Lexical-Semantic Modeling of Adjective-Noun Co-occurrences

Dissertation
zur
Erlangung des akademischen Grades
Doktor der Philosophie
in der Philosophischen Fakultät

der Eberhard Karls Universität Tübingen

vorgelegt von

Yana Strakatova

aus

Osch

2025

Gedruckt mit Genehmigung der Philosophischen Fakultät
der Eberhard Karls Universität Tübingen

Dekanin: Prof. Dr. Angelika Zirker

Hauptberichterstatter: Prof. Dr. Erhard. Hinrichs
Mitberichterstatter: Dr. Alexander Geyken

Tag der mündlichen Prüfung: 8. Januar 2025

Universitätsbibliothek Tübingen (TOBIAS-lib), Tübingen

Abstract

The main objective of this thesis is to classify the relations that hold between the elements of adjective-noun combinations. This classification makes a distinction between lexically free and lexically restricted phrases. For instance, in the free phrase *green dress*, the relation between the adjective *green* and the noun *dress* is `color`; in *light package* – `weight`. In a lexically restricted phrase such as *heavy smoker*, the relation is `intensity`; in *old friend*, the relation is temporal, more specifically, `duration`. This thesis claims that an adequate inventory of such relations can only be achieved through a synthesis of different theoretical frameworks. The investigation focuses on adjective-noun pairs as they have received considerably less attention in previous research than verbal constructions. The empirical basis for the research in this thesis is comprised of German data. The choice of language is motivated by the availability of rich linguistic resources developed for the German language that are suitable for data collection.

Using a sample of lexically varied data, the applicability of the following inventories in this task is investigated: Lexical Functions (Mel'čuk, 2012a), Qualia Roles (Pustejovsky, 1991), Frame Elements (Fillmore, 1982), and attributes (Barsalou, 1992; Löbner, 2014). The analysis reveals that each of these approaches alone is insufficient due to different reasons. Standard Lexical Functions can only be applied to a certain type of lexically restricted phrases, and Non-Standard Lexical Functions are too specific and fine-grained; Qualia Roles are coarse-grained and are limited to describing the semantics of concrete nouns. The inventory offered by Frame Elements is semantically broad and can accommodate nouns from various semantic fields. However, it has a disadvantage of being very fine-grained and, in part, inconsistent. Attributes overcome these issues, but they do not offer a specific inventory of labels that could be used for modeling. The semantic classification of adjectives in the German wordnet GermaNet (Hamp and Feldweg, 1997; Henrich and Hinrichs, 2010) is used as the basis for developing an adequate inventory of attributes.

The annotation scheme is developed in two stages: manually annotating the data with the attributes from GermaNet and modifying the labels for a better generalization relying on the insights from the above described frameworks. The resulting scheme consisting of 54 attributes is not exhaustive and can be extended depending on the data. In order to assess the adequacy of the scheme, an inter-annotator agreement is calculated, a qualitative analysis of the dataset is performed, and a series of machine learning experiments are conducted. Additionally, the scheme is successfully applied to an English dataset of adjective-noun phrases showing that the proposed inventory of relations is general and stable enough for describing different

and more diverse data. Finally, a sample of asymmetric adjective-noun pairs is compiled and the inventory of attributes is applied to it.

Contents

List of figures

List of tables

List of symbols, acronyms and abbreviations

Symbols and conventions:

?	anomalous expression
*	unacceptable expression

Abbreviations and acronyms:

AN	adjective-noun
FE	Frame Element
Fr.	French
Ger.	German
GL	Generative Lexicon
L2	second language
LF	Lexical Function
lit.	literal
LU	Lexical Unit
MTT	Meaning-Text Theory
MWE	multi-word expression
NLP	Natural Language Processing
NN	noun-noun
p.	page
pp.	pages
POS	part of speech
Rus.	Russian

Chapter 1

Introduction

> ...natural languages dissect reality differently and fill their words with different chunks of that reality.
>
> Raskin and Nirenburg (1995)

The main objective of this thesis is to classify the relations that hold between the elements of adjective-noun combinations. This classification makes a distinction between lexically free and lexically restricted phrases. For instance, in the free phrase *green dress*, the relation between the adjective *green* and the noun *dress* is `color`; in *light package* – `weight`. In a lexically restricted phrase such as *heavy smoker*, the relation is `intensity`; in *old friend*, the relation is temporal, more specifically, `duration`. This thesis claims that an adequate inventory of such relations can only be achieved through a synthesis of different theoretical frameworks. The investigation focuses on adjective-noun pairs as they have received considerably less attention in previous research than verbal constructions. The empirical basis for the research in this thesis is comprised of German data. The choice of language is motivated by the availability of rich linguistic resources developed for the German language that are suitable for data collection.

1.1 Object of study and motivation

The idea that vocabulary is not just a bag of words, but a system in which regular relations between its components can be established, emerged within structuralist semantics (Geeraerts, 2009). Such components are *word senses* and the relations between word senses are *lexical-semantic* relations. Traditionally, two types of lexical-semantic relations are distinguished: *paradigmatic* and *syntagmatic*.

Paradigmatic relations hold between similar elements that can be used in the same context and can replace each other. The most extensively studied relations include *hypernymy* (sometimes referred to as *hyponymy)*, *synonymy*, *antonymy*, and *meronymy*. Hypernymy holds between a more general and a more specific term. A hypernym is a more general term, for instance, *vehicle*; a hyponym is its subtype, for instance *car*: *car* is a type of *vehicle*. Synonymy

holds between senses that are identical or very similar and that are, thus, interchangeable in the same context. For instance, the adjectives *correct* and *right*. Antonyms have opposite meanings, such as *short* and *long*. Finally, meronymy is a *part-whole* relation that holds between *finger* and *hand* or *tree* and *forest*.

Syntagmatic relations hold between word senses in the same context. For instance, the relations that hold between the constituents of nominal compounds have been widely studied in recent decades.[1] This is illustrated in example (1) where the relations are INGREDIENT, OCCASION, APPEARANCE, and PRODUCTION METHOD.[2]

(1) a. *chocolate cake*: chocolate refers to an INGREDIENT of a cake
b. *birthday cake*: birthday refers to an OCCASION for a cake
c. *marble cake*: marble refers to an APPEARANCE of the cake
d. *sheet cake*: sheet refers to a PRODUCTION METHOD of the cake

Numerous inventories of compound relations have been developed, and there is no consensus among researchers on what the set of such relations is. The same issue holds for another type of syntagmatic relations, namely *semantic roles* (Gruber, 1965; Fillmore, 1968). Semantic roles (often referred to as *thematic roles*) reflect the relationships between predicates and "*the participants and objects involved in the event or state described by the predicate*" (Padó, 2007, p.4).

(2) [Leo]$_{agent}$ gave [the cake]$_{theme}$ to [Anja]$_{goal}$.

Example (2) illustrates typical semantic roles identified in a sentence in relation to the predicate: *agent*, *theme*, and *goal*. These roles map to the syntactic positions: *agent* to subject, *theme* to object, and *goal* to oblique object. The set of proposed semantic roles varies greatly from two *Proto-Roles* defined by Dowty (1991) to over one thousand *Frame Elements* distinguished in FrameNet.[3] It remains unresolved whether the set of semantic roles is limited and small, with each broad role mapping to one syntactic position, or they are more fine-grained, and multiple roles can be mapped to the same syntactic position (Ziegler and Snedeker, 2018). Large body of research is dedicated to semantic roles and has produced several lexical resources that contain semantic role information. Apart from FrameNets that are now available for a number of languages, there is a broad-coverage English verb lexicon VerbNet (Schuler, 2006) with 24 semantic roles and a large English corpus PropBank annotated with coarse-grained semantic roles (Gildea and Palmer, 2002; Kingsbury and Palmer, 2002; Palmer et al., 2005; Pradhan et al., 2022).

In his seminal work "The Generative Lexicon", Pustejovsky (1991, p.410) remarks "*The lexicon is not just verbs <...> we have little insight into the semantic nature of adjectival predication, and even less into the semantics of nominals*".

[1] For a comprehensive survey see Dima (2019).

[2] These compound relations stem from the dataset of German compounds described in Dima (2019).

[3] https://berkeleyfn.framenetbr.ufjf.br/, *last accessed on 04.02.2023.*

Even now, in comparison to verbs, there is still considerably less research on adjectives and nouns, and, more specifically, on lexical-semantic relations in adjective-noun pairs. As a result, there are fewer lexical resources that provide a broad and systematic coverage. FrameNet comprises both adjectives and nouns, but focuses primarily on event nouns. Noun predicates are annotated with semantic roles in PropBank's companion, NomBank (Meyers et al., 2004), where only nouns that take arguments were considered for annotation.

In order to understand what kinds of relations are typically described for adjective-noun combinations, consider the following examples:

(3) a. *agricultural equipment*: *equipment* USED IN *agriculture*
b. *heavy package*: *heavy* refers to the WEIGHT of the *package*
c. *heavy smoker*: *heavy* refers to the INTENSITY of *smoking*

(3a) is an example of a nominal phrase with a *relational* adjective as a modifier. Most relational adjectives are derived from nouns, and the nominal meaning is incorporated into their semantics. Thus, they do not denote qualities of their head nouns, but rather express objects: *agricultural* is related to *agriculture*. The relations between such adjectives and their head noun are similar to relations in nominal compounds.

Example (3b) illustrates the use of a typical *descriptive* adjective that assigns a value to a property or an *attribute* of a noun. Miller (1998, p. 48) describes this type of adjectives as follows:

> <...> to say *x is Adj* presupposes that there is an attribute A such that A(*x*)=*Adj*. To say *The package is heavy* presupposes that there is an attribute WEIGHT such that WEIGHT(*package*)=*heavy*. Thus, *heavy* and *light* are values for the attribute WEIGHT.

The adjective *heavy* is an acceptable value of the attribute WEIGHT for any *x* that can have that attribute: *cat*, *stone*, *bottle*, *etc.* Such phrases will be further referred to as *free phrases*.

Finally, (3c) is also an example of a descriptive adjective. However, it exhibits a behaviour that differs from that of *heavy* in example (3b). For one, the adjective does not directly specify an attribute of its head noun: a person cannot be characterized in terms of intensity. Rather, it describes the activity performed by the head noun, namely *smoking*. Secondly, the adjective does not combine freely with other nouns denoting persons by activity. A noun that is semantically very close to *smoker*, such as *drinker*, indeed can be modified by *heavy* referring to the INTENSITY of drinking: *heavy drinker* is a perfectly acceptable phrase in English. However, it is not possible to use the adjective in the same sense with, say, *worker*: the preferred adjectival modifier in this case is *hard*. Such adjective-noun phrases are examples of *collocations*, i.e. lexically restricted phrases that exhibit arbitrary behavior.

This thesis takes up the challenge of investigating what meanings are typically found in collocations (example 3c) and free phrases (example 3b)

by compiling a classification scheme of lexical-semantic relations that hold between the head noun and its adjectival modifier. The following subsections will elaborate on the notions of collocation and free phrase, clarify the motivation for studying collocations, and lay out the aims and contributions of this thesis.

1.1.1 The scope of study: collocations and free phrases

Adjective-noun (AN) **collocations** consist of a *base* (noun) and its *collocate* (adjective) where the choice of the collocate is *lexically restricted* by the base (Mel'čuk, 2012a). As an illustration of lexical restrictedness, consider the meaning of intensification mentioned above. In German, a large amount of adjectives can express intensification. Just to name a few examples: *helle Aufregung* 'great excitement' (lit. 'bright excitement'), *grober Fehler* 'grave mistake' (lit. 'coarse mistake'), *starker Raucher* 'heavy smoker' (lit. 'strong smoker'), *enger Freund* 'close friend' (lit.'narrow friend'), *stolzer Preis* 'high price' (lit. 'proud price'), *scharfe Strafe* 'severe punishment' (lit. 'sharp punishment'). The bases *Aufregung, Fehler, Raucher, Freund, Preis,* and *Strafe* can be intensified, but the speaker is not free to choose any adjective that expresses this meaning in German. The choice depends on the particular lexical unit that is the base of the collocation, hence, the choice is lexically restricted. Most attempts to swap the adjectives in these phrases would result either in awkward phrases such as (4a), (4b), and (4c), or in phrases in which the adjective has a completely different meaning such as (4d).

(4) a. ? *grobe Aufregung* lit. 'coarse excitement'
b. ? *heller Fehler* lit. 'bright mistake'
c. ? *enger Raucher* lit. 'narrow smoker'
d. *starker Freund* lit. 'strong friend'

Now consider the behaviour of an adjective in unrestricted or ***free* phrases**, for instance the adjective *grün* 'green' in the sense of 'green color'. Any object that can be described in terms of color can combine with the adjective *grün* in German or *green* in English: *dress*, *wall*, *pencil*, *liquid*, *field*, *hair*, *etc.* Based on this knowledge, it is possible to predict that, say, a *box* would also combine with the adjective *green*. In collocations, such predictions are either not possible, or limited.

Due to their unpredictable behaviour, collocations are often characterized as *idiosyncratic*. Translating a phrase into a different language serves as a classic illustration of collocations' idiosyncrasy. Consider one of the phrases listed above, *starker Raucher*. The literal meaning of the adjective *stark* is 'strong'. Example (5) shows that, in other languages, completely different adjectives combine with the noun *smoker* to express `intensity`.

(5) a. English: *heavy smoker* (lit. 'of great weight')
b. Russian: *zayadliy kurilshik* (lit. 'avid, enthusiastic')
c. Italian: *fumatore incallito* (lit. 'calloused')

d. Spanish: *fumador empedernido* (lit. 'hardened')

In each language, meanings are distributed among the adjectives differently. Each adjective contains different "chunks of reality", as the epigraph to this chapter says, and this is reflected in their collocational behavior.

There are two approaches to collocations: quantitative and qualitative, or, in terms of (Evert, 2004, p. 15), *distributional* and *intensional*, respectively. The former approach has grown immensely popular in recent decades. Within that framework, collocations are defined as *recurrent* word combinations. Whether a word combination is recurrent is determined by the frequency of the combination itself and the individual frequencies of its constituents which are calculated on the basis of a text corpus. Numerous statistical methods have been proposed for determining the degree of recurrence or *statistical association* of word combinations based on corpus frequencies. They will be discussed in detail in section 3.4. The qualitative approach, on the other hand, does not consider frequency a defining characteristic of lexicalized phenomena such as collocations, idioms, proverbs, proper nouns, etc. The status of a phrase is rather determined based on its meaning: idioms are characterized as *non-compositional*, free phrases are *compositional*, and collocations are usually placed somewhere in between. Opinions on the status of collocations vary among researchers depending on what their definition of compositionality is (this issue will be addressed in Chapter 3).

Evert (2004) claims that the two approaches are not necessarily orthogonal. Although it is not always explicitly mentioned within the intensional approach, collocations must be *recurrent* in order for them to be learned and stored by speakers as *ready-made* phrases. This thesis follows Evert's (2004) assumption and combines both approaches by defining collocations as recurrent word combinations in which one element (the base) is free and the second element (the collocate) is lexically restricted by the base. Free phrases, on the other hand, can be either recurrent or non-recurrent, and both the base and its modifier are lexically free (see section 3.5.2 for working definitions of collocations and free phrases). Both collocations and recurrent free phrases fall under the category of *co-occurrences* which will constitute the main focus of this thesis.

1.1.2 Why study collocations?

Collocations are traditionally studied in the context of multiword expressions (MWEs) along with idioms, compounds, proverbs, proper names, technical terms and others.[4] MWEs in general have always received a special treatment in language research and have always been a challenge due to their idiosyncratic syntactic and semantic properties. Collocations are probably the most notorious category of MWEs and there is a vast variety of interpretations of collocations. However, collocations are ubiquitous in language and play an important role in different fields of language study.

[4]There is no consensus among linguists on how MWEs and collocations are related to each other. This issue will be discussed in detail in section 3.2.

There are surprisingly few systematic studies of their semantics and this thesis aims at filling this gap.

Lexicography and lexical semantics. The initial motivation for this thesis comes from lexicography: an attempt to tackle the issue of grouping collocates of a given word based on semantic criteria. Lexicographers have always taken a special interest in collocations. They are mainly interested in two issues: how to select relevant collocations for a dictionary and how to represent them. In most dictionaries, the main underlying principle for grouping the selected collocates is based on their part of speech. A number of specialized collocation dictionaries list the collocates in a particular order that is semantically motivated, but the semantic categories they suggest are either non-systematic or highly coarse-grained.

In computational lexicography, research has focused mostly on improving the methods of automatic collocation extraction. Various statistical association measures have been proposed and successfully applied to retrieve collocations from corpora (Evert, 2008; Pecina, 2008; Rychly, 2008; Gries and Durrant, 2021). However, long unsorted lists of collocations do not satisfy the needs of dictionary users. Some sort of semantic classification of collocations is required.

A lexical-semantic theory that provides such a system was developed by Mel'čuk (1996), more specifically, the inventory of Lexical Functions (LFs) which will be discussed in detail in section 4.2.1. LFs were integrated into lexicographic works initially in Russian (Mel'čuk and Zholkovsky, 1984) and later in other languages. For instance, an online dictionary of Spanish collocations DiCE (Alonso Ramos et al., 2009; Vincze et al., 2011) and French Explanatory Combinatorial Dictionary (Mel'čuk et al., 2007). However, the inventory of standard Lexical Functions for adjective-nouns collocations is limited to five categories and needs to be expanded. This thesis investigates inter alia the limitations of the LF inventory and explores other theoretical frameworks that might remedy its shortcomings. From the point of view lexical semantics, the relevance of this research lies in the possibility of uncovering novel relations in lexically restricted phrases.

Second language (L2) learning. As illustrated by example (5), collocations are language-specific. A correct and frequent use of collocations is a sign of fluency and proficiency of a language speaker. Due to their arbitrariness, collocations, and especially their production, pose a serious challenge for language learners, and L2 learners make considerably more mistakes in collocations than in free combinations and idioms (Nesselhauf, 2003). In contrast to free phrases, the semantics of collocations is partially opaque, i.e. one of its constituents is used in a figurative sense which is not always interpreted correctly by the learners. In comparison to idioms, collocations are more flexible, and are perceived as less restricted than idioms by L2 learners (Fioravanti et al., 2021). For instance, one of the most typical mistakes of L2 learners in verbal collocations is the wrong choice of a verb as in **to make homework* instead of *to do homework* or **to carry out races* instead of *to hold races* (Nesselhauf, 2003). Since it is the verb that is used in a restricted

sense, Nesselhauf (2003) suggests that in teaching such collocations, the focus should be on the verb. For example, a learner should be aware that the verbal collocate *to run* combines with the nominal base *risk*, but not its synonyms *danger/peril*. Likewise, adjective-noun collocations may be very tricky for L2 speakers. For instance, for the native speakers of German or Russian, there is a danger of choosing the wrong adjectival modifier in English when describing the intensity of *rain*. A possible mistake would be choosing the adjective *strong*: **strong rain* instead of *heavy rain*. The corresponding collocation in Russian is *sil'nij dozhd* and in German *starker Regen* with literal meaning 'strong rain'. A reliable and robust semantic classification scheme of collocations is of benefit for teaching and learning collocations in a systematic way.

