (Libben 1998: 30). In this section, the conducted corpus-based study will be explained, in which average opacity rate of compounds in English and German was investigated. By using linguistic corpora, the study attempts to determine how transparent or opaque certain compounds are, offering the insights on the distinctions and parallels between the two languages.

5.1 RESEARCH QUESTIONS AND METHODOLOGY

As it can be seen from the section 4., semantic transparency is an important process for understanding and analysing compounds in English and German language. Both languages exhibit a vast number of compound words, yet their semantic transparency can differ greatly. Knowing how opaque or transparent these compounds are could help in understanding how various languages approach complex words (Libben 1998: 30-44). Since transparent compounds tend to be easier to understand and retain than opaque ones, the aim of this research is to investigate the variations in semantic transparency between English and German by comparing the average transparency scores of opaque compounds. By examining these variations, it can be determined whether one language has a tendency to have more opaque compound words than the other or whether there are notable changes between word production and interpretation processes in both languages.

The following questions aim to explore semantic opacity in both English and German:

1. What is the average opacity for compounds in the English language?
2. What is the average opacity for compounds in the German language?
3. Is there a difference between the two languages?

The corpus data used in this part of the paper were collected by using the digital corpora for English and German available via the Sketch Engine platform – the British National Corpus (BNC) for English and the German web corpus (deWaC) for German. BNC contains over 100 million tokens created using samples of written (newspapers, academic books, letters, essays) and spoken language (conversations, radio shows) of British English from the later part of the 20th century. The deWaC corpus contains over 1 billion tokens and was created using the methodology described by Kilgarriff et al. (2010). Although both corpora (BNC in particular) are less then impressive in terms of size, as Sketch Engine contains multiple corpora with over 5 billion words, and despite both corpora being over 15 years old, they were still considered relevant enough to be used in this research as the analysis was done on the top 100 most frequent compounds, which tend to be stable over time.

The data were collected using the queries like those in (1a) for English and (1b) for German. Compounds in English can be written as single words (e.g. greenhouse), two words joined with a hyphen (word-formation) and two separate words (dog owner), with the latter pattern being the most productive (Plag et al. 2008.). I choose to analyse the data only from the two-token compound patterns as the one in (1a). German compounds are written only as single words, which is why the single-token pattern was used for it.

(1a)

[tag="N.*" & tag!="NP"] [tag="N.*" & tag!="NP" & lemma="knife"]

(1b)

[tag="N.*" & tag!="NP" & lemma=".*messer"]

5.2 RESULTS AND ANALYSIS

The data of the semantic transparency research conducted on the top 100 most common compounds are organized into two sections: one representing English compounds (table 1) and the other representing compounds in German language (table 2).

Table 1: semantic transparency - English

	Bell	Book	Brush	Chain	Cloth	Frame	Holder	Knife	Shelf	Table
transparent	57	87	78	36	93	70	88	90	80	60
opaque	43	13	22	64	7	30	12	10	20	40
AverageOpaque	26,1									

Table 2: semantic transparency - German

	Buch	Burste	Glocke	Halter	Kette	Messer	Rahmen	Regal	Tisch	Tuch
transparent	81	91	46	85	20	69	27	87	65	89
opaque	19	9	54	15	80	31	73	13	35	11
AverageOpaque	34									

All the dataset followed the same criteria: transparent being actual, physical items with literal meaning and opaque being abstract representations or metaphorical meanings. Nouns that were the right-hand constituent of English compounds are *bell, book, brush, chain, cloth, frame, holder, knife, shelf, table*, as well as their German equivalents *Glocke, Buch, Burste, Kette, Tuch, Rahmen, Halter, Messer, Regal* and *Tish*. For every word pair, the number of opaque and transparent compounds are listed first, as well as the average opacity score. The following are some observations. In the table 1, the first row represents the number of compounds classified as transparent in the English language. The row that follows represents the number of opaque

compounds in the English language. Lastly, the row labelled as “AverageOpaque” contains an average value for opaque objects. Similarly, data for the German language are shown in table 2. The first row exhibits the number of transparent compounds, followed by the number of opaque compounds and the average opacity score is summed in the final row.