Natural Language Processing. Multi-word expressions in general are of relevance in all NLP tasks and applications concerned with semantic processing such as sentiment analysis, word sense disambiguation, semantic parsing, machine translation, etc. (Cordeiro et al., 2019). Due to their idiosyncrasies, MWEs remain a challenge even for the modern distributional models. Shwartz and Dagan (2019) show that, although state-of-the-art word representations reflect idiosyncrasies in meaning, they struggle with revealing implicit meanings such as relations in nominal compounds and adjective-noun phrases.

This issue remains understudied, and especially little attention has been paid to collocations. The findings in Espinosa-Anke et al. (2019) support the hypothesis that, compared to standard lexical-semantic relations such as hypernymy or meronymy, the relations between the constituents of collocations are especially challenging. Espinosa-Anke et al. (2019) investigate the performance of different language models at understanding collocations. Their results show that even context-aware models struggle with collocations.

Research in NLP requires reliable gold standard data for evaluation. This thesis provides two manually annotated datasets for German that can be used in various machine learning experiments concerned with collocations or adjective-noun phrases in general.

1.1.3 Objective and contributions

The main goal of this thesis is to define a semantically diverse inventory of lexical-semantic relations in adjective-noun collocations and free phrases. This will be achieved through a synthesis of relation inventories provided by different theoretical frameworks. Chapter 4 presents four theories that might provide the required inventory. Each of them will be applied to a semantically varied sample of semi-automatically selected German adjective-noun pairs to identify gaps in their inventories of relations. The following theories will be investigated:

- the Meaning-Text Theory (Mel'čuk, 1996, 1999, 2015, 2023) with the inventory of Lexical Functions;

- the Generative Lexicon (Pustejovsky, 1991) and the Qualia Roles;
- Frame Semantics (Fillmore, 1982) and Frame Elements;
- Concept frames (Barsalou, 1992; Löbner, 2014) and attributes based on the semantic classification of adjectives as adopted in the German wordnet GermaNet (Hamp and Feldweg, 1997; Henrich and Hinrichs, 2010).

An adequate relation inventory should be able to accommodate semantically varied data independently of the language and be flexible in terms of granularity. Based on the insights from the small-scale study described above, an initial set of relations will be proposed. In order to test the adequacy of the proposed inventory on a large scale, a dataset of German adjective-noun phrases will be compiled. The construction of the data set is performed in joint work with colleagues from the Brandenburg Academy of Sciences (BBAW) Isabel Fuhrmann, Dr. Alexander Geyken, and Dr. Lothar Lemnitzer. The robustness of the inventory will be tested by measuring inter-annotator agreement, conducting machine-learning experiments, and applying the inventory to a different dataset in English.

The main contributions of this thesis are the following:

- Comparison of four theoretical frameworks with respect to lexical-semantic modeling of adjective-noun collocations and free phrases.
- Manually annotated gold standard data set of German adjective-noun co-occurrences consisting of collocations, free phrases, idioms, proper names and technical terms.
- Guidelines on how to differentiate between free phrases, idioms, and several types of collocations.
- Semantically broad inventory of relations that hold between the constituents of adjective-noun phrases for both collocations and free phrases. The differences between the two types of phrases in terms of the relations between their constituents are outlined.
- Two gold standard data sets of adjective-noun collocations and free phrases manually annotated with the relations from the new inventory: a novel German data set and a re-annotated English dataset.
- Corpus study of directional statistical association measures. Such measures are effective at extracting asymmetric collocations, technical terms and proper names.

1.2 Chapter guide

The thesis is divided into nine chapters. **Chapter 2** provides a short theoretical background on the semantics of adjectives in general and gives a concise grammatical description of German adjectives.

Chapter 3 introduces the concept of collocation in detail and discusses different aspects of the phenomenon. In the first part of the chapter, various theoretical approaches to collocations and MWEs in general are presented. The second part focuses on the statistical properties of collocations. Working definitions of collocation and free phrase are given based on the collected insights from the literature. The chapter ends with a survey of existing collocation resources.

Chapter 4 discusses the theoretical frameworks considered for constructing an annotation scheme for adjective-noun collocations and free phrases. It focuses on the Meaning-Text theory, the Generative lexicon, Frame Semantics, and concept frames as well as the semantic classification of adjectives in English and German wordnets. The chapter presents a pilot study conducted on a sample of semantically diverse data to evaluate the theories. The limitations and advantages of each theory are discussed.

Chapter 5 presents the details of the construction of the dataset GerCo (German Collocations) that contains adjective-noun collocations and free phrases. First, the data collection and the annotation scheme are described, followed by the discussion of the annotation results evaluated via an inter-annotator agreement study. A qualitative analysis of the final dataset with respect to the assigned labels provides further insights about the differences between various types of phrases identified in the dataset. Finally, the robustness of the binary classification is tested in a series of machine-learning experiments.

In **Chapter 6**, the dataset GerCo described in Chapter 5 is further annotated with respect to the relations that hold between the base and its collocate. The annotation scheme and guidelines are based on the findings of the pilot study presented in Chapter 4. Similarly to the previous chapter, the set of labels used for annotation is evaluated based on the inter-annotator agreement, the qualitative analysis of the dataset, and in a number machine-learning experiments.

In **Chapter 7**, the robustness of the developed annotation scheme is tested by applying the set of labels to a different dataset of adjective-noun phrases: the semantically diverse English dataset HeiPLAS.

In **Chapter 8**, the discussion turns to *statistically asymmetric* adjective-noun combinations. For instance, in the asymmetric phrase *himmelweiter Unterschied* 'huge difference', the adjective is more predictive of the noun than vice versa. A few statistical association measure have been proposed to measure such asymmetries. The chapter will present a corpus study that investigates the directional association measure ΔP applied to adjective-noun pairs in German.

Chapter 9 concludes the thesis with a discussion about the relation between the inventory compiled in this thesis and the theories that inspired it, gives a short summary of the work, and outlines possible future work.

Chapter 2

Adjectives: theoretical background

This chapter provides a short overview of theoretical backgound of adjectives. First, various approaches to describing adjectival semantics will be presented in section 2.1. In the second part of the chapter (section 2.2), adjective as a part of speech in the German language will be discussed.

2.1 Semantics of adjectives

Boleda (2006) provides a comprehensive overview of various approaches to adjectival semantics. This section will summarize Boleda's (2006) theory review focusing on various semantic classifications of adjectives. There are three major groups of approaches to semantics of adjectives: within formal semantics, within descriptive grammar, and within computational linguistics.

Formal semantics

In formal semantics, three categories of adjectives are distinguished based on their entailment pattern: **intersective, subsective**, and **non-subsective** (Boleda, 2006). For intersective adjectives (also referred to as *absolute* or *restrictive*), *"the set of things denoted by the adjective-noun phrase is the intersection of the set of things denoted by the adjective and the set of things denoted by the noun"* (Peters and Peters, 2000, p.2). For instance, a *yellow umbrella* belongs to the set of yellow things and to the set of umbrellas. The meaning of subsective (or *relative*) adjectives is dependent on the noun they modify: *a small book* belongs to the set of books, but the notion of *small* is relative: a small book may not belong to a set of small things. The distinction between the two categories is quite vague (Boleda, 2006) and sometimes they are treated as one category, as in Pavlick and Callison-Burch (2016).

Finally, non-subsective (or *intensional*) adjectives do not refer to a subset of things their head noun denotes. For instance, *a former president* is not a president anymore. Pavlick and Callison-Burch (2016) further divide this class into two categories: **privative** and **plain**. Privative adjectives such as *fake* or *former* are truly non-subsective because a fake thing is not actually that thing anymore: *fake gold* is not actually gold, and *former president* is not a president anymore. Plain subsective adjectives, on the other hand, allow an interpretation where the set of things denoted by the adjective-noun phrase

intersects with the set of things denoted by the noun. For instance, an *alleged criminal* may or may not be an actual criminal.

Descriptive grammar

In descriptive grammar, adjectives are classified into **relational** (also referred to as relative, denominal, or pertainyms) and **descriptive** (or qualitative) adjectives (Warren, 1984; Raskin and Nirenburg, 1998; Boleda, 2006).

Most relational adjectives are derived from nouns, so the nominal meaning is incorporated into the semantics of such adjectives. Thus, they do not denote qualities of their head nouns, but rather express objects:

(6) a. *financial* – is related to *finances*
b. *agricultural* – is related to *agriculture*

Typically, relational adjectives are rarely found in predicative position as in example (7b), they are non-gradable as in example (7c) and are not used in coordination with descriptive adjectives (Warren, 1984) as in example (7d):

(7) a. agricultural workers
b. * workers who are agricultural
c. * very agricultural workers
d. * smart and agricultural workers

Boleda (2006) remarks that being derived from a noun is not obligatory for an adjective to be considered relational. For instance, the adjectives *urban* and *rural* are not denominal. However, they are relational adjectives as they incorporate the nominal meaning in their semantics: 'related to the city' and 'related to the country' respectively.

Descriptive adjectives are referred to by Hartung (2015) as *attribute-denoting adjectives*. Miller (1998, p. 48) describes this type of adjectives as follows:

> <...> to say *x is Adj* presupposes that there is an attribute A such that A(*x*)=*Adj*. To say *The package is heavy* presupposes that there is an attribute WEIGHT such that WEIGHT(*package*)=*heavy*. Thus, *heavy* and *light* are values for the attribute WEIGHT.

The adjective *heavy* is an acceptable value of the attribute WEIGHT for any *x* that can have that attribute: *cat*, *stone*, *bottle*, *etc.* Another example of a descriptive adjective is *yellow*: in the phrase *yellow umbrella*, the adjective is the value of the attribute `color` of the noun *umbrella*. Other properties or attributes that an *umbrella* can have are:

(8) a. `size` – *small umbrella*
b. `weight` – *heavy umbrella*
c. `function` – *broken umbrella*

Examples (7) and (8) illustrate the difference between the nature of relation between the modifier and the head for relational and descriptive adjectives: *agricultural workers* are workers who *work in* agriculture, the adjective in this case does not describe any property of the noun. Rather, there is a relation *work_in* between the two words. This makes adjective-noun phrases with relational adjectives as modifiers semantically close to nominal compounds. Vivid illustrations for this observation are found in languages that do not rely on compounding as much as English and German, for instance Russian. The role of nominal modifiers in Russian is frequently taken over by relational adjectives as in (9a) and (9b). In the first example, the relation between the head and its modifier is of PART-WHOLE type: a *handle* is a PART of a *door*, in the second one, the relation is INGREDIENT: *milk* is an INGREDIENT of a *cocktail*.

(9) a. *dver-n-aya* *rutchka*
door-ADJ-INFL handle
'door handle'

b. *moloch-n-iy* *kokteyl*
milk-ADJ-INFL cocktail
'milk cocktail'

Computational linguistics

Raskin and Nirenburg (1998, pp.167–170) propose a microtheory of *Ontological semantics* which distinguishes three classes of adjectives:

- **scalar**, based on property ontological concepts; for instance, the adjective *easy* whose meaning is placed on the scale EASY/DIFFICULT
- **deverbal**, based on process ontological concepts; for instance, the adjective *eager* whose lexical entry is based on that of the verb *want*
- **denominal**, based on object ontological concept; for instance, the adjective *provincial* in the sense *of a province*, whose lexical entry is derived from that of the noun *province*.

Note that Raskin and Nirenburg (1998) do not equate the term *scalar* with the term *gradable*. The latter refers to a characteristic that is typical of various adjectives from any of the three classes presented above: gradable adjective are those that can be used comparatively. Thus, there are scalar adjectives such as *male* that are not gradable, and gradable adjectives that are not scalar, for instance the denominal adjective *administrative*.

(10) His style is more administrative than hers.

Raskin and Nirenburg (1998, pp.172) claim that there it is not difficult to come up with contexts in which such an adjective can be used comparatively as in example (10).

A similar classification is proposed by Boleda (2006). Based on the results of a series of unsupervised experiments Boleda (2006, pp.173–174), proposes the following classification of adjectives:

- **basic** adjective express attributes, for instance the adjective *yellow* denotes a attribute COLOR, and *heavy* denotes WEIGHT;
- **event-related** adjectives such as *comfortable* express a relationship to an event;
- **object-related** adjectives such as *political* and *biological* express a relationship to an object.

Hartung (2015) conducts a number of annotation studies relying on this classification and comes to the conclusion that the distinction between basic and event-related adjectives is too fine-grained. For instance, in the phrase *fast horse*, the adjective can either be interpreted as basic as it refers to the SPEED of a horse, or it can be also considered event-related because a fast horse is a horse that *runs* fast. Due to a large number of such ambiguities, the two categories had to be conflated into one class of property-denoting adjectives leading to a binary distinction into **attribute-denoting** and **object-related** adjectives.

Lexical-semantic resources such as the Princeton WordNet Fellbaum (1998) and GermaNet Hamp and Feldweg (1997); Henrich and Hinrichs (2010) also distinguish between two classes of adjectives: **descriptive** such as *heavy* and *yellow* and **relational** (also referred to as *pertainyms*) such as *musical* and *dental*. Descriptive adjectives are basically values of attributes, whereas relational adjectives *"function as classifiers"* and thus "play a role similar to that of a modifying noun" (Miller, 1998, p.59).

Summary

This section presented approaches to semantic classifications of adjectives in formal semantics, descriptive grammar, and computational linguistics. The latter two approaches are very similar as they distinguish two main groups of adjectives based on the type of relation between the adjective and the modified noun. Descriptive adjectives describe properties or attributes of modified nouns, whereas relational adjectives behave similarly to nouns and denote objects. Adjective-nouns combinations with a relational adjective as modifier are thus similar to noun-noun compounds. As mentioned in Chapter 1, there is extensive research on relations in noun-noun compounds and numerous relation inventories have been suggested. It seems plausible that such inventories can also accommodate AN combinations with relational adjectives and it is outside of the scope of this thesis to test this hypothesis. The main focus of the thesis lies on the descriptive adjectives. However, relational adjectives can acquire a different meaning and function as descriptive adjectives. For instance, the adjective *wooden* has a second meaning "awkward/stiff" as in *wooden performance*. For this reason, relational adjectives will be considered in this thesis if they function as descriptive ones.

2.2 Attributive adjectives in German

Grammatical characteristics of attributive adjectives in German are defined through *number* (singular or plural), *gender* (masculine, feminine, neuter), and *case* (nominative, genitive, dative, accusative). These properties are expressed in adjectives by means of inflection.

(11) a. *klein**es** Kleid* 'small dress'
b. *etwas klein**es** Kleid* 'somewhat small dress'
c. *das klein**e** Kleid* 'the small dress'
d. *dieses klein**e** Kleid* 'this small dress'
e. *ein klein**es** Kleid* 'a small dress'
f. *kein klein**es** Kleid* 'no small dress'
g. *dein klein**es** Kleid* 'your small dress'

Three types of inflection of attributive adjectives are traditionally distinguished in German based on their syntactic surroundings (Zifonun et al., 1997; Eisenberg, 2020):

1. *strong* (Ger. *stark*): an adjective takes over the role of a determiner and the inflection corresponds to that of a definite article (see table 2.1). This holds in the following cases: there is no article preceding the adjective as in example (11a) or the preceding determiner is not inflected as in example (11b);

case	singular			plural
	masc	fem	neut	
Nominative	klein**er**	klein**e**	klein**es**	klein**e**
Genitive	klein**en**	klein**er**	klein**en**	klein**er**
Dative	klein**em**	klein**er**	klein**em**	klein**en**
Accusative	klein**en**	klein**e**	klein**es**	klein**e**

Table 2.1: Strong inflection of German attributive adjectives for the adjective *klein* 'small'. The inflections are in bold.

2. *weak* (Ger. *schwach*): an adjective is inflected either with **-e** or **-en** as illustrated in table 2.2. This holds when an adjective is preceded by a definite article as in (11c), a demonstrative pronoun as in (11d), and a few other cases;

case	singular			plural
	masc	fem	neut	
Nominative	klein**e**	klein**e**	klein**e**	klein**en**
Genitive	klein**en**	klein**en**	klein**en**	klein**en**
Dative	klein**en**	klein**en**	klein**en**	klein**en**
Accusative	klein**en**	klein**e**	klein**e**	klein**en**

TABLE 2.2: Weak inflection of German attributive adjectives for the adjective *klein* 'small'. The inflections are in bold.

3. *mixed* (Ger. *gemischt*): the inflection is shown in table 2.3. This holds when an adjective is preceded by an indefinite article as in (11e), the pronoun *kein* 'none' as in (11f), or a possessive pronoun as in (11g).

case	singular			plural
	masc	fem	neut	
Nominative	klein**er**	klein**e**	klein**es**	klein**en**
Genitive	klein**en**	klein**en**	klein**en**	klein**en**
Dative	klein**en**	klein**en**	klein**en**	klein**en**
Accusative	klein**en**	klein**e**	klein**es**	klein**en**

TABLE 2.3: Mixed inflection of German attributive adjectives for the adjective *klein* 'small'. The inflections are in bold.

The regular comparative form of German adjectives is built by means of the inflection **-er**; the regular superlative form inflection is **-st** as presented in example (12a). The same inflection rules described above apply to those forms when used attributively as illustrated in example (12b).

(12) a. *klein – klein**er** – klein**st***
'small' – 'smaller' – 'smallest'

b. *klein**eres** Kleid – das klein**ere** Kleid*
'smaller dress' – 'the smaller dress'

Lemmatized adjectives in preprocessed corpora do not have any inflections. The convention in this thesis will be to give all examples of German adjective-noun phrases using the strong declension in nominative case singular. This implies that for all the neuter nouns, the modifying attributive adjective will be inflected with **-es**: *kleines* 'small', *teueres* 'expensive', *kluges* 'smart'. For all the masculine nouns, the modifying attributive adjective will be inflected with **-er**: *kleiner*, *teuerer*, *kluger*. Finally, for all the feminine nouns, the modifying attributive adjective will be inflected with **-e**: *kleine*, *teuere*, *kluge*. In the cases where a plural form is needed, also strong declension in nominative case will be used unless specified otherwise: *kleine*, *teuere*, *kluge*.

Participle I and II

There is an ongoing discussion whether participles should also be considered adjectives in German and if so, under which conditions. The general agreement is that participle II in attributive use as in *genähtes Kleid* 'sewn dress' should be considered an adjective (Eisenberg and Schöneich, 2020).

As for participle I, first of all, there is a large number of lexicalized examples: *reizend* 'lovely' derived from the verb *reizen* 'to appeal/excite', *bedeutend* 'important' derived from the verb *bedeuten* 'to mean', *strahlend* 'bright, radiant' derived from the verb *strahlen* 'to radiate/beam'. Such participles are unequivocally adjectives. This is reflected in dictionary entries, for instance, the DWDS provides entries for all the three participles mentioned above. Non-lexicalized participles such as *lesend* 'reading' or *schwimmend* 'swimming' do not have dedicated dictionary entries, and their meaning is derived from the corresponding verb. Irrespective of the level of lexicalization, any participle I used attributively is considered an adjective in German. The verbal origin of participle I is reflected in their ability to take arguments that they inherit from the corresponding verb, for instance, direct objects as in *ein Buch schreibende Person* 'person writing a book' or prepositional phrases as in *im See schwimmende Menschen* 'people swimming in the lake' (Eisenberg and Schöneich, 2020).

Chapter 3

Concept of collocation

This thesis makes a distinction between adjective-noun collocations and free phrases. But what does the term *collocation* stand for? What is the difference to a *free phrase*? Collocations have received a lot of attention in different areas of language research, but the only aspect everyone seems to agree upon is that collocation is a fuzzy concept. This is reflected in a wide and confusing variety of definitions and yet absence of a systematic linguistic study of this phenomenon.