The assumption was that German compounds would be, on average, more opaque than English ones. Let us dig deeper into some word pairs in order to analyse the transparency. As it is the case with *bell* as right-hand constituent, it could be examined that 57 out of 100 most common compounds are classified as transparent. Compound words such as *church bell, door bell* or *hand bell* all exhibit the instances of transparent compounds since they all are actual object that can be either seen or held in hands and are, in fact, bells, unlike instance *leper bell*. Even though, at first, it might seem as semantically transparent compound, the meaning is rather opaque as *leper bell* has some historical background. It was not an actual bell, but its purpose was to warn others about the possibility of a person suffering from leprosy nearby. Furthermore, most compounds containing *brush* as right-hand constituent are classified as transparent as they have concrete, physical forms and purposes. This includes examples like *paint brush, hair brush* or *makeup brush*. Opaque compounds like purpose brush or tiredness brush are symbolic tools rather than actual objects, in fact, they symbolize abstract concepts. *Purpose brush* could serve as a tool or technique to help one concentrate on certain goals or *tiredness brush* may represent a figurative object used to handle tiredness. Unlike examples mentioned earlier, compound words with chain as their right-hand constituent showed greater number of opaque than transparent meanings. Compounds *iron chain, paper chain* or *safety chain* show that these word pairs have practical uses, whereas compounds *mountain chain, supermarket chain* and *food chain* are used figuratively to describe a sequence of something, in this case of mountain, supermarkets and food. The least opaque compounds are found under *cloth*. Everyday items like *tea cloth* or *velvet cloth* and *silk cloth*, as well as terms used in certain occupations like *spinnaker cloth* and *billiard cloth* are semantically transparent. Only a small number of compounds are opaque, those are *chin cloth* meaning it could be emotional barrier one puts up to hide insecurities and weaknesses since chin can be considered as symbol of strength and if one puts a cloth on it, it could be understood like that, or *fondant cloth* which could refer to as a surface used for aesthetic purposes if we consider the fact that fondant is usually used as cake decorating icing that is smooth and adjustable.

In the German language, the most instances of compounds with opaque meaning are *Kette* and *Rahmen* compounds. The ones containing Kette, have rather opaque meaning which shows that German tends to use more semantically opaque compounds, with most terms having meanings beyond their literal components, like *Gebirgskette, Ladenkette* or *Ursachenkette*. Similar to the

English language, they represent several sets of things like mountains, shops and causes. Even though they could be thought of as actual objects, here they are classified as opaque as they are not actual chains one could wear as fashion accessory or for other purposes. Another example with large number of opaque compounds is *Rahmen*. While only a small number of compounds like *Türrahmen, Bilderrahmen* or *Holzrahmen* can be seen as transparent, a lot of compounds such as *Zeitrahmen, Zielrahmen* and *Gesprächsrahmen* are not actual, real-life frames but conditions under which something is determines such as time, goal and conversation. The dataset for *Burste* and *Halter* exhibited more transparent than opaque compounds in contrast to previously mentioned *Kette* and *Rahmen*. Examples like *Zahnbürste, Haarbürste* and *Nagelbürste* have clear meaning, and they serve exactly the left-hand constituent implies, namely, "hairbrush" and "nail brush". Further example that contains more transparent compounds is *Halter*. The reason behind these data is that most nouns ending in -er indicate a person, so the whole term is transparent. A large number of compounds in this data, although transparent, is not an actual object but a living person, which shows that further research in this area is needed. However, compound *Fackelhalter* is considered opaque because he left constituent is a term that denotes a torch and when it is paired with the noun Halter, "holder" the meaning changes to a lamp. Although it is a real object, the exact meaning of this compound is not entirely clear.

The assumption was partially confirmed, although the difference in average numbers is small, opaque compounds make up on average 26,1 of the selected compounds, whereas this average in German is 34. In most instances, the number of semantically transparent compounds is significantly higher than the number of opaque ones. For example, *bell* has 43 opaque and 57 transparent compounds, but *book* only has 13 opaque ones. This suggests that the meaning of a large number of these compounds can be understood directly from their constituents. The compounds containing *book, frame, knife* and *holder* exhibit this predictable pattern. Although there are generally fewer opaque compounds, words like *chain* and *cloth* show a higher portion of opacity. This implies that compounds made up of these words might have developed meanings that are more difficult to determine from their constituent parts. Similarly, German compounds have a great degree of transparency. For instance, *Tuch* contains 89 transparent compounds and only 11 opaque, but *Buch* includes 81 transparent compounds and only 19 opaque. This shows that German compounds maintain distinct semantic relationship to their constituents for a great number of everyday objects or nouns. However, compounds with more opaque meaning like the ones containing *Glocke, Kette* and *Rahmen* indicate that their meaning is less predictable in certain contexts and supports the claim that German is semantically rich language that mainly relies on abstract meanings of compound words.