This chapter presents an overview of the existing theories that deal with collocations and addresses the most common questions that are crucial in any discussion concerned with the topic of collocations:

1. Where does the term *collocation* come from? (section 3.1)
2. What is the relation between collocations and multi-word expressions in general? (section 3.2)
3. What is *compositionality*, *decomposability*, and *transparency*? (section 3.3)
4. How to extract collocations from language corpora? (section 3.4)
5. How to define collocations and free phrases? (section 3.5)
6. What collocation resources are available? (section 3.6)

As stated in the introduction chapter (section 1.1.1), there are two approaches to collocations: intensional and distributional. Purely intensional approach is covered in section 3.3. Distributional approach is presented in 3.4. Finally, the insights from both approaches are combined in section 3.5 where working definitions of the notions *collocation* and *free phrase* will be given.

3.1 A short history of collocations

As described in Bartsch (2004), the term *collocation* had already been used in linguistics in the 1930-es. Palmer (1938, pp.x-xi) defined collocations as "*a succession of two or more words that may be best learnt as if it were a single word*." However, the establishment of the term is traditionally ascribed to

the British linguist J. R. Firth. The concept of collocation is thus inseparably associated with the *Firthian linguistics* also referred to as *London School of Linguistics* or *British Contextualism*.[1] Firth introduced the concept of *meaning by collocation* which he illustrated by the phrase *dark night*: "*one of the meanings of night is its collocability with dark, and of dark, of course, collocation with night*"(Firth, 1957, p.196). Thus, a word's meaning is determined by the words with which it co-occurs, i.e. it is defined on the syntagmatic and not only on paradigmatic level. Although Firth's vagueness about the notion of collocation and collocational meaning has been criticised (Lehr, 1996), his ideas are fundamental to the empirical studies of language.

Firth did not equate the concepts *lexical* and *semantic* (Halliday, 1966). This idea was further developed by his students who are commonly referred to as *Neo-Firthians*. One of his students, M.A.K. Halliday, advocated the recognition of a 'lexical level' along with grammar. His notion of *lexicogrammar* implies the interdependence of lexis and grammar, and collocations can be considered the basis for the lexical analysis of language (Pecina, 2008).

Another student of Firth, John Sinclair (1966), also believed that one should look at the internal patterns of language[2] from two aspects: *grammar* and *lexis*. Grammar accounts for those patterns in language that form certain systems within language, whereas lexis is concerned with "*the tendencies of items to collocate with each other*"(Sinclair, 1966, p.411).

Neo-Firthians believed that lexical patterns can be studied based on very large samples of text under consideration of their probabilities (Halliday, 1966) which has become the fundamental idea for corpus-based research.[3] Over the last few decades, a lot of studies emerged concerned with statistical properties of collocations in order to develop methods of automatic extraction of collocations (see section 3.4). Some of those techniques have been successfully used for compiling collocation dictionaries.

Collocations have always received a lot of attention from lexicographers who are mainly concerned with two questions: which collocations should be included in dictionaries and what is the best way of representing collocations in a dictionary? Specialized collocation dictionaries are sometimes called *combinatory* or *combinatorial* dictionaries. Earlier learners' dictionaries included a lot of free combinations of words such as *neues Haus/Auto/Hemd* 'new house/car/shirt'. Hausmann (1985), who worked mostly with German and French languages and was mainly interested in compiling a practical learner's dictionary, criticized such an approach. He considered such combinations unnecessary and redundant because they are not specific and not typical: they do not pose any problems to a language learner. Hausmann's

[1] The term "British Contextualism"has been mostly used by German scholars (German: *Britischer Kontextualismus*) (Lehr, 1996).

[2] Sinclair (1966) recognizes external patterns that may have an influence on language form as well, for instance, social patterns or natural organization, but does not take them into account.

[3] For an in-depth and detailed description of Firthian and Neo-Firthian linguistics see Lehr (1996)

principles of compiling combinatory dictionaries were applied in the BBI Combinatory Dictionary of English (Benson, 1989).

Another type of a collocation dictionary, Explanatory Combinatorial Dictionary (ECD), was proposed by Mel'čuk and Zholkovsky (1984) and their colleagues. The ECD's main function is to store information about the semantic and combinatorial relations[4] of individual words in a formalised manner which make this type of dictionary theory-oriented. The ECD is rather meant for linguists than for language learners and is meant to connect lexicography and theoretical linguistics. Every entry in the ECD can be considered a "*self-contained linguistic article on the corresponding < ... > lexeme*" (Mel'čuk and Zholkovsky, 1984, pp.39, 43). A considerable part of such an article is dedicated to combinatorial preferences of each lexeme: the collocates are listed and grouped according to the semantic relations that hold between them and the given lexeme. The framework for describing the semantic relations in collocations is based on the notion of a *Lexical Function* which will be discussed in detail in section 4.2.1. At the present moment, the inventory of Lexical Functions is the most comprehensive and systematic framework describing the semantic properties of collocations.

3.2 Collocations and Multi-Word Expressions

As already mentioned in the beginning of this chapter, the terminology concerned with collocations and other multi-word expressions is not homogeneous. This is reflected in the multitude of terms used to refer to multi-word expressions: *set expressions, fixed expressions and idioms, lexical combinations, phrasemes, fixed combinations, multiword lexical units*. There is also no general agreement on the relation between MWEs and collocations. The approaches to this issue can be roughly classified into three categories that will be presented in more detail in the following subsections:

1. Collocations are a more general concept than MWEs (section 3.2.1)
2. Collocations are a subclass of MWEs (section 3.2.2)
3. Collocations and MWE are two disjoint concepts (section 3.2.3)

In each subsection, the typologies of MWEs are presented in the chronological order with the focus on collocations. All the typologies rely on three factors for distinguishing between different types of MWEs: flexibility of form, compositionality of meaning, and statistical idiosyncrasy which will be discussed at the end of the section (section 3.2.4).

3.2.1 MWEs as a subclass of collocations

Kjellmer (1991): set expressions

[4]Both paradigmatic and syntagmatic ones.

Strictly speaking, in this classification, set expressions are not considered to be a subgroup of collocations but are rather used synonymously to collocations. For Kjellmer (1991), set expressions (or ready-made phrases) are *"combinations of words that customarily co-occur"* that account for a large portion of our mental lexicon. Collocations are defined as *"structured patterns that recur in identical form"* (Kjellmer, 1991) and are thus equated with *ready-made phrases*. Similarly to many other scholars, Kjellmer emphasizes that different classes of set expressions may sometimes overlap and thus the distinction between them is unclear. He distinguishes three groups of ready-made expressions:

1. **Fossilized phrases**: combinations in which one constituent predicts the other one. Such combinations allow very little variation, e.g. inflection of the predicted word. Examples of such phrases are *aurora borealis, bonsai tree, ad infinitum*;

2. **Semi-fossilized phrases**: one constituent predicts a limited number of words. This category includes most idioms such as *to go off the right/wrong foot* and phrases like *moot point/question, nodding acquaintance/terms*. The constraints on the choice of the component words in such phrases are lexical rather than functional;[5]

3. **Variable phrases**: their constituents tend to co-occur frequently, however, the prediction in such case should be interpreted more loosely. There are two kinds of variable phrases:

   – phrases consisting of two or more lexical words and some function words: *close friend, glass of water, loud and clear*;

   – established sequences of one lexical word and one or more function words with somewhat restrained variability: *a number of, for a change, out of action*.

Moon (1998): FEIs

Moon (1998) adopts the term **FEIs** (fixed expressions and idioms) to refer to holistic units consisting of two or more words. Moon (1998) names three factors based on which an expression is considered FEI:

– **institutionalization**: based on statistical co-occurrence. It is a "process by which a string or formulation becomes recognized and accepted as a lexical item"(Moon, 1998, 7).

– **lexicogrammatical fixedness**: quality of having some lexicogrammatical idiosyncrasy.

[5]For instance, the adjective *blue* can select any word with certain functional properties such as concrete, noun which is not enough for an expression to be considered semi-fossilized as this criterion is very general and allows a lot of options (basically, all concrete nouns).

- **semantic non-compositionality**: the meaning of the string does not arise from the meanings of its constituents, such expressions are typically considered metaphorical.

The term **collocation** is used to denote any co-occurrence of items and is thus more general than the term FEI. There are three subtypes of FEIs covered in Moon (1998):[6]

1) Anomalous collocations are syntagmatically or paradigmatically deviant.

- **Ill-formed collocations** are grammatically idiosyncratic: *by and large, of course, at all*
- **Cranberry collocations** contain unique elements not found in other collocations: *to and fro, kith and kin*
- **Defective collocations** are not fully compositional because one constituent has a shifted meaning not found in other contexts: *to beg the question, a foregone conclusion*
- **Phraseological collocations** are those that have alternative variants, but "*the structure is not productive*"(Moon, 1998, p.21): *to a – degree* and *to a – extent*, *on show* and *on display*

2) Formulae are "specialized pragmatically", most of them are compositional (Moon, 1998, p.21)

- **Simple formulae** are syntagmatically fixed expressions with a discoursal function: *I'm sorry to say, not exactly*
- **Sayings** include quotations, catch-phrases, and truisms: *an eye for an eye, that's the way the cookie crumbles*
- **Proverbs** are "*maxims with deontic function*"(Moon, 1998, p.22): *enough is enough, every cloud has a silver lining*
- **Similes** are opaque institutionalized comparisons: *as old as the hills, live like a king*

3) Metaphors are non-compositional phrases with different degrees of transparency

- **Transparent metaphors** can be decoded relying on world knowledge: *behind someone's back, pack one's bags*
- **Semi-transparent metaphors** can be decoded, but this requires more specialised knowledge; they may be interpreted in various ways: *throw in the towel, grasp the nettle*
- **Opaque metaphors** are almost impossible to decode compositionally: *shoot the breeze, red herring*

[6] FEIs not covered in Moon (1998) are compounds, phrasal verbs, foreign phrases and multi-word inflectional forms of verbs, adjectives, and adverbs (Moon, 1998, p.2).

Sag et al. (2002): multi-word expressions

Sag et al. (2002, p.2) give a very general definition to multi-word expressions: "*idiosyncratic interpretations that cross word boundaries (or spaces)*". The notion of collocation is even more general and is used to refer to any statistical co-occurrence and thus includes not only MWEs, but also compositional phrases with high frequency that do not exhibit any semantic or lexical idiosyncrasies. They distinguish between two classes of MWEs: **lexicalized** and **institutionalized** phrases:

1) Lexicalized phrases are (at least partially) semantically or syntactically idiosyncratic:

- **fixed expressions**: grammatically idiosyncratic expressions such as *by and large, in short*. They are fully lexicalized and do not allow any modifications.
- **semi-fixed expressions** allow some modification, e.g. inflection. They include non-decomposable idioms such as *shoot the breeze*, compound nominals (*part of speech*), proper names (*the San Francisco 49ers*).
- **syntactically-flexible expressions**: syntactically more variable. They comprise verb-particle constructions (*to look up*), decomposable idioms (*spill the beans*), light verbs (*to give a speech*).

2) Institutionalized phrases are fully compositional, but they exhibit statistical idiosyncrasy. Thus, in the compound *traffic light*, both constituents are semantically transparent. However, this concept cannot be expressed by any other compositional combination of words because *traffic light* has been strongly institutionalized.

3.2.2 Collocations as a subclass of MWEs

Benson et al. (1986) and van der Wouden (1997): lexical combinations

Benson et al. (1986, p.252) define lexical combinations as expressions that "*exhibit varying degrees of cohesiveness*"; collocations form a subgroup in their classification. The same typology of lexical combinations is also used in van der Wouden (1997).

1) Free combinations exhibit the lowest degree of cohesiveness, their constituents are free to combine freely with a considerable number of words: e.g. *to examine/forget/report/romanticize a murder*

2) Idioms are fixed non-compositional expressions that sometimes allow variability, especially grammatical one: *we'll kill (or: we killed) two birds with one stone*

3) Collocations are "*loosely fixed combinations*" (Benson et al., 1986, p.253) between idioms and free combinations. They exhibit the following characteristics: (1) collocations are compositional, but restricted: their constituents cannot be easily substituted by synonyms, and (2) collocations are frequent

and thus "*psychologically salient*" (Benson et al., 1986, p.253). This group includes adjective-noun collocations such as *pure chance, grave concern* and two types of verb-noun collocations: (1) CA (a verb denotes creation/activation): *make an impression* and (2) EN (a verb denotes eradication/nullification): *to dispel fear, to lift a blockade*

4) Transitional combinations are in the grey area between idioms and collocations, they are more fixed than collocations, but their meaning is not entirely opaque: *to catch one's breath, to be in a tight spot*

5) Proverbs are more frozen than idioms, but their meaning can be literal. They convey folk wisdom or general truths expressed by a complete sentence: *An apple a day keeps the doctor away.*

6) Compounds are completely frozen with no variations, they include adjective-noun combinations such as *floppy disk, definite article*, noun-noun combinations such as *hand luggage, fire escape*, and phrasal verbs such as *add up, break through*.

Baldwin and Kim (2010): multi-word expressions

Baldwin and Kim (2010) define MWEs as "lexical items that: (a) can be decomposed into multiple lexemes; and (b) display lexical, syntactic, semantic, pragmatic and/or statistical idiomaticity" The taxonomy of MWEs presented in Baldwin and Kim (2010) follows the one described in Sag et al. (2002). However, their view on collocations is different: collocations are considered a subclass of MWEs. At the same time, collocations as such are not mentioned in the taxonomy of MWEs: the term **simple collocation** is used as a synonym for **institutionalised phrases** – statistically idiosyncratic phrases. Moreover, Baldwin and Kim (2010) emphasize that there should exist a direct syntactic relation between the constituents of collocations: for instance, adjective-noun or verb-noun collocations.

Mel'čuk (2012a): phrasemes

Following the Russian linguistic tradition, Mel'čuk refers to MWEs as "*phraseological expressions*"or "*phrasemes*" According to Mel'čuk (2012a, p.32), phrasemes are "*linguistic expressions formed by several (at least two) lexemes syntactically linked in a regular way <...> featuring some unpredictable properties*" and thus they are linguistically constrained as opposed to free phrases. Mel'čuk's taxonomy of phrasemes is based on the idea of freedom of selection: in the process of speech production, the speaker is faced with a choice of lexical items. If the freedom is violated in some way, the resulting phrase is lexically constrained and is considered a phraseme. In his taxonomy, there are two major classes: **lexical phrasemes** and **semantic-lexical phrasemes** (alternatively, they are referred to as semantic and pragmatic phrasemes respectively (Mel'čuk, 1998)):

1) Lexical phrasemes are those, in which the choice of the starting meaning is free, but the choice of some of the lexical components is constrained by

the language. There are two subgroups of lexical phrasemes:

- **Idioms** are non-compositional. They include full idioms, such as *by heart*; semi-idioms (*private eye*), and quasi-idioms or weak idioms (*to start a family*)
- **Collocations** are compositional. They include standard collocations, such as *to give an order, deeply committed*, and non-standard collocations such as *black coffee, leap year* (see a more detailed review of Mel'čuk's theory in section 2.1)

2) Semantic-lexical phrasemes are those in which both the meaning and the lexical components are constrained by the language depending on the situation.

- **Clichés** are compositional. They include pragmatically constrained clichés (pragmatemes) such as *Return to sender*, and pragmatically non-constrained (normal) clichés such as proverbs (*a watched pot never boils*) and complex proper names (*The Old Testament, The Moonlight Sonata*).

3.2.3 Collocations and MWEs are disjoint

Hausmann (1984): fixed combinations[7]

Hausmann (1984) emphasizes that the preference of certain word to combine with each other cannot not be explained purely semantically. Based on words' combinability, the following typology of word combinations is proposed where collocations form a separate from fixed expressions subclass:

1) Fixed combinations include compounds and idioms

2) Non-fixed combinations include the following subclasses:

- **Free combinations** (*co-creations*): are composed creatively by a speaker according to the rules of a language, there is an unlimited number of possibilities how words can be combine
- **Collocations** (*collocations*): their constituents have restricted combinability and an affinity to combine with each other. They are combined according to semantic rules, but not composed creatively. Collocations can be considered "semi-finished products"of a language.
- **Words with restricted combinability** are combined with words to which they have a negative affinity (*counter-creations*): they are composed creatively, but break semantic rules. Such combinations are often encountered in literary texts

[7] Ger.: Fixierte Wortverbindungen/Wortkombinationen

Savary et al. (2018): multi-word expressions

Savary et al. (2018, p.92) define MWEs as "*sequences of words which contain at least two component words which are lexicalised <...> [and which] display some degree of lexical, morphological, syntactic and/or semantic idiosyncrasy*". The idiosyncrasy in collocations is, on the other hand, of pragmatic or statistical nature only. Examples of collocations are such phrases as *the graphic shows, drastically drop, all aboard*. Thus, as statistical and pragmatic idiosyncrasies are not included in the definition of MWEs, the authors conclude that "*MWEs and collocations are seen as disjoint sets of linguistic objects*"(Savary et al., 2018, p.93).

Kosem et al. (2020): multiword lexical units

In this classification, the concept of the Multiword lexical units (MLU) is used synonymously to the multi-word expression. By using the word *unit* instead of *expression*, Kosem et al. (2020) emphasize that MLUs function as holistic units that are semantically independent. Collocations are distinguished both from free combinations and MLUs:

1) free combinations: have high frequencies, but are less informative than collocations and therefore are lexicographically irrelevant (*nisem vedel* 'I didn't know')

2) collocations have the following characteristics: statistical relevance, syntactic structure, semantic transparency and meaningfulness (*najbližji sorodniki* 'closest relatives')

3) multiword lexical unit: holistic units that belong in a dictionary. Their meaning is at least to some degree idiomatic.

- **compounds**: the meaning of the whole cannot be derived from the meanings of its parts. However, the meaning of a compound is not metaphorical, but rather technical, e.g. *topla greda* 'greenhouse' or 'green-house effect', *črna luknja* 'black hole'.
- **phraseological units**: they are similar to compounds, but have a metaphorical meaning and are expressive. This category includes proverbs and sayings (*time is money*), expressions with pragmatic or evaluative role (*damn*), expressions with adverbial role (*more or less*), and expressions with communicative role (*good evening*).
- **lexico-grammatical units** have grammatical and function words as their constituents and their role is to organize text. They include:
  - phrasal verbs
  - light-verb constructions
  - reflexive verbs
  - syntactic combinations

3.2.4 Summary

The above presented typologies illustrate that, in spite of all the attempts to define collocations and multi-word expressions, the decision about the status of a complex expression is often subjective and depends on the framework adopted for a particular task. The issues of subjectivity and the overall 'fuzziness' of terminology and methodology surrounding multi-word expressions have not been solved yet.

Statistical idiosyncrasy. There is no agreement on whether MWEs are statistically idiosyncratic.

- All MWEs exhibit statistical idiosyncrasy (Kjellmer, 1991; Moon, 1998; Sag et al., 2002).
- MWEs have high frequency (Benson et al., 1986; van der Wouden, 1997).
- MWEs may exhibit statistical idiosyncrasy (Baldwin and Kim, 2010; Kosem et al., 2020).
- Statistical properties are not mentioned (Hausmann, 1984; Mel'čuk, 2012a).
- Clear distinction between statistically idiosyncratic phrases (*collocations*) and MWEs that are only lexically, semantically, or syntactically idiosyncratic (Savary et al., 2018).

Resolving the complex issue of terminology is outside of the scope of this thesis. The typologies presented in this section are meant to be used a reference material for creating a database of collocations that will serve as an empirical basis for the study of semantic properties of collocations in this thesis. Although the amount of views may seem overwhelming and confusing, there are common themes in various frameworks that can be identified. Such commonalities touch upon both the form and the meaning of MWEs.

Form. MWEs exhibit various degrees of morphosyntactic flexibility (the terminology follows the one suggested by Sag et al. (2002)):

- **Fixed**: grammatically idiosyncratic expressions that do not allow any modifications: *by and large, in short, in line* (Moon, 1998; Sag et al., 2002; Baldwin and Kim, 2010).
- **Semi-fixed**: allow some modification. Idioms are often considered to be semi-flexible (Sag et al., 2002; Benson et al., 1986): *He spill**ed** the beans*.
- **Flexible**: MWEs that can undergo a wide range of syntactic changes. This group includes verb-particle constructions, e.g. *to look up*; light verb constructions, as in *to make a mistake* (Sag et al., 2002); adjective-noun collocations, as in *grave mistake*.

Meaning. MWEs exhibit various degrees of semantic compositionality:

- **Compositional**: collocations (Mel'čuk, 2012a; Benson et al., 1986; van der Wouden, 1997) and proverbs (Benson et al., 1986; van der Wouden, 1997; Moon, 1998; Mel'čuk, 2012a).
- **Non-compositional**: idioms (Benson et al., 1986; van der Wouden, 1997; Moon, 1998; Mel'čuk, 2012a).

Flexibility of form does not seem to cause any misunderstanding among linguistic scholars, whereas the properties referring to meaning, such as compositionality and transparency, vary extremely throughout literature. The goal of the next section is to discuss semantic properties of MWEs in more detail and provide a working definition for each of them.

3.3 Overview: Compositionality, transparency, decomposability

There are three terms that inevitably make their way into any discussion about complex word combinations: *transparency*, *compositionality*, and more rarely *decomposability*. These terms are frequently encountered in literature and are sometimes used interchangeably which may lead to confusion. In order to be able to navigate among different terms and concepts in the realm of multi-word expressions, it is crucial to define these notions.

3.3.1 Transparency

The term *semantic transparency* is frequently encountered in works on morphology, especially in the context of compounds, and in research on idioms.

For compounds, there are two approaches to defining semantic transparency in literature: *predictability* and *analysability* (Bell and Schäfer, 2013). The first approach to defining semantic transparency is by linking it to **meaning predictability**. The main idea of this approach is that if it is possible to predict the meaning of a word from the meaning of its constituents based on the rules of word-formation then the word is semantically transparent. Bell and Schäfer (2013) argue that, in this case, to consider a compound *milkman* transparent, it would be necessary to have a word-formation rule "*x who delivers y to people's houses*."

The second approach is based on **analysability**, where the main criterion is whether the meaning of the compound is related to the meaning of its constituents. Zwitserlood (1994) classifies compounds into three groups based on transparency: *fully transparent* – when the relation between the compound as a whole and its constituents is clear as in *milkman*; *partially opaque* – when the meaning of one constituent word is related to the meaning of the compound as in *jailbird*; and *fully opaque* – when the meaning of the constituent words is not related to the meaning of the compound as in *blackguard*. A similar definition of transparency is adopted in Libben

et al. (2003), also in context of compounds. Depending on the level of semantic transparency, a compound can either be interpreted in terms of its constituents or it has to be stored in the lexicon as a holistic unit. Thus, the meaning of the semantically transparent compound *car-wash* is combined from the meanings of *car* and *wash*. In contrast, the compound *hogwash* (= *nonsense*) is semantically opaque because its meaning cannot be derived from the meanings of its constituents and thus the word has to be stored in the lexicon with its idiosyncratic meaning (Libben et al., 2003).

In research on idioms, semantic transparency is defined as a direct and traceable relation between the literal and figurative meaning of an idiom. For instance, the idiom *to pull strings* is considered transparent because there is a clear analogy to controlling a puppet by pulling its strings (Sailer, 2021). An example of an *opaque* idiom is *to kick the bucket* (fig. 'to die') in which the literal meanings of its components have no relationship with the figurative meaning of the idiom. Thus, a person who does not know the meaning of this idiom cannot guess it.

Semantic transparency is commonly considered *gradable*, i.e. there is a continuum from fully opaque to fully transparent linguistic units. This feature is normally measured in terms of a 1 to 5 or 1 to 7 Likert scale (Pagliai, 2023).

3.3.2 Compositionality and decomposability

The Principle of Compositionality is also known as *Frege's principle*. It implies that "*the meaning of a complex expression is a function of the meaning of its constituent parts and of the mode of their combination*"(Baroni et al., 2014, p.242). In formal semantics, the principle suggests that the meaning of a sentence is constituted from the meanings of its components. It allows the speakers to combine simpler meanings to compose more complex ones and vice versa – decompose complex sentences into simple meanings based on a finite lexicon (Baroni et al., 2014).

Compositionality is not limited to studying language on a sentence level, the concept is also frequently encountered in works on lexical semantics. For instance in Mel'čuk's (2012a) framework, the notion of compositionality is essential for distinguishing between different types of complex word combinations or, in Mel'čuk's terminology, *complex linguistic signs*.

Definition 1
A linguistic sign is a triplet comprising *the signified* (= linguistic meaning), *the signifier* (= phonemes or characters), and *the syntactics* (= information about the sign necessary for combining it with other signs) (Mel'čuk, 2012a).

Simple linguistic signs can be combined into complex ones, and an important property of complex signs is *compositionality* defined in the following way (Mel'čuk, 2012a, p.35):

Definition 2
A complex linguistic sign **AB** is compositional iff **AB = A⊕B** .

The symbol ⊕ stands for *the Operation of Linguistic Union* that combines simple linguistic signs (the signified, the signifier, and the syntactics) into those of higher order according to the specific rules of a given language. As Mel'čuk (1998, p.27) notes, this operation is more complex than simple addition because "*it presupposes observing all the general combination rules [of a language], and doing this in conformity with the nature of items being united*." The operation of linguistic union A⊕B thus implies an automatic union of the items A and B according to the general combination rules of a language considering the properties of A and B (Mel'čuk, 2012b). For instance, the regular way of forming plural of nouns in English is adding the affix **-s**: *cat – cats*, *book – books*, *willow – willows*.

A different approach is presented in Reddy et al. (2011) who study compositionality of English compounds. Bell and Schäfer (2013) note that the definition of compositionality adopted in that study actually corresponds to the notion of semantic transparency based on predictability (although the term *transparency* is never mentioned in the paper): "*A compound is compositional if its meaning can be understood from the literal (simplex) meaning of its parts*" (Reddy et al., 2011, p.211). In the annotation set up of Reddy et al. (2011), compositionality is thus equated with *literality*. This approach is later adopted by Bell and Schäfer (2013). A similar view on this phenomenon is advocated in Manning and Schütze (1999); Schone and Jurafsky (2001); McCarthy et al. (2007) where literality is not always mentioned explicitly, but is implied:

- Manning and Schütze (1999, p. 111): "*<...> does not obey the principle of compositionality by which the meaning of the whole can be strictly predicted from the meaning of the parts.*"
- Schone and Jurafsky (2001, p. 1): "*As an example of semantic non-compositionality, consider "compact disk": one could not deduce that it was a music medium by only considering the semantics of "compact" and "disk."*"
- McCarthy et al. (2007, p. 369): "*it is essential to take the words of non-compositional phrases together as a unit because the meaning of a phrase cannot be obtained straightforwardly from the constituent words.*"

Unlike transparency, compositionality gives rise to disagreements with respect to its gradability. The results of the annotation in Reddy et al. (2011) show that there is a *continuum of compositionality*. This finding goes in line with the common assumption that compositionality is a continuum or a scale from fully non-compositional idioms to compositional (Moon, 1998; Bartsch, 2004; McCarthy et al., 2007; Biemann and Giesbrecht, 2011; Pagliai, 2023). However, there also exists an opposite view on that issue. Mel'čuk (2023, p.32) emphasizes that partial compositionality does not exist and "compositionality does not admit degrees". It is a discrete notion and all expressions in a language fall under two categories: compositional or non-compositional, and there is no in between.

Finally, research on idioms introduces a third term into the discussion, namely, *decomposability*. Decomposable idioms are idioms in which the

meaning is composed of *"idiomatic meanings of its parts"* (Nunberg et al., 1994, p. 507). In such phrases, the meanings of the constituents are figurative, but they can be mapped onto parts of the figurative meaning of an idiom. For instance, in the German idiom *bitteres Erwachen* 'bitter awakening' which means 'unpleasant realisation', *bitter* is mapped to *unpleasant* and *awakening* is mapped to *realisation*. Both constituents add to the meaning of the whole phrase, therefore, this idiom is decomposable. In contrast, the idiom *trip the light fantastic* that means "dance lightly" is non-decomposable since its figurative meaning cannot be split between its elements "trip", "the light", and "fantastic".

3.3.3 Conclusion

This section discussed three variables that characterise semantic properties of complex linguistic signs: transparency, compositionality, decomposability. There is no general consensus on how to define these terms. They are frequently used interchangeably and are operationalized in different ways.

The discussion in the subsections above relied to a large extent on the notion of *literal meaning* and this thesis will also make use of this term. However, note that apart from *literal* meaning, in literature on lexical semantics and corpus linguistics, the terms *core* meaning (Sinclair, 1991), *salient* meaning (Philip, 2011), *basic* meaning (Pragglejaz Group, 2007), *normal* meaning (Mel'čuk, 2023) are encountered.

Core meaning is the meaning of a word that first comes to mind, without any context given. Sinclair (1991, p.113) hypothesizes that the core meaning of a word is "*the most frequent independent sense of a word*". Philip (2011) refers to that phenomenon as *salience* and defines *salient meaning* of a word or phrase as the most dominant one for a particular person. Thus, salience is a subjective notion and it is not possible to empirically establish the salient meaning of a word using quantitative methods. According to Philip (2011), salient meaning should not be equated with the literal meaning, the insight backed up by the findings in psycholinguistics that idioms have a salient meaning as a whole. There is evidence that in the phrases allowing both a literal and a non-literal reading, the latter is actually more salient than the literal one (Philip, 2011).

A similar view is advocated by Pragglejaz Group (2007) where instead of *literal*, the the notion of *basic* meaning is used. Basic meaning is not the most frequent meaning of a word. In Strakatova et al. (2020) a basic meaning of a word is referred to by a German term *Grundbedeutung* which can be roughly translated as *basic, primary, core*. Mel'čuk (2023, p. 114) uses the expression *normal meaning* that stands for "context-free".

This thesis will refrain from using the terms *compositionality*, *transparency*, and *decomposability*. Instead, it will make use of the term *literal meaning* which is necessary for distinguishing between different types of MWEs and free phrases. Literal meaning will not be equated with salient: compare the definitions 3 and 4.

Definition 3
Literal meaning is the most basic meaning of a word (German: *Grundbedeutung*). Following the definition by Pragglejaz Group (2007, p.3): the most concrete, imaginable, historically older meaning, it is precise and is related to a bodily action.

Definition 4
Salient meaning is the most frequent meaning of a word.

This section covered the semantic criteria applied to MWEs in general and to adjective-noun MWEs in particular. Those criteria reflect the phraseological approach to MWEs which is only one half of the whole story. As stated in sections 1.1.1 and 3.2, another crucial characteristic of MWEs, and collocations in particular, is recurrence that is traditionally operationalized as *statistical idiosyncrasy* that will be discussed in the next section. Only after discussion of statistical idiosyncrasy, formal definitions of the key concepts in this thesis will be given.

3.4 Collocation extraction methods

There is a growing body of research concerned with developing effective and suitable statistical measures based on the corpus frequencies of the collocations and their constituents. Such measures usually test whether the occurrence of a word pair is statistically significant. This section gives a short overview of the research on automatic collocation extraction and presents several association measures commonly used for this task. Section 3.4.1 introduces the main concepts and presents the traditional association measures. Section 3.4.2 addresses the topic of asymmetry in collocations and its statistical interpretation.

3.4.1 Statistical association measures

The traditional approach to measuring association strength between any two units (in this case, an adjective used attributively and its head noun) is based on frequencies of co-occurrences that are commonly represented in the form of a **contingency table** (table 3.1) (Manning and Schütze, 1999; Evert, 2004; Pecina, 2008).

	word2	**¬word2**	**total**
word1	a	b	a+b
¬word1	c	d	c+d
total	a+c	b+d	a+b+c+d=n

TABLE 3.1: Contingency table of observed frequencies. *a* is the frequency of a bigram, *a+c* is the frequency of the noun and *a+b* is the frequency of the adjective

Table 3.2 gives an example of such a contingency table for the phrase *starker Kaffee* 'strong coffee'.[8] The table is interpreted in the following way:

- the frequency of the phrase *starker Kaffee* is 478 (a);
- the overall frequency of *stark* as a modifier is 380,302 (a+b);
- the overall frequency of *Kaffee* used with an adjectival modifier is 15,750 (a+c);
- the overall amount of all AN bigrams is 194,715,517 (a+b+c+d).

	word2 *Kaffee*	**other nouns**	**total**
word1 *stark*	478	379,824	380,302
other adjectives	15,272	194,319,943	194,335,215
total	15,750	194,699,767	194,715,517

TABLE 3.2: Contingency table of observed frequencies for the bigram *starker Kaffee* 'strong coffee'

The idea that simple raw frequencies are not enough for the task and a statistical interpretation is required comes from the intuition that, first of all, co-occurrence might be just a coincidence. Apart from that, the numbers for the raw frequencies stem from a particular corpus, but the aim is to be able to make generalisations about a language as whole (Evert, 2004). Traditionally, collocation extraction methods are mathematically expressed as **lexical association measures** (AMs) that indicate the degree of association between the constituents of collocations (Pecina, 2008). Such association scores are computed for word pairs and, as a result, a list of ranked pairs is created. The higher the association score is, the more likely it is that the given word pair is a collocation.

A typical pipeline for extracting collocations using AMs is the following (Evert and Kermes, 2003; Evert, 2004; Pecina, 2008):

1. **Step 1**: pre-processing the corpus (lemmatization, POS-tagging, syntactic annotations).
2. **Step 2**: extracting a list of candidate pairs with their frequencies. The candidates are chosen according to the criteria depending on the task (usually minimal frequency and certain syntactic patterns such as verb-object constructions or adjective-noun phrases). Further filtering of the list if necessary.
3. **Step 3**: computing association measure scores for the word pairs.

[8] Frequencies were calculated based on a large corpus consisting of the treebanks Wikipedia 2017 and Wikipedia 2018 (de Kok and Pütz, 2019) and decow16ax (Schäfer and Bildhauer, 2012; Schäfer, 2015)

4. **Step 4**: ranking the candidates or classifying the candidates as collocations and non-collocations. In the second case, there are two options: applying a cut-off threshold or creating **n-best list** by accepting the n highest ranking candidates (Evert, 2004).

In most cases, the task is regarded as a binary classification task, where the two discrete categories are **collocation** and **non-collocation**. The extracted candidates (n-best lists) are given to human annotators who further classify the word pairs following certain guidelines.

Lexical association measures are borrowed from different fields such as mathematical statistics, information theory, information retrieval. Over the last few decades, various measures have been proposed, modified, discovered and re-discovered. According to Evert (2004, pp.76-77), there are four major groups of AMs:

1. **significance of association group** (likelihood measures, exact hypothesis tests, asymptotic hypothesis tests);
2. **degree of association group** (MI, Dice coefficient, Jaccard coefficient, odds ratio);
3. **information-theory group** (MI, PMI);
4. **heuristic formulae** (co-occurrence frequency, MI^2, MI^3).

Numerous studies have shown that there is no such thing as one universal measure that works best for extracting collocations in any task from any data. As the extraction principles behind different AMs vary, there are different patterns in their behaviour when applied to the task. For instance, PMI (pointwise mutual information) measure is known for its *low-frequency bias*: it tends to give higher scores to low-frequency events (Bouma, 2009). This is not necessarily a disadvantage as it can help to identify rare collocations. The performance of individual AMs depends on different aspects, therefore, they should be evaluated on a specific task and data. Evert (2004) conducts a comparative evaluation study of several AMs in the task of collocation identification. The results show that t-score and simple co-occurrence frequency work best on the preposition-noun-verb collocations dataset (Krenn, 2000) followed by log-likelihood and chi-squared. However, applying the same AMs to a different data sample, this time adjective-noun collocations, yields surprising results: the best performance is achieved by MS measure and Dice coefficient. The insight that different dependency patterns behave differently regarding various measures is also confirmed by Garcia et al. (2019). Pecina (2008) conducted a comprehensive study of over 80 association measures and their applicability to the task of automatic collocation extraction. Apart from testing individual measures, Pecina conducted a few experiments to show that a combination of several AMs can work even better than each of them separately.

The results of the automatic collocation extraction can be affected not only by the choice of association measures, but also by several other factors:

- **Corpus size**: the results of the experiments in the dissertation by Pecina (2008) indicate that a larger corpus has a positive effect. The same observation is reported in Evert et al. (2017).
- **Corpus quality**: results for clean and balanced corpora are better (Evert et al., 2017).
- **Corpus domain**: the intuition suggests that depending on the corpus domain, different kinds of collocations will be extracted, e.g. certain collocations are characteristic of newspaper style. Moreover, Evert et al. (2017) show that the scores gained based on movie subtitles perform better than those based on the British National Corpus (BNC), presumably because the language there is closer to spoken language.

The above described factors make it difficult to create a general universal tool for automatic collocation extraction that would be suitable for any corpus size and word frequency. An additional problem is concerned with the interpretation of the most AM scores: as mentioned previously, the most common procedure is sorting the word pairs by their scores and taking the n-best lists. Each score by itself, however, is not very informative because of the lack of a fixed upper bound (Bouma, 2009), so it is not possible to compare the scores computed from different corpora. The solution is proposed by Rychly (2008). It is based on the Dice coefficient measure eq. (3.1) that comes from information retrieval (based on the notation in table 3.1):

$$D = \frac{2a}{(a+c)+(a+b)} \tag{3.1}$$

Rychly (2008) notes that the main problem of the Dice coefficient is that its values are usually very small numbers. The issue is solved in the following way:

$$logDice = 14 + log_2 D = 14 + log_2 \frac{2a}{(a+c)+(a+b)} \tag{3.2}$$

The new logDice measure has the following features (Rychly, 2008):

- its theoretical upper bound is 14, most values are <10
- a negative value means there is no statistical significance of the co-occurrence
- the score is independent of the corpus size

Thus, for the phrase *starker Kaffee* 'strong coffee' (table 3.2), the logDice score of 5.3 is computed:

$$logDice(starker\ Kaffee) = 14 + log_2 \frac{2 * 478}{15,570 + 380,302} \approx 5.3 \tag{3.3}$$

The idea to use corpora frequencies for describing words' grammatical and collocational behaviour is at the basis of the so-called **word sketches** (Kilgarriff et al., 2004). A word sketch relies on grammatical relations rather than just looking at n words surrounding the head word. The list of collocates is thus provided for each relation a word participates in. Kilgarriff et al. (2004) developed a tool that generates such word sketches given a corpus in any language – Sketch Engine[9]. Sketch Engine *"is a corpus query system which allows the user to view word sketches, thesaurally similar words, and 'sketch differences' < ... >"* (Kilgarriff et al., 2004, p.4). Generation of collocate lists in the Sketch Engine is based on the logDice measure (eq. (3.2). For instance, for the noun *coffee*, the following lists of collocates are generated in Sketch Engine: modifiers of "coffee", nouns and verbs modified by "coffee", verbs with "coffee"as object, verbs with "coffee"as subject, "coffee"and/or ..., prepositional phrases, adjective predicates of "coffee", etc. A similar tool has been developed for German as well – the DWDS[10] word profile system the Wortprofil[11] (Geyken et al., 2008). The lists of statistically salient co-occurrences are generated based on the logDice-scores computed for certain grammatical relations.