6. CONCLUSION

In conclusion, compounding is an important yet complex process in Germanic languages, with different criteria and meanings making it difficult to distinguish compounds from phrases. Both English and German frequently contain noun-noun compounds, however German compounds are longer. They can take on many forms based on grammatical cases and numbers, but they are usually written as a single word with the occasional hyphen or infix added. Adjective compounds are more limited in formation compared to noun compounds, but they serve as intensifiers and offer more specific meaning. Both languages exhibit a large number of compounds which can be classified according to the same or similar rules. The analysis of the semantic transparency in English and German in the last section shows that German has slightly more opaque compounds. Transparent compounds are typically easier to understand, but some compounds, particularly those involving abstract concepts tend to be rather opaque, highlighting semantic richness and complexity of a language. In order for data to be more precise, further research is suggested with greater number of participants, as more objective classification of compounds may result in different data.

7. REFERENCES

Adams, V. (2013). Complex Words in English. Routledge.

Bauer, L. (2020). Compounds and Compounding.

Benczes, R. (2006). Creative Compounding in English: The Semantics of Metaphorical and Metonymical Noun-Noun Combinations. John Benjamins Publishing Company.

Berg, T. (2017). Compounding in German and English: A quantitative translation study.

Borgwaldt, S., Lüttenberg, D. (2010). Semantic transparency of compound nouns in native and non-native speakers.

Clark, E. V., Fran Hecht, B., Mulford, R. C. (1986). Coining complex compounds in English: affixes and word order in acquisition.

Dixon, R. M. W. (2014). Making New Words: Morphological Derivation in English. Oxford University Press.

Donalies, E. (2005). Die Wortbildung des Deutschen: Ein Überblick. Gunter Narr Verlag.

Duden: Die Grammatik (2009), 8.Auflage, Band 4, Mannheim, CPI books GmbH

Frisson, S., Niswander-Klement, E., Pollatsek, A. (2008). The role of semantic transparency in the processing of English compound words. The British Psychological Society. Vol. 99,

Gagné. C., Spalding. T. (2016). Processing English compounds: Investigating semantic transparency. Canada: SKASE Journal of Theoretical Linguistics, vol. 13, no.2
Kilgarriff, A., Reddy, S., Pomikálek, J. and Avinesh PVS. A corpus factory for many languages. In LREC workshop on Web Services and Processing Pipelines, Malta, May 2010.
Libben, G. (1998). Semantic Transparency in the Processing of Compounds: Consequences for Representation, Processing, and Impairment. Brain and Language, 61, 30-44.
Lieber, R., Štekauer, P. (2011). The Oxford Handbook of Compounding. Oxford University Press.
Lohde, M. (2006). Wortbildung des modernen Deutschen. Narr Francke Attempto Verlag GmbH & Co.KG.
Mitib Altakhaineh, A. R. (2016): What is a compound? The main criteria for compoundhood. De Gruyter
Olagunju, Ezekiel, O. (2010). Eine vergleichende Analyse der deutschen und Yoruba Substantivbildung. Cuvillier Verlag.
Plag, I. (2002). Word-Formation in English. Cambridge University Press.
Plag, I., Kunter, G., Lappe, S., & Braun, M. (2008). The role of semantics, argument structure, and lexicalization in compound stress assignment in English. Language, 84, 760–794.
Podgorski, E. (2008). Word Formation Types: Compounding. GRIN Verlag.
Sahel, S., Nottbusch, G., Grimm, An., Weingarten, R. Written production of German compounds: Effects of lexical frequency and semantic transparency.
Schlücker, B., Gaeta, L. (2012). Das Deutsche als kompositionsfreudige Sprache: Strukturelle Eigenschaften und systembezogene Aspekte. De Gruyter.
Figure 1: Plag, I. (2002). Word-Formation in English. Cambridge University Press.
Figure 2, 3 and 4: Donalies, E. (2005). Die Wortbildung des Deutschen: Ein Überblick
Table 1 and 2: made by author