Association measures such as logDice and PMI are measures of *mutual* attraction between words. However, the nature of idiosyncratic word combinations is very often asymmetric, i.e. one constituent is more predictive of the other. This quality of multi-word expressions is discussed in the next section.

3.4.2 Asymmetry in collocations

Most of the previous studies on statistical properties of collocations focus on bidirectional measures and do not take into consideration the fact that one of the constituents can be more predictive of its partner. This issue has been addressed in a few studies and has been getting more attention recently.

Early observations on this topic are found in the works of psychologists and cognitive scientists studying similarity. Amos Tversky (1977) argued that similarity is not a symmetric relation: the fact that a resembles b does not entail that b likewise resembles a (Tversky, 1977). In this case, a is the subject or *variant* and b is the referent or *prototype*, and according to Tversky, the referent is more *salient* or *prototypical* than the subject. He conducted several experiments to compare human judgements of similarity between countries, figures, letters, and signals, and the results support his theory that the subject is more similar to the referent than vice versa, e.g. North Korea is more similar to China than China to North Korea or an ellipse is more similar to a circle than vice versa (Tversky, 1977).

[9] https://www.sketchengine.eu/, *last accessed on 16.03.2021*

[10] Digitales Wörterbuch der deutschen Sprache. Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart, hrsg. v. d. Berlin-Brandenburgischen Akademie der Wissenschaften, <https://www.dwds.de/>, *last accessed on 30.10.2020.*

[11] https://www.dwds.de/wp, *last accessed on 30.10.2020*

With regard to collocations, Sinclair (1966) noticed that the significance of collocational constituents is not equal. He illustrates this idea with the example *a good omen*: for *omen*, its co-occurrence with *good* is more significant than for *good* the co-occurrence with *omen*. This intuition is based on the fact that the adjective *good* has a very high frequency and a great number of co-occurring nouns modified by it, so the role of *omen* as one of the many nouns is not that significant.

Kjellmer (1991) distinguishes three categories regarding asymmetry in collocations: **right-and-left predictive**, **left-predictive**, and **right-predictive**. In some fossilized phrases (section 3.2.1) one element suggests the presence of the other one to the same extent as vice versa, e.g. $\overleftrightarrow{aurora\ borealis}$; such phrases are **right-and-left predictive**. He emphasizes, however, that it is more common for fossilized phrases to be unidirectional: for instance, in collocations such as $\overrightarrow{artesian\ well}$, $\overrightarrow{bonsai\ tree}$, $\overrightarrow{wellington\ boots}$, *etc.* the first element is highly predictive of the second one. These collocations are **right-predictive**. In the **left-predictive** phrases, the second word suggests the first one, e.g. $\overleftarrow{ad\ infinitum}$. In semi-fossilized phrases, several words compete for the spot in a collocation. However, the options are very limited: e.g. the word *caesarian* has only two noun options to modify: *operation* and *section* (Kjellmer, 1991). Most of semi-fossilized phrases are right-predictive, but there are also some examples of the left-predictive ones, e.g. $\overleftarrow{inferiority/Oedipus/persecution\ complex}$. Similar ideas are encountered in Stubbs (2001), who points out that the asymmetry from left to right is more common than from right to left ($\overrightarrow{bonsai\ tree}$, $\overrightarrow{cushy\ job}$).

These observations are based on comparing raw frequencies of the constituents of collocations. Similarly to the development of bidirectional measures, different attempts have been made to develop statistical measures for evaluating the directionality and the degree of asymmetry in collocations.

One of the first studies of that kind is described in Michelbacher et al. (2007). The focus of the study is on paradigmatic word pairs such as *man/woman*, *pepper/salt*, *happy/sad*. The authors propose two measures of asymmetry to evaluate the association between the words based on the corpus data: rank measure and conditional probability. The results are compared to the values of psychological association norms obtained from the USF Free Association Norms database[12] (Nelson et al., 1998).

The proposed rank measure is based on calculating χ^2 score (eq. (3.4) based on the notation in contingency table 3.1) for all the combinations $word_1 word_2$, creating n-best lists of collocates for each word and assigning a rank to each collocate.

$$\chi^2 = \frac{n \times (ad - bc)^2}{(a+b) \times (a+c) \times (b+d) \times (c+d)} \tag{3.4}$$

[12] http://w3.usf.edu/FreeAssociation/, *last accessed on 23.10.2020*

The second proposed measure is a simple conditional probability:

$$P(w_2|w_1) = \frac{a}{a+b} \qquad P(w_1|w_2) = \frac{a}{a+c} \tag{3.5}$$

Although both measures correspond to some extent to the human association scores between the words, their accuracy is still quite low. Moreover, the conditional probability measure performs poorly in predicting symmetric relations in contrast to the rank measure (Michelbacher et al., 2007).

In the follow-up study, Michelbacher et al. (2011) extend their investigation of asymmetry to syntagmatic relations between words using the two above mentioned unidirectional methods and additional rank measures (calculated from G^2, t-score and raw frequencies). They focus on adjective-noun constructions such as *wishful thinking* and noun-noun phrases such as *Christmas decorations*. The experiments demonstrate that both measures are good at predicting the asymmetry in syntagmatic phrases, but perform rather poorly when dealing with symmetric combinations. The results also show that right-predictive asymmetry (as in $\overrightarrow{bonsai\ tree}$) is more common than left-predictive, which goes in line with the observations by Stubbs (2001) and Kjellmer (1991).

The origin of another asymmetry measure, ΔP, goes back to psychological research concerned with associative learning theory. According to Ellis (2006), the unidirectional ΔP is better suited for measuring the dependency between a pair of events (*cue* and *outcome*) than the two-way dependency measures such as χ^2. ΔP is defined as follows:

$$\Delta P = P(outcome|cue) - P(outcome|\neg cue) \tag{3.6}$$

ΔP is the difference between two conditional probabilities and takes values within the range [-1,1]. Basically, it is an adjusted conditional probability. The proximity to zero means very low covariation between the cue and the outcome, values close to 1 indicate that the presence of the cue increases the probability of the outcome and negative values imply that the cue reduces the likelihood of the outcome[13] (Ellis, 2006).

Gries (2013) adopts the ΔP measure in collocations research. For each collocation pair, two values are calculated and compared:

$$\Delta P_1 = \Delta P(w_1|w_2) = \frac{a}{a+c} - \frac{b}{b+d} \tag{3.7}$$

$$\Delta P_2 = \Delta P(w_2|w_1) = \frac{a}{a+b} - \frac{c}{c+d} \tag{3.8}$$

[13] See also Allan (1980) for the comparison of ΔP, χ^2, and Δd.

As illustration of this method, consider as example the collocation *ungebetener Gast* 'uninvited guest' (table 3.3).[14]

	Gast	**other nouns**	**total**
ungebeten	1,113	927	2,040
other adjectives	93,893	194,619,584	194,713,477
total	95,006	194,620,511	194,715,517

TABLE 3.3: Contingency Table of observed frequencies for *ungebeten* 'uninvited' and *Gast* 'guest

Based on the observed frequencies, the two ΔP values are computed (eq. (3.7) and eq. (3.8):

$$\Delta P_1 = \Delta P(ungebeten|Gast) = 0.0117$$
$$\Delta P_2 = \Delta P(Gast|ungebeten) = 0.5451$$

The low value of ΔP_1 and the high value of ΔP_2 suggest that this collocation is highly asymmetric: *ungebeten* is a much better predictor of *Gast* than vice versa. In Stubbs (2001)'s and Kjellmer (1991)'s terminology, it is a right-predictive collocation: $\overrightarrow{ungebetener\ Gast}$.

In his study, Gries (2013) makes a comparison between the performance of ΔP and several bidirectional measures commonly used in the task of automatic collocation extraction. The main difference is that bidirectional measures manage to detect strong association between two components, but fail to reveal that, in numerous examples, this attraction is not entirely mutual, but rather asymmetric. For example, the two phrases *de facto* and *ipso facto* score highly in such bidirectional measures as MI, t, Dice, and G^2. However, those high scores do not reveal the directionality in these phrases.

phrase	$\Delta P_1(w_1 \mid w_2)$	$\Delta P_2(w_2 \mid w_1)$
de facto	0.75	0.01
ipso facto	0.25	1.00

TABLE 3.4: ΔP values for the phrases *de facto* and *ipso facto* from Gries (2013, p. 147)

In the case of *facto*, there is little competition for the slot before it: *de* or *post* or *ipso* which is reflected in a rather high ΔP_1 score of 0.75. ΔP_2 for *de facto* is very low which indicates that there are a lot of candidates for the slot after *de*. On the contrary, *ipso* can only be combined with *facto* and there is no competition at all for the slot after *ipso*, the ΔP_2 for *ipso facto* equals 1.0. In terms of directionality, $\overleftarrow{de\ facto}$ is left-predictive and

[14] Frequencies were calculated based on a large corpus consisting of the treebanks Wikipedia 2017 and Wikipedia 2018 (de Kok and Pütz, 2019) and decow16ax (Schäfer and Bildhauer, 2012; Schäfer, 2015)

ipso $\overrightarrow{facto}$ is right-predictive. To better understand the nature and properties of asymmetric combinations, it is necessary to take into account their type-token distributions and entropies (Gries, 2013).

The type/token distribution and information theoretical measures such as *entropy* and *relative entropy* and their relevance for collocation studies have been getting more attention recently. The idea stems from research in corpus linguistics and naive discriminative learning. According to Baayen (2011), ΔP shows a strong correlation with a measure that comes from information theory, *relative entropy* or *the Kullback-Leibler Divergence*. KL-divergence (D_{KL}) measures the difference between two probability distributions p and q, where p is the distribution of a stimulus in the presence of the cue and q is the general distribution of the stimulus. It is defined as follows:

$$D_{KL} = \sum_i p_i log_2 \frac{p_i}{q_i} \tag{3.9}$$

The D_{KL} measure is proposed by Gries and Durrant (2021) as an alternative to ΔP for measuring directionality in collocations. Basically, D_{KL} compares the distribution of a word when it is a constituent of a collocation and the general distribution of this word. For the earlier described example *ungebetener Gast* 'uninvited guest' (table 3.3), the following scores are computed (based on the notation in contingency table 3.1):

$$D_{KL_1} = \frac{a}{a+c} \times log_2 \frac{a \times n}{(a+b) \times (a+c)} + \frac{c}{a+c} \times log_2 \frac{c \times n}{(a+c) \times (c+d)} = 0.10185 \tag{3.10}$$

$$D_{KL_2} = \frac{a}{a+b} \times log_2 \frac{a \times n}{(a+b) \times (a+c)} + \frac{b}{a+b} \times log_2 \frac{b \times n}{(a+b) \times (b+d)} = 5.00838 \tag{3.11}$$

Similarly to the results obtained from ΔP scores, the values of D_{KL} confirm that this collocation is right-predictive: the presence of the word *ungebeten* gives us more information than the presence of the word *Gast*. The D_{KL} measure has not yet been studied much in connection with collocations. Coming from the information theory,[15] it has an extra advantage in the interdisciplinary approaches (Gries and Durrant, 2021).

One of the questions arising in the discussion about directional measures is whether they are better suited for automatic collocations extraction. As already discussed in section 3.4.1, multiple studies confirm that there

[15] Another application of a measure that comes from information theory – entropy – for studying collocations is described in Kilgarriff (2006), where it serves a different purpose: not evaluating the degree of association between the constituents, but measuring the *collocationality* of words to identify the words that tend to appear in collocations.

$$H(x) = -\sum p(x) \times log_2(p(x))) \tag{3.12}$$

Words that tend to co-occur with a limited set of words rather then combine freely with many different words would have lower entropy. The potential use of such a measure is identifying words that should be listed in dictionaries.

is no universal method for this task. Performance of different association measures depends on various factors. The situation is not different for the directional measures. Schneider (2020) compares ΔP to G and MI and does not find any evidence that an approach using a directional measure is better or worse than the one based on bidirectional AMs. A more recent study by Garcia et al. (2019) shows that bidirectional measures yield better results than the directional ones in spite of the asymmetric nature of a large number of collocations. An interesting point made by Garcia et al. (2019, p.54) is that it is expected for ΔP_2 to perform poorly because it "encodes the directionality from the collocate to the base"which is the opposite of the theoretical description of collocations. The results of the experiments in Garcia et al. (2019) show that ΔP_1 indeed slightly outperforms ΔP_2. The theoretical description of collocations adopted by Garcia et al. (2019) follows Mel'čuk's theory: when producing a text, the speaker first selects a base of the collocation freely and then, depending on the base, chooses the collocate (Mel'čuk, 2012a). At the same time, Kjellmer (1991); Stubbs (2001); Michelbacher et al. (2011) observe that right-predictive collocations (from collocate to base) are more common which suggests that ΔP_2 should be more successful at revealing collocations. It has not been studied yet whether directional measures reflect the way collocations are constructed.

It seems most plausible that such generalizations cannot be made for all types of collocations. Thus, in AN combinations such as *starker Regen* 'heavy rain', the adjective undergoes the process of delexicalization, i.e. its meaning is becoming more and more vague and the adjective can modify a large number of nouns. The entropy of such adjectives is expected to be very high and thus such AN pairs are probably left-predictive or are close to symmetric, and conventional measures such as logDice would be successful at extracting them. At the same time, there is another type of AN combinations in which the adjective is highly specific and combines only with a very limited set of nouns, e.g. *sintflutartiger Regen* 'torrential rain', for which ΔP_2 or D_{KL_2} are expected to reflect this asymmetry. This topic will be addressed in more detail in Chapter 8 that describes a corpus study of directionality in adjective-noun combinations.

3.5 Defining collocations and free phrases

Previous sections addressed terminological issues that inevitably arise in the discussions about collocations. Semantic and statistical criteria relevant for collocations have been discussed in detail. In this section, all the pieces of the puzzle are put together in order to provide working definitions of *collocation* and *free phrase* that will be used throughout this thesis.

3.5.1 Structure of collocations

Typically, collocations are regarded as binary expressions: they consist of two elements where one element is considered the **base** and the other one

– its **collocate**. The base is freely chosen by the speaker independently of the collocate, whereas the choice of the collocate is restricted depending on the base: the collocate is *"chosen as a function of the base to express a given meaning bearing on the base"* Mel'čuk (2012a, p.39). In the collocation *starker Regen* 'heavy rain' (lit. 'strong rain'), the noun *Regen* 'rain' is the base and the adjective *starker* 'strong' is the collocate that expresses the meaning of intensification.

Hausmann (1984), who treats collocations primarily from the lexicographic point of view, also recognizes two elements of a collocation – *Basis* 'base' and *Kollokator* 'collocate'. The two elements have different status: the base is independent and autonomous, the collocate is dependent on the base. In the process of creating an utterance, the speaker always starts with a base.[16] In most cases, that means that a collocation is constructed from a noun to an adjective or to a verb: when a speaker plans to say something about *a mistake*, it is highly unlikely that they first have in mind what modifier they want to use, for instance *grave*, and only then search for an appropriate noun that can be described by that modifier. Rather, they start off with the noun *mistake* and then search in the mental lexicon (or a dictionary) for a suitable collocate (Hausmann, 1984). The same observation is made in Fillmore et al. (2003): in support verb constructions, the noun selects the verb and not the other way around. As illustrated in example (13a), *prayer* selects *say* as a support verb, whereas *speech* selects *give* (13b). Swapping the verbs is not possible as in examples (13c) and (13d).

(13) a. *say a prayer*
b. *give a speech*
c. **give a prayer*
d. **say a speech*

The process is similar when translating an expression into a foreign language. For instance, when translating the collocation *a grave mistake* into German, one begins with the base: *mistake* → *Fehler*; this stage is usually not complicated as the base retains its literal meaning. Choosing the correct translation for the collocate poses a challenge: depending on the noun it modifies, the adjective is translated differently. With *mistake* being the base, the correct translation of *grave* is *grob* 'coarse/rough'. For a different collocation, *a grave danger*, the noun's translation is straightforward: *danger* → *Gefahr*, however, the collocate is again problematic. The correct translation is not *grob* as in the previous example, but *akut* 'acute' or *groß* 'big/great'. Thus, the collocate does not always keep the same meaning outside of a collocation, whereas the meaning of the base is independent.

Binary collocations described above are, however, not the only possible constellation of several words co-occurring together frequently. One word can simultaneously be a constituent of two collocations. Such expressions

[16] For this reason, collocates should be listed in a dictionary under their bases (Hausmann, 1984).

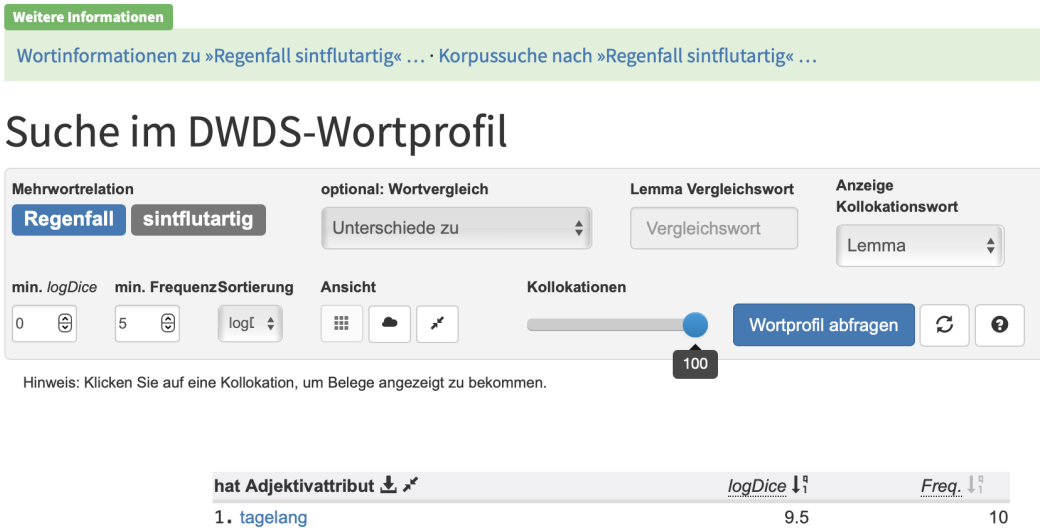


FIGURE 3.1: Word profile for the AN pair *sintflutartiger Regen* for the attribute relation.

are termed *collocation chains* by Alonso Ramos and Wanner (2007). For instance, in the phrase *to make a huge mistake*, the noun *mistake* is the object of *to make* and is modified by the adjective *huge*. The collocation consisting of three elements can be decomposed into two binary collocations: *to make a mistake* and *a huge mistake*. Such a pattern is rather frequent. However, there are also collocation chains that cannot be broken down into two separate collocations. Consider the expression *to follow the rules rigidly*: here the adverb modifies the collocation *to follow the rules* as a whole rather than just the verb *to follow* on its own (Alonso Ramos and Wanner, 2007).

Collocation chains have not been studied extensively, the only known systematic study addressing this issue in detail is described in Alonso Ramos and Wanner (2007). Collocation chains are formally defined as sequences consisting of three elements where collocational relations hold either between the elements individually, or between one element and two other elements taken as a holistic unit. This thesis focuses primarily on binary adjective-noun combinations as stated in section 1.1. However, as a possible direction for future work, it would be interesting to investigate what semantic groups of adjective-noun collocations have an affinity to occur in collocation chains. Especially, since the current version of the Wortprofil has an option of constructing a word profile for MWEs. For instance, the AN pair *sintflutartiger Regen* 'torrential rain' is marked as MWE in the Wordprofil and for this MWE, a separate word profile can be constructed (see Figure 3.1)[17] It shows that the phrase *sintflutartiger Regen* co-occurs with the adjective *tagelang* 'lasting for days' as its modifier.

[17]DWDS-Wortprofil for „Regenfall sintflutartig", created via the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/wp/sintflutartig/?limit=100&mwe=-17495355>, *last accessed on 12.09.2023.*

3.5.2 Working definitions

In order to make the following discussion easier to follow, the definition of literal meaning given in section 3.3.3, is presented here:

Definition 3
Literal meaning is the most basic meaning of a word (German: *Grundbedeutung*). Following the definition by Pragglejaz Group (2007, p.3): the most concrete, imaginable, historically older meaning, it is more precise and is related to a bodily action.

This section provides working definitions of the concepts *collocation* and *free phrase* that reflect three aspects: morpho-syntactic form, statistical properties, and meaning.

Form. Adjective-noun (AN) pairs are bigrams with a noun as head and an attributive adjective as modifier.

Statistical behaviour. This thesis focuses on recurrent AN combinations. Identifying such combinations goes beyond their simple raw frequencies. Previous section (3.4) thoroughly reviewed the common methods employed for calculating the level of recurrence of binary word combinations and presented the available tools: Sketch Engine for English and the Wortprofil for German.

Definition 5
An AN pair is considered a **co-occurrence** (following the terminology in Evert (2004) if the statistical association between its components is strong.

There are two ways of defining what *strong* means in this case and both are rather arbitrary: either to determine a particular numeric cut-off threshold of an association score or to determine a cut-off on the number of word pairs in an n-best list, i.e. to take into consideration the top-n word pairs ranked based on the chosen association measure. In AN co-occurrences, the head (noun) is considered the *base* and its modifier (adjective), the *collocate*.

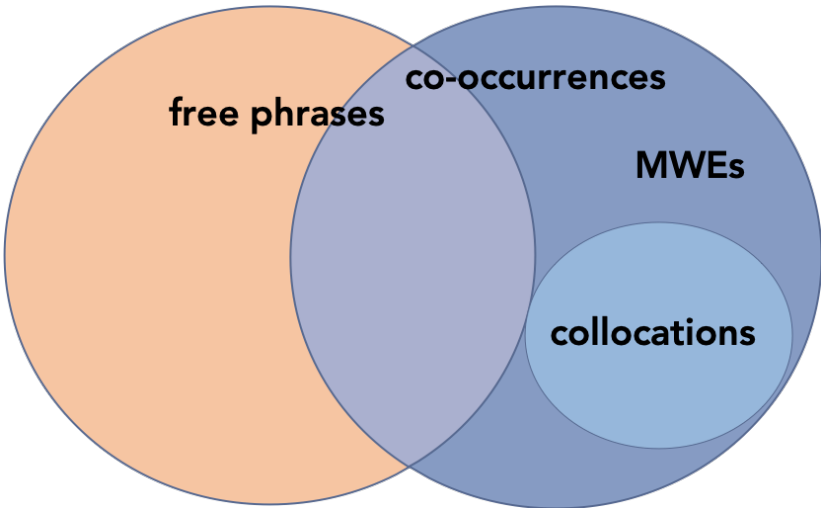


FIGURE 3.2: Relation between co-occurrences, collocations, MWEs in general, and free phrases.

Figure 3.2 roughly illustrates the relation between co-occurrences, collocations, MWEs in general, and free phrases. MWEs are a subset of co-occurrences; collocations are a subset of MWEs. This approach follows Evert (2004): collocations (and MWEs in general) must be recurrent in order for them to be learned and stored by speakers. Some free phrases can also be recurrent, for instance, *scharfes Messer* 'sharp knife'; hence, there is an overlap between free phrases and co-occurrences. The difference between recurrent free phrases and MWEs is based on lexical-semantic criteria. More specifically, in MWEs, the choice of at least one of its elements is constrained (Mel'čuk, 2023).

Meaning. There are many types of AN co-occurrences that can make their way into an n-best list; remember the various typologies discussed in section 3.2. The following discussion will focus on two types of AN co-occurrences distinguished based on semantic properties: free phrases and collocations. In free phrases, both elements are chosen by a speaker freely as long as the phrase is semantically well-formed. For instance, in *grünes Kleid* 'green dress', the adjective describes the color of the dress and is chosen freely by a speaker. Any object that can be described in terms of color, can be modified by the adjective *grün*: *wall*, *pencil*, *liquid*, *field*, *hair*, *etc.* Building on this knowledge, a speaker can predict that, say, a *box* would also combine with the adjective *green* in the same meaning. In collocations, the head is freely chosen by a speaker and the choice of the collocate is lexically restricted, i.e. the adjective can combine in this particular sense only with certain nouns.

This approach follows to a large extent Mel'čuk (2012a) and Nesselhauf (2003). Nesselhauf (2003) proposes two criteria to identify a restricted sense of a collocate:

1. The collocate cannot be used in this sense with all words that are syntactically and semantically possible.

2. The sense of the collocate is so specific that it only allows its combination with a small set of words.

The first criterion corresponds to Mel'čuk's (1996; 2012a) notion *standard collocation*, for instance *helle Aufregung* 'extreme excitement'. The adjective *hell* expresses intensification, in this sense it combines with a number of nouns: *Freude* 'joy', *Empörung* 'outrage', *Begeisterung* 'enthusiasm', etc. – nouns expressing emotions. However, not all emotion nouns allow intensification by *hell*: *Angst* 'fear', *Ekel* 'disgust', *Sympathie* 'sympathy', *Heiterkeit* 'amusement', etc. Therefore, the use of the adjective *hell* as intensifier is lexically-restricted, but is acceptable in combination with a considerable number of nouns.

The second criterion can be compared to the notion *non-standard collocation* in Mel'čuk's theoretical framework. For instance, in *schwarzer Kaffee* 'black coffee', the adjective has a very specific sense "without addition of dairy products". In this sense, the adjective does not combine with any other nouns (Mel'čuk, 2012b,a, 2015). Such lexemes are termed *quasi-unilexemes* by Mel'čuk (2023). A different example of a non-standard collocation discussed in Mel'čuk (2015, p.176) is *artesian well* in which the adjective combines with very few nouns: *well*, *water*, *pressure*. Such lexemes are referred to as *unilexemes* by Mel'čuk (2023). The difference to the example *black coffee* lies in the fact that the adjective *artesian* does not have any other senses, whereas *black* is a highly polysemous adjective (14 senses in WordNet). In the case of *black*, a meaning shift has taken place: from color sense to the way a coffee is prepared (without milk). The process of meaning shift also takes place in the examples of standard collocations given above: the adjective *hell* freely combines with various nouns in the sense 'bright'. After a meaning shift to 'intense', the combinability of the adjective becomes restricted.

A final issue that needs to be clarified is how to determine whether a collocate's meaning is restricted. As discussed above, in standard collocations and some non-standard collocations, the modifier undergoes a meaning shift. If an adjective in its literal sense (as defined in (3), combines freely with all syntactically and semantically possible nouns, a given AN pair is a free phrase. In AN collocations such as *helle Aufregung* and *schwarzer Kaffee*, the sense of the adjective is not its literal sense. Finally, in collocations such as *artesian well*, no meaning shift takes place. Both the head and the collocate retain their literal meaning, but the collocate's combinability is extremely limited.

Building on the previous discussion, adjective-noun collocations and free phrases are defined in this thesis in the following way:

Definition 6
Adjective-noun collocations are recurrent binary word combinations consisting of a base (noun) and a collocate (adjective) where the base is used in its unrestricted sense and the collocate's meaning is restricted. There are two types of restricted collocates:

Type 1: the collocate's sense is not literal. Type 1 collocations have two subtypes:

Type 1.1: the collocate combines in this sense with a large number of nominal bases, for instance, *helle Aufregung* 'extreme excitement' (lit. 'bright').

Type 1.2: the collocate combines in this sense with very few nominal bases, for instance, *schwarzer Kaffee* 'black coffee'.

Type 2: the collocate's sense is literal and it can be combined only with very few nominal bases, for instance, *artesian well*.

Definition 7
Adjective-noun free phrases are binary word combinations consisting of a base (noun) and a modifier (adjective) where both the base and the modifier are used in their literal senses, for instance, *grünes Kleid* 'green dress'. Some free phrases can exhibit strong statistical association between the base and the modifier (as illustrated in Figure 3.2).

Advanced techniques of automatic collocation extraction developed over the last few decades boosted the compilation of specialized collocation dictionaries for language learners. Apart from that, empirical studies on the statistical properties of collocations produced a few datasets suitable for further collocation research. Next section will present such resources that might be valuable for the study undertaken in this thesis.

3.6 Collocation resources

There are two types of data collections of collocations. First of all, as collocations play an important role in lexicography and language learning, several dictionaries specializing only in collocations have been compiled for a variety of languages. Secondly, the extensive research on statistical properties of collocations requires datasets annotated according to certain guidelines driven by the research questions and hypotheses. This section introduces both kinds of resources and gives an overview of the existing data collections that prima facie seem suitable as the empirical basis for studying the semantics of German and/or English adjective-noun collocations undertaken in this thesis.

3.6.1 Collocation dictionaries

Collocations are language specific and due to their arbitrariness pose considerable problems to non-native speakers. Lexicographic work on this topic produced several collocation dictionaries. In such dictionaries, the lists of collocates are presented under their bases (head words) following the approach described in Hausmann (1984).

Oxford Collocations Dictionary for Students of English

The Oxford Collocations Dictionary (McIntosh et al., 2009) approaches collocations in a pragmatic way. It follows a very general definition to collocations: "the way words combine in a language to produce natural-sounding speech and writing" (McIntosh et al., 2009, p.v), and includes the full continuum of collocations from very weak ones that according to the definition adopted in this thesis are considered free combinations to the most restricted ones. The main objective for compiling this dictionary is to support a language learner. The collocations are grouped based on grammatical relations (e.g adjective+noun, noun+verb, adverb+adjective, etc), but no information about the semantics is provided apart from the general logic that groups of synonymous collocations are presented from the 'strongest' to the 'mildest'.[18]

Macmillan Collocations Dictionary for Learners of English

Macmillan Collocations Dictionary (Rundell, 2010) is a dictionary aimed at helping English learners to speak fluently and sound natural where collocations are defined as two or more words that frequently co-occur. In addition to organizing the collocates of a headword according to grammatical patterns, the dictionary provides additional semantic information. The collocates of each word are grouped into *sets* based on their meaning. Each set is defined, either using a paraphrase or the most prototypical collocate. For instance, in the entry for the noun *argument*, for the relation adjective+noun, the following sets are listed:

1. **strong** compelling, convincing, persuasive, powerful, strong, valid
2. **about a particular subject** ethical, legal, philosophical, political, theological

The meaning of the first set is defined by one of the collocates, whereas for the second set, a paraphrase is used. Such a system is very practical for the purposes of learning a foreign language, however, from the theoretical point of view, the described approach to collocation classification lacks system. For instance, the same meaning of intensification expressed as 'strong' in the above described entry for *argument* is defined differently for various headwords: *control – absolute, contribution – big, magnitude – great, rain – heavy*.

Wörterbuch der Kollokationen im Deutschen

The Dictionary of Collocations in German (Quasthoff, 2011) defines collocations as recurring word pairs and the main purpose of the dictionary is to help to combine German words correctly. The order of presenting the

[18] For instance, in the entry for the head noun *acid*, its modifiers are presented in the following way: *concentrated, strong; dilute, weak*.

collocates is based, first of all, on their syntactic structure. The second criterion for grouping the collocates is of semantic nature and is similar to the approach in the Oxford Collocations Dictionary (McIntosh et al., 2009):

- **verbal collocates** are sorted according to the temporal aspect, the so-called *life cycle*: start → end. For instance, for the base *Liebe* 'love' as a subject: *begin* → *inspire* → *last* → *end* → *die*.

- **adjectival collocates** are sorted according to the adjectives' evaluative meaning: positive/strong → neutral → negative/weak. For instance, for the base *Liebe* 'love': *complete* → *happy* → *passionate* → *mutual* → *unrequited* → *faded*. At the end the collocates that refer to a domain are listed: *physical* → *platonic*.

This kind of semantic grouping is rather coarse-grained and there are no clear distinctions between the groups.

Conclusion

Collocation dictionaries are valuable resources with the main purpose of supporting language learners and helping them sound natural. This being the main criterion for considering an expression a collocation results in a lack of a clear-cut definition of collocation as a linguistic phenomenon. Thus, using a sample of data from a dictionary is an option that requires further filtering and/or additional corpus search of candidates that comply with the more strict definition of collocations adopted in this thesis. The next subsection introduces collocational databases that were created as part of linguistic studies on the topic.

3.6.2 Collocation gold standards

Since collocations started receiving attention in corpus linguistics and computational linguistics, there appeared a need for gold standard datasets on which the experiments can be conducted. Most of such experiments focus on the task of automatic collocation extraction, for instance based on the German adjective-noun collocations (Evert, 2008), however, there has also been some research on the semantic properties of collocations (LexFunc Dataset (Espinosa-Anke et al., 2019).

German adjective-noun collocations (Evert, 2008)

The database was created as a resource for evaluation of automatic methods of extracting multi-word expressions. The dataset contains 1,252 German adjective-noun expressions derived from the Frankfurter Rundschau corpus.[19] The combinations were annotated by professional lexicographers

[19] http://www.elsnet.org/eci.html, *last accessed on 30.11.2020*

with the main criterion for an expression to be evaluated as a **true collocation** being its *usefulness* for a bilingual dictionary. The adjective-noun combinations were classified into six categories:

1. **True collocations**: partly lexicalized expressions: (*brennendes Problem* 'burning problem', *grüne Politik* 'green politics', *böses Blut* 'bad blood')
2. **Habitual combinations**: semi-compositional expressions with some idiosyncratic properties (*neuer Anlauf* 'another go', *ehemaliger Minister* 'former minister')
3. **Familiar combinations**: free phrases with high frequency (*bewaffneter Räuber* 'armed robber', *lange Geschichte* 'long story')
4. **Unclear status**: not directly relevant for a dictionary, but may still be of use for lexicographers (*nahe Zukunft* 'near future', *große Hoffnung* 'great hope')
5. **Non-collocational**: free phrases that might be used to understand or explain the headword, but are irrelevant for a dictionary (*erstes Semester* 'first semester', *hohe Preise* 'high prices')
6. **Trash**: combinations tagged or lemmatized incorrectly (*verliebte Wolke* 'cloud love' - name of a stage play)

In a more coarse-grained approach, the categories 1–3 can be treated as **true collocations** and the other phrases as **non-collocations**. There are two reasons why this dataset is not suited as the empirical basis in this thesis. First of all, the criteria for annotating an expression as a collocation are quite vague. For instance, no distinction is made between a collocation and an idiom (the idioms *böses Blut* 'bad blood' and *eiserne Hand* 'iron hand' are annotated as category 1). Besides, it is not clear how the distinction between free phrases and collocations is made: fully transparent phrases *illegale Einwanderung* 'illegal immigration' and *kleiner Fisch* 'small fish' are considered collocations, whereas the combination *hohe Preise* 'high prices' is annotated as a free phrase.

LexFunc Dataset (Espinosa-Anke et al., 2019)[20]

Most of the available datasets of collocations contain lists of word combinations annotated based on their status: idiom, collocation, free phrase etc. `LexFunc` Dataset is a different kind of database. It contains 10,077 lexical function (LF) instances of English which also include collocations (see section 4.2 for a more detailed overview of the theory behind LFs). The phrases in `LexFunc` are manually annotated by Igor Mel'čuk, all the keywords are disambiguated with the corresponding Babelnetid (Navigli and Ponzetto, 2012).[21] The dataset contains LFs represented by at least 30 instances and therefore does not include any examples of non-standard LFs.

[20] `https://www.upf.edu/web/taln/english-collocations`, *last accessed on 15.11.2020*
[21] `https://babelnet.org/`, *last accessed on 18.11.2020*

Adjective-noun collocations account for about 1/4 of the dataset and represent 7 standard lexical functions[22]. For instance, for the head word *question* the following collocates are listed:

- `AntiBon2` (subjective negative evaluation): *embarrassing*;
- `Magn` (intensifier): *big, difficult, haunting, serious, tough, tricky*;
- *Ver* (objective qualifier): *fair, proper*.

The dataset provides valuable information about the semantic relations between the constituents of collocations, however, it does not contain free combinations or idioms (or at least does not make such a distinction between the expressions). Apart from that, it is restricted to standard LFs only and lacks semantic variety: any combinations considered as instances of non-standard LFs (such as *black coffee*) are not included. One of the objectives of this thesis is modelling the cases that are not covered by standard Lexical Functions, therefore, this dataset cannot serve as the empirical basis.

3.6.3 Conclusion

This section provided an overview of the existing data collections that could potentially be used as empirical basis in the study of adjective-noun collocations and free phrases undertaken in this thesis. However, a closer analysis of the data shows that none of the existing resources comply with the criteria defined in this thesis. The collection should (1) cover different semantic classes, both for adjectives and nouns; (2) contain both collocations and free phrases; (3) be large enough to make general conclusions based on a quantitative analysis and to be used in machine learning experiments. Therefore, a major effort in this thesis will be dedicated to creating such a data set. The data creation process will be described in Chapters 5 and 6. Before moving to the practical part, a thorough comparison of theoretical frameworks will be performed in order to define an initial set of relations that can be applied to adjective-noun phrases.

[22] Magn, AntiMagn, Bon, AntiBon, AntiBon2, Ver, AntiVer

Chapter 4

Lexical-semantic modeling of adjective-noun co-occurrences

What do the German adjectives *stolz* 'proud' and *hoch* 'high' have in common? Or the adjectives *scharf* 'sharp' and *schwer* 'heavy'? or *dick* 'thick' and *eng* 'narrow'? Without any context – not much, but consider adding a head noun to each modifier: *Preis* 'price' in the first case, *Strafe* 'punishment' – in the second, and finally *Freund* 'friend' in the last one. All these adjectives, when collocating with these particular nouns, express the same meaning, namely `intensity`. Are there more examples like that in German and are there other adjectives that can express this meaning? The answer is yes, and plenty of them. Just to name a few: *helle Aufregung* 'great excitement' (lit. 'bright'), *starker Raucher* 'heavy smoker' (lit. 'strong'), *grober Fehler* 'gross mistake' (lit. 'coarse'), *unsinnige Freude* 'great joy' (lit. 'absurd'). Are there meanings apart from `intensity` expressed by collocations and if yes which ones? There are numerous examples from various semantic fields: *scharfer Hund* 'vicious dog' (lit. 'sharp'), *strenger Winter* 'severe winter' (lit. 'strict'), *weiche Droge* 'soft drug', *brisantes Thema* 'controversial topic' (lit. 'explosive').