8. SUMMARY AND KEY WORDS

ANALYZING COMPOUND STRUCTURES IN ENGLISH AND GERMAN
Although definitions of compounds differ, compounding is a crucial step in the development of new words in Germanic languages. While English compounds can be written as single words, hyphenated words, or phrases, German compounds are often written as single words. Both German and English have a large number of compounds that differ in their semantic

transparency, which affects how easily their meanings may be understood from their parts. Transparency has no effect on reading time for compounds that are written as two words, but it does when readers must piece together meanings for the compounds. Lastly, the data for the semantic transparency research were collected by using the digital corpora for English and German and then later classified into transparent and opaque depending on whether they are actual objects with literal meanings or objects with abstract meaning. Even though there are many opaque compounds in both languages, the German language exhibited greater number.
Key words: compound, semantic transparency, IS A test, English, German

9. SAŽETAK I KLJUČNE RIJEČI

ANALIZA SLOŽENICA U ENGLESKOM I NJEMAČKOM JEZIKU

Iako se definicije složenica razlikuju, tvorba složenica ključan je korak u razvoju novih riječi u germanskim jezicima. Dok se složenice u engleskom jeziku mogu pisati keo jedna riječ, riječi sa spojnicom ili kao fraze, složenice u njemačkom jeziku često se pišu samo kao jedna riječi. Oba jezika imaju velik broj složenica koje se razlikuju u svojoj semantičkoj transparentnosti, što utječe na to Koliko se lako može razumijeti njihovo značenje. Transparentnost nema utjecaja na vrijeme čitanja složenica koje su napisane kao dvije riječi, ali ima kada čitatelji moraju sastaviti značenje složenica. Podaci za istraživanje o semantičkoj transparentnosti prikupljeni su korištenjem digitalnog korpusa za engleski i njemački jezik i kasnije klasificirani kao transparentni ili neprozirni prema tome jesu li stvarni predmeti s doslovnim značenjem ili imaju apstrakno značenje.
Ključne riječi: složenica, semantička transparentnost, IS A test, engleski, njemački

10. APPENDIX

Arbeitsamt, das – employment office
armenisch-deutsche-Beziehung, die – Armenian-German relation
Bilderrahmen, der – picture frame
bitterkalt – bitterly cold
Buch, das – book
Burste, die – brush
CO_2-Ausstoß, der - CO_2 emissions
Determinativekompositum, das – endocentric compounds
Dichter-Diplomat, der – poet diplomat

Edelmarzipankonditortorte, die – sweet and sour pastry cake / fine marzipan pastry cake

eisfrei – free of ice

Fackelhalter, der - lamp

Fischfrau, die – fish woman

Garten- und Campingmöbel, das – garden and camping furniture

Gartenhaus, das – garden house

Gartenhausidylle, die – garden house style

Gebirgskette, die – chain of mountains

Gesprächsrahmen, der – conditions of a conversation

Glocke, die - bell

Haarbürste, die – hairbrush

Halter, der – holder

Hauptstadt, die – capital city

hellblau – light blue

Holzrahmen, der – wooden frame

Hosenrock, der – culottes

Hutschachtel, die – hatbox

Kette, die – chain

Kfz-Steuer, das – motor vehicle tax

Kindergeburtstagsfeier, die – children's birthday party

Königsmantel, der – the king's coat

Königsmantelfabrikant, der– the king's coat manufacturer

KopfHörer, der - headphone

Kopulakompositum, das – copulative compound

Kurtisanenschuh, der – courtesan shoe

Ladenkette, die – chain of stores

Maler-Ornithologe, der – painter-ornithologist

Messer, das - knife

Metallgardinenstange, die – metal curtain rod

Nagelbürste, die – nail brush

Rahmen, der - frame

Regal, das - shelf

Rindfleischetikettierungsüberwachungsaufgabenübertragungsgesetz, das – beef labelingsupervision task transfer act

rot-grün – red green

schwarz-rot-gold – black red golden

Schweffelwasserstoff, der – hydrogen sulfide

Straßenbahnfahrer, der – tram driver

Tish, der – table

Tuch, das – cloth

Tüllgardinenstange, die – curtain railing

Türrahmen, der – door frame

Ursachenkette, die – chain of causes and effects

wasserdicht – waterproof

y-Achse, die – y-axis

Zahnbürste, die - toothbrush

Zeitrahmen, der - time span

Zielrahmen, der – targeted framework