This chapter discusses theoretical frameworks that can shed light on the semantics of adjective-noun phrases. Each theory provides a semantic classification that could be used as an inventory of relations that hold between the constituents of such phrases. The following theoretical frameworks will be investigated:

- The **Meaning-Text Theory** (MTT) (Mel'čuk, 1996, 1999) is the only theory that provides an inventory of relations specifically for collocations. The first study in this chapter will be based on the MTT (section 4.2);
- The theory of **Generative Lexicon** (GL) (Pustejovsky, 1991) gives an account of the semantics of nouns in terms of qualia roles (section 4.3);
- **Frame Semantics** (Fillmore, 1982) describes the semantics of content words in terms of semantic roles – Frame Elements. The lexical database FrameNet provides a broad semantic coverage (section 4.4);
- **Concept frames** (Barsalou, 1992; Löbner, 2014) that describe the semantics of nouns in terms of attribute-value matrices (section 4.5).

The approach to modeling with concept frame differs from the first three approaches in terms of the annotation target. Instead of defining all the possible properties that a head noun has and selecting the property expressed by a given adjective, the annotation is performed directly for the adjectival modifier.
At the beginning of each section, the basic idea of a theory will be introduced, followed by the presentation of the relations inventory within that theory that could be applied to adjective-noun phrases. Then, the inventory will be applied to the sample of carefully selected AN pairs described in section 4.1. The advantages and limitations of each approach will be discussed at the end of each study.

Each inventory will be evaluated in terms of its semantic **coverage** and the **level of granularity** it offers. The aim is to establish an inventory of relations that gives an account of the semantics for nouns from various semantic fields and has flexible granularity.

4.1 Data sample for case studies

In order to evaluate and compare the theoretical frameworks, a data sample has been created based on the study by Strakatova and Hinrichs (2019). The main criterion for selecting the nominal bases for the sample was their semantic variety. The semantic classes were assigned based on the German lexical-semantic network GermaNet (Hamp and Feldweg, 1997; Henrich and Hinrichs, 2010). In GermaNet, 23 semantic classes of nouns are distinguished. The initial data sample in Strakatova and Hinrichs (2019) contained nouns from eleven semantic classes. In this chapter, the sample is expanded to cover all 23 semantic classes.

The selected nouns are listed in table 4.1. All the nouns have medium to high frequency in the DWDS corpus[1] and less than five senses in GermaNet. The latter criterion has been taken into consideration in order to maintain a manageable level of ambiguity. High ambiguity would make it much more challenging to select corresponding collocates for the chosen sense and reduce the number of the available collocates for a given sense. Apart from that, in the selection of the nominal bases for the sample, their level of concreteness has been taken into account. A study by Naumann et al. (2018) shows that the contexts of abstract and concrete words differ consistently. Thus, it is expected that abstract nouns exhibit a different combinatorial behaviour compared to concrete nouns. The concreteness ratings stem from

[1] According to the DWDS frequency scale (https://www.dwds.de/d/worthaeufigkeit, *last accessed on 29.12.2022*): scales 2 to 6 which corresponds to at least 11,154 tokens in the DWDS corpus.

lexical unit	semantic class in GermaNet	concreteness
Schokolade 'chocolate'	*Nahrung* 'food'	7.848
Schuh 'shoe'	*Artefakt* 'artefact'	7.578
Hund 'dog'	*Tier* 'animal'	7.155
Fels 'cliff'	*natGegenstand* 'natural object'	6.797
Wald 'forest'	*Ort* 'location'	6.747
Baum 'tree'	*Pflanze* 'plant'	6.672
Bein 'leg'	*Körper* 'body'	6.433
Regen 'rain'	*natPhaenomen* 'phenomenon'	6.203
Freund 'friend'	*Mensch* 'person'	5.852
Winter 'winter'	*Zeit* 'time'	5.308
Droge 'drug'	*Substanz* 'substance'	5.267
Team 'team'	*Gruppe* 'group'	5.079
Kurve 'curve'	*Form* 'shape'	4.983
Objekt 'object'	*Tops* 'Tops'	4.938
Geruch 'smell'	*Attribut* 'attribute'	4.839
Preis 'price'	*Besitz* 'possession'	4.155
Strafe 'punishment'	*Geschehen* 'event'	3.888
Thema 'topic'	*Kommunikation* 'communication'	3.244
Ähnlichkeit 'similarity'	*Relation* 'relation'	3.160
Anzahl 'amount'	*Menge* 'quantity'	3.154
Ziel 'aim'	*Motiv* 'motive'	3.07
Angst 'fear'	*Gefühl* 'feeling'	2.785
Interesse 'interest'	*Kognition* 'cognition'	2.338

TABLE 4.1: The selected nominal bases for the sample sorted by their concreteness scores in the descending order. The corresponding semantic classes come from GermaNet. The concreteness scores are from Köper and Schulte im Walde (2016).

the database described in Köper and Schulte im Walde (2016).[2]

After the sample of nominal heads had been determined, adjectival collocates were selected for each noun. The initial data sample from Strakatova and Hinrichs (2019, p. 105) contains only "true AN collocations", i.e. AN pairs in which the modifier is lexically restricted by its head. The scope of this thesis is broader as it considers not only collocations, but also recurrent free phrases.

[2]In the original study by Strakatova and Hinrichs (2019), the concreteness ratings were taken from the English MRC database (Wilson, 1988). This decision had been motivated by the necessity to use English translations of the base nouns in order to utilize the Berkeley FrameNet. However, two LUs from the sample are absent from the English database: *price* and *similarity*. As there exists an alternative database for German, it has been decided to rely on the concreteness ratings from the German database by Köper and Schulte im Walde (2016) in this chapter.

The most concrete word in the database is the noun *Mühlestein* 'millstone' with the score of 9.877 and the least concrete one is the verb *herabspielen* 'to play down' with the score of 0.

Remember the definition of a recurrent AN pair or a *co-occurrence* from the previous chapter (definition 5): *an AN pair is considered a co-occurrence if the statistical association between its components is strong.* The components of co-occurrences are a *base* – a nominal head, and a *collocate* – an adjectival modifier. This chapter will make no distinction between free phrases and true collocations: the theoretical frameworks will thus be tested on both types of phrases. In order to select adjectival collocates for the chosen nominal bases, the first step is to obtain a list of adjectival modifiers co-occurring with those nouns. The second step is to select only those adjectival collocates which are of interest to the this thesis, namely, *descriptive* adjectives (as defined in section 2.1).

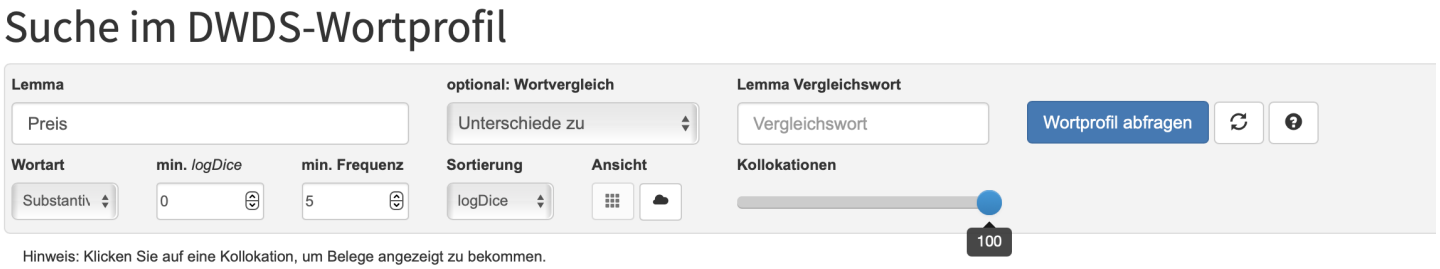
Step 1. The first step can be easily automated. As described in section 3.4, various statistical association measures have been developed for this purpose over the decades of collocation research. None of the measures is considered "the best" as there are many factors that influence their performance. However, some of those measures gained more popularity than others and became standard in collocation extraction tools. For instance, logDice (Rychly, 2008) is used in Sketch Engine for constructing word sketches (see section 3.4.1 for a detailed description).

For German, such a tool was developed as part of the digital dictionary of the German language, the DWDS. The tool is called the Wortprofil (Geyken et al., 2008; Didakowski and Geyken, 2014) and, inspired by Sketch Engine described in the previous chapter in section 3.4.1, it also makes use of the logDice association measure. The Wortprofil has several advantages over Sketch Engine. First of all, it is based on a large high-quality newspaper corpus of German: "Wortprofil 2021" uses 4.8 billion words whereas Sketch Engine provides only web corpora of comparable size which do not offer the same level of quality. Secondly, it is directly connected to the DWDS dictionary and thus indirectly also connected to GermaNet as some word senses in GermaNet were aligned with those in the DWDS dictionary (Henrich et al., 2014). Lastly, the Wortprofil is an open-access tool.

The Wortprofil provides a word sketch or a *word profile* by presenting an overview of statistically significant syntagmatic relations for a given word. For nouns, the following relations are listed: "has an adjectival attribute", "is predicative complement of", "is an object of", "is a subject of", "in coordination with", "is in prepositional group", "has genitive attribute", "has predicate", "has prepositional group", etc. For each relation, the Wortprofil can extract up to 100 collocates.

For extracting adjectival collocates of the selected nouns, the relation "has an adjectival attribute" has been considered. Figure 4.1 shows a part of the interface of the Wortprofil website: the top-20 adjectival collocates for the noun *Preis* 'price' sorted by the logDice values.[3]

[3]DWDS-Wortprofil for "Preis", created via the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/wp/Preis>, *last accessed on 03.10.2022.*



hat Adjektivattribut	logDice	Freq.
1. dotiert	9.7	8633
2. günstig	9.4	11870
3. steigend	9.3	10268
4. niedrig	9.3	13598
5. sinkend	8.9	5294
6. überhöht	8.9	4602
7. hoch	8.8	50333
8. fair	8.1	3155
9. stabil	8.0	3421
10. fallend	8.0	2511
11. gestiegen	7.6	2216
12. groß	7.6	35900
13. angemessen	7.5	2172
14. symbolisch	7.4	1885
15. erschwinglich	7.4	1593
16. attraktiv	7.4	2185
17. stolz	7.3	2422
18. halb	7.2	3553
19. vernünftig	7.0	1525
20. moderat	7.0	1360

FIGURE 4.1: A screenshot from the Wortprofil: top-20 adjectival collocates for the noun *Preis* 'price' sorted by the logDice values. The columns in the table are the following: the adjectival modifier, logDice value, frequency in the corpus.

For the current study, all the 100 available collocates are considered. The minimum frequency is set to five. For each noun, twenty attributive adjectives from the total of 100 are randomly selected.

Step 2. The randomly selected collocates are manually filtered: pertainyms are excluded from the study leaving only descriptive adjectives in the sample. The following AN pairs will be considered in the case study:

- ***Schokolade* 'chocolate'**: *schwarz* 'dark', *flüssig* 'liquid', *lecker* 'tasty', *dunkel* 'dark', *süß* 'sweet', *teuer* 'expensive', *weiß* 'white', *edel* 'premium', *heiß* 'hot', *hochwertig* 'of high quality', *knackig* 'crunchy'
- ***Schuh* 'shoe'** : *hochhackig* 'high-heeled', *flach* 'flat', *seiden* 'silk', *bequem* 'comfortable', *braun* 'brown', *schmutzig* 'dirty', *naß* 'wet', *schwarz* 'black', *rot* 'red', *weiß* 'white', *rosa* 'pink', *blau* 'blue', *teuer* 'expensive', *spitz* 'pointy', *elegant* 'elegant'
- ***Fels* 'cliff'** : *nackt* 'bare', *blank* 'bare', *schroff* 'craggy', *steil* 'steep', *glatt* 'smooth', *rötlich* 'reddish', *grau* 'grey', *naß* 'wet', *hart* 'hard', *schwarz* 'black'
- ***Hund* 'dog'**: *arm* 'poor', *süß* 'sweet', *scharf* 'vicious', *bissig* 'biting', *tollwütig* 'rabid', *treu* 'loyal', *aggressiv* 'agressive', *dick* 'fat', *klein* 'small', *niedlich* 'cute', *riesig* 'huge', *jung* 'young', *klug* 'smart', *dumm* 'stupid'
- ***Wald* 'forest'** : *licht* 'sparse', *tief* 'deep', *dicht* 'dense', *feucht* 'humid', *immergrün* 'evergreen', *unberührt* 'untouched', *trocken* 'dry', *nah* 'close', *riesig* 'huge', *herrlich* 'magnificent', *endlos* 'endless', *wild* 'wild'
- ***Baum* 'tree'**: *mächtig* 'mighty, massive', *kahl* 'leafless', *immergrün* 'evergreen', *morsch* 'rottern', *krank* 'sick', *alt* 'old', *gesund* 'healthy', *tot* 'dead', *riesig* 'huge', *heilig* 'sacred', *prächtig* 'magnificent', *jung* 'young', *hoch* 'high',
- ***Bein* 'leg'**: *staksig* 'spindly', *schwer* 'heavy', *wacklig* 'shaky', *nackt* 'naked', *müde* 'tired', *schlank* 'slim', *dünn* 'thin', *gesund* 'healthy', *lang* 'long', *kurz* 'short', *dick* 'fat', *schwach* 'weak', *stämmig* 'sturdy'
- ***Regen* 'rain'**: *sintflutartig* 'torrential', *stark* 'heavy', *strömend* 'pouring', *sauer* 'acid', *heftig* 'heavy', *warm* 'warm', *leicht* 'light', *spärlich* 'sparse', *kräftig* 'heavy', *kalt* 'cold', *fein* 'fine', *ewig* 'endless', *endlos* 'endless', *kühl* 'cool'
- ***Freund* 'friend'**: *alt* 'old', *wahr* 'true', *echt* 'real', *treu* 'loyal', *gut* 'good', *eng* 'close', *falsch* 'fake', *neu* 'new', *lebenslang* 'lifelong', *mächtig* 'powerful', *dick* 'close', *reich* 'rich', *nah* 'close', *langjährig* 'longtime', *verlässlich* 'reliable'
- ***Winter* 'winter'**: *streng* 'severe', *hart* 'harsh', *tief* 'deep', *mild* 'mild', *kalt* 'cold', *warm* 'warm', *bitterkalt* 'bitter cold', *eiskalt* 'ice cold', *letzt* 'last',

feucht 'humid', *trocken* 'dry', *lang* 'long', *kühl* 'cool', *ewig* 'endless', *trist* 'dull, depressing'

- ***Droge* 'drug'** : *hart* 'hard', *weich* 'soft', *leicht* 'soft', *illegal* 'illegal', *legal* 'legal', *körpereigen* 'endogenous', *gefährlich* 'dangerous', *harmlos* 'harmless', *flüssig* 'liquid', *tödlich* 'lethal', *billig* 'cheap', *schädlich* 'harmful', *sanft* 'soft',

- ***Team* 'team'**: *stark* 'strong', *schwach* 'weak', *eingespielt* 'coordinated, experienced',*schlagkräftig* 'powerful', *toll* 'excellent', *jung* 'young', *spanisch* 'Spanish', *deutsch* 'German', *gut* 'good', *komplett* 'entire', *klein* 'small', *erfahren* 'experienced', *homogen* 'homogenous', *neu* 'new'

- ***Kurve* 'curve'**: *scharf* 'sharp', *sanft* 'gentle', *steil* 'steep', *zackig* 'jagged', *blau* 'blue', *konvex* 'convex', *weich* 'soft'

- ***Objekt* 'object'**: *obskur* 'obscure', *dreidimensional* 'three-dimensional', *wertvoll* 'valuable', *entfernt* 'distant', *interessant* 'interesting', *hell* 'bright', *sakral* 'sacred', *zahlreich* 'numerous', *hochwertig* 'of high quality', *kompakt* 'compact', *teuer* 'expensive', *dekorativ* 'decorative'

- ***Geruch* 'smell'**: *beißend* 'acrid', *stechend* 'pungent', *unangenehm* 'unpleasant', *übel* 'foul', *modrig* 'musty', *muffig* 'musty', *süß* 'sweet', *streng* 'strong', *penetrant* 'pungent', *widerlich* 'disgusting', *ranzig* 'rotten', *brenzlig* 'burnt', *ekelhaft* 'disgusting', *intensiv* 'intense'

- ***Preis* 'price'**: *hoch* 'high', *fest* 'fixed', *horrend* 'horrendous', *stolz* 'stiff', *erschwinglich* 'affordable', *vernünftig* 'reasonable', *niedrig* 'low', *stabil* 'stable',*attraktiv* 'attractive', *moderat* 'moderate, *tief* 'low', *toll* 'great', *heutig* 'of today', *konstant* 'konstant', *rückläufig* 'descreasing'

- ***Strafe* 'punishment'**: *scharf* 'harsh', *streng* 'strict', *symbolisch* 'symbolic', *unmenschlich* 'inhumane', *drakonisch* 'draconian', *hart* 'harsh', *empfindlich* 'severe', *saftig* 'stiff', *deftig* 'severe', *niedrig* 'weak', *hoch* 'high', *mild* 'mild', *schwer* 'heavy', *schlimm* 'severe', *lebenslang* 'lifelong'

- ***Thema* 'issue/topic'**: *groß* 'big/major', *lokal* 'local', *heikel* 'delicate', *drängend* 'pressing' , *sperrig* 'unwieldy', *wichtig* 'important', *brennend* 'urgent', *heiß* 'hot', *unbequem* 'uncomfortable', *leidig* 'vexed', *sensibel* 'sensitive', *brisant* 'controversial', *beherrschend* 'dominant', *zentral* 'central', *aktuell* 'current', *interessant* 'interesting', *ewig* 'eternal'

- ***Ähnlichkeit* 'similarity'**: *frappierend* 'striking', *äußerlich* 'external', *frappant* 'striking', *groß* 'strong', *leise* 'slight', *stark* 'strong', *flüchtig* 'fleeting', *erstaunlich* 'surprising', *vage* 'vague', *unheimlich* 'uncanny'

- ***Anzahl* 'amount'**: *stattlich* 'stately', *groß* 'large', *gering* 'small',*schier* 'sheer', *beträchtlich* 'large', *hoch* 'high', *riesig* 'huge', *klein* 'small', *bescheiden* 'modest', *überschaubar* 'manageable', *beliebig* 'arbitrary', *erklecklich* 'considerable', *bedeutend* 'substantial', *variabel* 'variable', *exakt* 'precise'

- ***Ziel* 'aim'**: *ober* 'ultimate', *klar* 'clear', *hochgesteckt* 'ambitious', *groß* 'great', *langfristig* 'long-term', *hoch* 'high', *fest* 'firm', *primär* 'primary', *selbstgesteckt* 'self-imposed', *zentral* 'central', *ehrgeizig* 'ambitious', *hehr* 'noble', *gemeinsam* 'joint', *wichtig* 'important', *kurzfristig* 'short-term'
- ***Angst* 'fear'**: *groß* 'strong/big', *tief* 'deep', *pur* 'pure', *höllisch* 'hellish', *krankhaft* 'pathological', *blank* 'sheer', *nackt* 'pure', *diffus* 'vague', *dumpf* 'vague', *furchtbar* 'tremendous', *ständig* 'constant', *ewig* 'eternal', *stet* 'constant'
- ***Interesse* 'interest'**: *hoch* 'strong', *ureigen* 'own', *widerstreitend* 'conflicting', *groß* 'strong/big', *stark* 'strong', *lebhaft* 'lively', *echt* 'genuine', *wahr* 'genuine', *breit* 'broad', *handfest* 'tangible', *elementar* 'fundamental', *vital* 'vital', *massiv* 'massive', *rege* 'active', *gering* 'little'

This data sample of AN phrases will be used to test each of the theoretical frameworks that are described in detail in the next four sections.

4.2 The Meaning↔Text Theory

According to Mel'čuk (1999, p.9), natural language is a converter that transforms **meanings** into the corresponding **texts** and vice versa. **Text** is a special construct that represents the empirical reality of linguistics – speech. The number of samples of speech in natural languages is unlimited. **Meaning** is used synonymously to Fregean notion of *Sinn* 'sense' as opposed to *Bedeutung* 'reference'. The **Meaning↔Text Theory** (MTT) provides a system of rules that secure a transfer from each meaning to the corresponding text and the other way around. The analysis of meaning itself is outside of the scope of the MTT model. For instance, from the point of view of the English language, the phrase *a square circle*[4] is absolutely acceptable: it is formed in accordance with the rules of the language: an NP with an adjective as modifier. Translating the phrase into other languages would not eliminate the anomaly that is on the level of world knowledge. This type of anomaly is of no interest for the MTT and should be handled by a different model that Mel'čuk (1999) calls a "Reality↔Meaning" model. The Meaning-Text model is thus only sensitive to the "incorrectness of text" often referred to as *grammaticality*. According to Mel'čuk (1999), incorrectness can be **absolute** or **relative**. The former refers to grammatical violations of language, for instance, the incorrect use of the singular pronoun in the phrase *this women*. The second type of incorrectness is relative to the implied meaning. For example, the sentence *He carries a lecture* is an incorrect English translation of the German *Er hält einen Vortrag*. The correct one would be: *He gives a lecture*. However, if the implied meaning is that someone is moving and

[4]The original example used in Mel'čuk (1999), *kruglij kvadrat* 'round square', has been changed for more clarity: *a round square* is an acceptable phrase in English due to the ambiguity of the noun *square*: a geometrical shape cannot be round, but an open area in a city can. In Russian, this ambiguity does not apply as the word *kvadrat* has only the meaning of a geometrical form.

physically *holding* a lecture, then the initial sentence is perfectly correct (for a more detailed discussion see Mel'čuk (1999, pp. 24-25).

As Mel'čuk (1999) points out, the correspondence between meanings and texts is of many-to-many type which is manifested through synonymy (one meaning to many texts) and homonymy (one text to many meanings). Due to synonymy, there are thousands and even millions of possible paraphrases for any sentence in a language. As both synonymy and homonymy are ubiquitous in natural languages, a direct transition from meaning to text is nearly impossible. Therefore, the transition should be partitioned into separate consecutive phases which implies that a number of intermediate meaning representations must be introduced. The Meaning-Text Model has seven levels of meaning representation (Mel'čuk, 1999; Wanner, 1996):

1. Semantic Representation [SemR] – **meaning**
2. Deep-Syntactic Representation [DSyntR]
3. Surface-Syntactic Representation [SSyntR]
4. Deep-Morphologic Representation [DMorphR]
5. Surface-Morphologic Representation [SMorphR]
6. Deep-Phonetic Representation [DPhonR]
7. Surface-Phonetic Representation [SPhonR] – **text**

All the levels of meaning representation with the exception of [SemR] have two sublevels: *deep* and *surface*. The deep levels are oriented towards meaning, whereas the surface ones – towards text. In the context of the relations in collocations, only the first three levels are relevant:

(i) The transition from meaning to text starts with the Semantic Representation [SemR] at the semantic level. The main component of the [SemR] is a semantic graph where the nodes are the elementary semantic units corresponding to words senses (Wanner, 1996).

(ii) The Deep-Syntactic Representation [DSyntR] is a dependency tree where the nodes represent **generalized lexemes** that include real lexemes of a language, phrasemes, and Lexical Functions.

(iii) Finally, the Surface-Syntactic Representation [SSyntR] contains the actual word forms of the phrase including functional words organized as a dependency tree.

To make the transition from [SemR] to [DSyntR], the Speaker has to choose the corresponding lexical units (LUs) that express the intended meaning. Mel'čuk (2015) distinguishes between two types of LUs:

- **Free LUs** are chosen by the Speaker directly from the mental lexicon based only on their meanings and are independent of other lexical choices. Thus, the choice is **semantically-driven**.
- **Restricted LUs** are selected based on the other LUs chosen by the Speaker. This kind of choice is **lexically-driven**.

Lexically-driven choices of the Speaker can be **paradigmatic** or **syntagmatic**. Paradigmatic choice is of the type "either – or" or "instead of", for example, choosing a noun when referring to the type of emotion: *joy, fear* or *surprise*. Syntagmatic choices are of the type "and – and" or "together with", such as combinability of a head noun with an adjectival modifier. For instance, one can describe *joy* as *utter*, *great*, *sheer*, but not as *deep*. In contrast, the noun *fear* can be combined with all the listed adjectives including *deep*.

The dependencies between the LUs are described in the MTT systematically in terms of **Lexical Functions** (LFs). The next subsection is devoted to the notion of a Lexical Function.

4.2.1 Lexical Functions

Lexical Functions are functions in the mathematical sense: $f(x) = y$, where f is an abstract general meaning that is expressed by an LU y depending on the LU x. In the terminology of the Meaning-Text Theory adopted in this thesis, x is referred to as the *base* and y is its *collocate*. The formal definition of a Lexical Function is the following (Mel'čuk, 2015, p.166):

> "Formally, an LF **f** is a function that associates with a given LU L a set $\{L'_i\}$ of lexical expressions that express, contingent on L, the meaning 'f' associated with **f** and bearing on the meaning 'L': **f**(L) = $\{L'_i\}$, such that 'f'⇔ L'_i. L is called the argument, or keyword, of f; f(L) – the set $\{L'_i\}$ – is the value of **f** applied to L; and an L'_i is an element of this value."

Consider an example of the LF `Magn` that expresses the meaning of *intensification* applied to the noun *price* in English: `Magn` (*price*) = {*high, horrendous, stiff*}, where function **f** = `Magn` that means 'strong, very', lexical unit L = *price*, set of lexical expressions contingent on L $\{L'_i\}$ = {*high, horrendous, stiff*}. This is interpreted as follows: the adjectives *high, horrendous, stiff* express the meaning 'strong, very' for the lexical unit *price*.

Among all the possible LFs in a language, the MTT is only interested in those LFs that are expressed phraseologically, i.e. in a restricted way. All the meanings that are expressed in a non-restricted way are referred to as **LF-constants** and are not covered in the MTT (Mel'čuk, 1999). The inventory of LFs includes both syntagmatic and paradigmatic relations. The paradigmatic ones will not be considered in this thesis as they are not relevant in the context of collocations.[5] The syntagmatic LFs describe the relations that hold between the constituents of collocations. Mel'čuk (1999, 2015) distinguishes two large categories of LFs: **standard** and **non-standard**.

Standard LFs. There are two conditions that have to be satisfied in order to consider a Lexical Function standard: (1) broadness of the **f**'s **domain**; (2) broadness of **f**'s **range**. The first condition implies that the meaning corresponding to a given LF is general enough – it should combine with a large number of other meanings. The second criterion demands a rich variety

[5]Examples of paradigmatic LFs are `synonym`: `Syn`(*policeman*) = *cop*; `adjectival derivative`: A_0(*city*)=*urban*; `collective noun`: `Mult`(*ship*)=*fleet*; `standard metaphor`: `Figur`(*remorse*)=*pangs [of]*, etc. In total, there are 26 paradigmatic LFs.

of ways to express the meaning 'f' in a language. A typical example of a meaning that satisfies the two conditions is the meaning of *very* (intensification). Consider the following examples in English: 'very'(*illness*) = *grave*, 'very'(*smoker*) = *heavy*, 'very'(*headache)* = *bad*, 'very' (*wind*) = *strong*, 'very' (*appetite*) = *ravenous*, 'very' (*friend*) = *close*, etc. They illustrate that the meaning 'very' can be applied to a large variety of bases, i.e. its domain is broad – the first condition is thus satisfied.

The second condition of broadness of range is also met for the meaning 'very' as it is expressed in the English language by a large variety of lexical units, and the choice of the correct collocate is contingent on the base. Thus, the meaning 'very' has both a broad domain and a broad range, and therefore a corresponding standard LF (`Magn`) has been introduced (it will be discussed in more detail further in this section). This holds true not only for English, but also for many other languages: see Mel'čuk (2015, p.167) for ways to express intensification of the meaning 'skinny' in 23 different languages.

Now consider another example: the meaning 'having great weight'. It obviously meets the first requirement as it combines with a large number of keywords. Nearly any physical entity can be described in terms of its weight: *table, car*, *animal*, *rock*, *tomato*, *person*, *sofa*, *etc*. However, the lexical options for expressing this meaning are limited. In English, that would be the adjective *heavy* that can freely combine with a large variety of nouns (in some cases also the adjectives *massive and hefty*). The domain of the meaning is broad due to the almost unlimited amount of possible keywords, but the range is limited to only a few adjectives that express this meaning in English. From the point of view of the MTT, there is no need to introduce a standard LF for the meaning 'of great weight' and it is considered an *LF-constant* (Mel'čuk, 1999).

Standard LFs have formal fixed names that are abbreviations derived from Latin words. The above described meaning 'very' corresponds to the LF `Magn` from Latin *magnus* 'big, great'. Other examples are:

- `Cap` from Latin *caput* 'head' as in `Cap` (*university*) = *rector, president*; `Cap` (*aircraft*) = *captian, pilot*
- `Oper` from *operāre* 'to do, carry out' as in `Oper` (*order*) = *give*; `Oper` (*analysis*) = *carry out*
- `Manif` from *manifestāre* 'manifest' as in `Manif` (*distrust*) = *lurks*; `Manif` (*scorn*) = *drips*

Altogether, the inventory of LFs comprises 64 standard LFs as listed in Mel'čuk (2015): 26 paradigmatic LFs and 38 syntagmatic ones. The syntagmatic ones are further divided into 4 groups based on the part of speech of the collocates: nominal (1 LF), adjectival (5+2 LFs),[6] prepositional (6 LFs), and verbal (24 LFs).

[6]Two adjectival LFs are comparison markers and can only be encountered in combinations with other LFs (Mel'čuk, 2015)

Non-standard LFs express very specific meanings that combine only with one or very few semantically close bases. There exists a subtler division of non-standard LFs into *normal non-standard* and *degenerate* LFs, where degenerate LFs refer to LFs that are combined only with one base, for instance *artesian well* and *aquiline nose*. However, the distinction between the two categories is not entirely clear and, as Mel'čuk (2015) notes, is not relevant in practice.[7] In contrast to standard LFs, the non-standard ones do not have fixed names. As their meanings are very specific and non-general, each function is defined very precisely as in the above mentioned examples, and this definition serves as the name of the function.

A typical example of a non-standard LF is the meaning 'without addition of dairy product' applied to *coffee*: the corresponding collocation is *black coffee*. Another example of a non-standard LF, this time in Russian, is the meaning 'made from rye flour' applied to the noun *bread*: *tschyornij khleb* lit. 'black bread'. The combination with the adjective *tschyornij* 'black' in this sense is only possible for bread, but not other baked goods. The combinability is restricted, but both the domain and the range of this meaning is very limited. Therefore, for the both examples, non-standard LFs should be introduced (Mel'čuk, 1999):

`without addition of a dairy product` (*coffee*)=*black*

`made from rye flour` (*khleb* 'bread')=*tschyorniy* 'black' (Rus.)

Non-standard LFs are very numerous, yet, are highly non-systematic. There have been attempts to formalize non-standard LFs. For instance, Jousse (2007) proposes an approach to standardising nominal and adjectival LFs. The main assumptions of the proposed approach is that (1) all entities have *attributes* such as `size, color, shape, function, appreciation`, etc.; and (2) that such attributes are mainly expressed by adjectives. Jousse (2007) compares such attributes with *Frame Elements* from *Frame Semantics* (see section 4.4 for a detailed discussion of Frame Elements). The analysis of a sample of non-standard LFs conducted by Jousse (2007) supported the hypothesis: all the examples of non-standard LFs could be classified according to the corresponding attributes. Jousse (2007) gives the following examples:

- `color`: fr. *barbe poivre et sel* 'grey beard' (lit. 'pepper-and-salt beard')
- `cause`: fr. *célibat forcé/ obligé* 'compelled/forced celibacy'
- `consequence`: fr. *coup de feu mortel* 'fatal gunshot'

Simple and complex LFs. All the above discussed LFs are examples of simple LFs. In general, **simple** Lexical Functions are elementary, i.e. they are not further divisible into other LFs. However, there are some examples of simple LFs that can be represented in terms of other LFs, but they were

[7] Moreover, Jousse (2007) also mentions the category **semi-standard** LFs that are written in natural language but include a standard LF.

introduced as separate LFs because the meaning they represent is highly frequent in natural languages. Consider the LF `Liqu` (Lat. *liquidāre* 'liquidate') = 'do something so that something else does not take place'. This LF can be represented as `CausFin` – 'to do something so that something else takes place' + 'cease' (Mel'čuk, 2015, p.179). **Complex** LFs are combinations of two or more simple ones: for example, the complex LF `AntiMagn` denotes a meaning opposite to `Magn`.

Adjectival standard LFs

As the focus of this thesis lies on adjective-noun collocations, only syntagmatic adjectival Lexical Functions will be further be considered. The MTT introduces the following five simple standard LFs for adjectival modifiers:

- **`Magn`** from Latin *magnus* 'big, great' is interpreted as "very, to a (very) high degree, intense(ly), completely" (Mel'čuk, 2015, p.207). As already mentioned previously in this section, this LF is very productive and is found in a great number of languages. This meaning combines with LUs that already have a gradability element in their meaning. The LUs can be any part of speech. For adjective-noun phrases, some examples are:

  `Magn`(*rain*) = *heavy* `Magn`(*danger*) = *grave*
  `Magn`(*headache*) = *bad* `Magn`(*appetite*) = *ravenous*

- **`Ver`** from Latin *verus* 'real, genuine' interpreted as "such as [L] should be" (Mel'čuk, 2015, p.208). This LF is an objective qualifier of the base, it describes lexical units in terms of certain norms.

  `Ver`(*surprise*) = *genuine* `Ver`(*guess*) = *correct*
  `Ver`(*instrument*) = *precise* `Ver`(*joke*) = *appropriate*

- **`Bon`** from Latin *bonus* 'good' (Mel'čuk, 2015, p.209). It expresses a subjective evaluation, a praise or an approval from the Speaker. It combines with a great variety of LUs as it can be applies to anything that can be evaluated.

  `Bon`(*career*) = *shining* `Bon`(*opportunity*) = *golden*
  `Bon`(*lie*) = *white* `Bon`(*behaviour*) = *excellent*

- **`Pos`** from Latin *positīvus* 'positive'. It is interpreted as a"positive evaluation" (Mel'čuk, 2015, p.209) or a "received praise" of the referent of $DSyntA_{II}$ from the $DSyntA_{I}$. It is required that the meaning of the base have at least two actants ($DSyntA_{II}$ and $DSyntA_{I}$) which makes this LF quite rare.

Pos(*esteem*) = *high* Pos(*review*) = *positive*

According to Mel'čuk (2015), *positive review* is not the same as *good/brilliant/excellent review* which should be described by the LF Bon because a *brilliant review* can still be either positive or negative, and at the same time be *brilliantly written*.

- **Epit** from Latin *epitheton* interpreted as "redundant clichéd modifier" (Mel'čuk, 2015, p.207). The meaning of the modifier is almost empty and contributes little or nothing to the meaning of the base. There are two kinds of this LF: restrictive (Epith^{restr}) and qualificative (Epith^{qual}). The former one restricts the meaning of a very polysemous LU outside of context. The second one amplifies some mandatory feature of the base; in a way, it is semantically redundant.

Epith^{restr}(*current*) = *electrical* Epith^{qual}(*drizzle*) = *light*

Epith^{restr}(*machine*) = *vending* Epith^{qual}(*opinion*) = *humble*

All of the above presented LFs are semantically similar. In some cases, the classification according to those LFs may pose a challenge. For instance, if intensity is an central quality of something, then Magn would be equal to Ver and possibly to Bon if the intensity quality is perceived as 'good'. Mel'čuk (2015) illustrates this problem with the the phrase *good health* to which Magn, Bon and Ver are applicable. This challenge has been also corroborated by the results of the experiments in automatic classification of collocations reported in Espinosa-Anke et al. (2019). The authors observe a high error rate for the collocations labeled with the LFs Magn and Bon. The hypothesize that the confusion between the two categories is due to their semantic similarity.

The simple LFs can be combined into complex ones. The most common combination is with the LF Anti (antonym): e.g. AntiMagn as in *weak hope*; AntiBon as in *poor example*; AntiVer as in *empty threat*. A more complex arrangement of LFs is also possible, for instance AntiMagn + AntiBon as in *dull colour* or Magn+Bon=Ver as in *genuine art* (Mel'čuk and Zholkovsky, 1984). Sometimes, standard LFs are combined with non-standard ones, e.g. Ver + 'contains important details' as in *precise description*.

The more in-depth discussion of the inventory of Lexical Functions will be presented in the next subsection. The data sample of adjective-noun collocations introduced in 4.1 will be modeled in terms of Lexical Functions.

4.2.2 Case study: Lexical Functions

This section presents the first case study with the data sample described in section 4.1. The relations that hold between the constituents of the selected co-occurrences have been manually modeled based on the inventory of standard Lexical Functions listed in the previous section. In the cases where a standard LF could not be applied, a non-standard LF has been provided

as a definition of the corresponding sense of an adjective, e.g. "poor – inciting pity". The definitions are based on those provided by the Princeton WordNet (Fellbaum, 1998), Wikipedia, the DWDS,[8], and the Duden.[9] The discussion starts with the most concrete nouns followed by more abstract ones in the order presented in table 4.1.

Schokolade 'chocolate'
`Bon`: *lecker* 'tasty'
`with high percentage of cocoa and little/no milk and sugar`: *schwarz, dunkel* 'dark'
`made of cocoa butter, without cocoa solids`: *weiß* 'white'
`of high quality` or `Bon`: *edel* 'premium, fine', *hochwertig* 'of high quality'
`capable of flowing`: *flüssig* 'liquid'
`tasting like sugar`: *süß* 'sweet'
`high in price`: *teuer* 'expensive'
`firm and crisp`: *knackig* 'crunchy'
`beverage made from cocoa powder, milk and sugar`: *heiß* 'hot'

One collocate that could be modeled as a standard LF is *lecker* 'tasty' as it is a positive subjective evaluation of chocolate. None of the remaining phrases with the base *Schokolade* 'chocolate' could be modeled with standard LFs. A possible exception is the meaning "having high quality" which is semantically close to the LFs `Bon` or `Ver`. However, `Bon` is a subjective evaluation, whereas *fine/premium chocolate* is a commonly accepted notion. It expresses a certain type of chocolate the production of which involves many different factors. In that sense, it is rather an objective qualifier. The standard LF `Ver` corresponds to a similar meaning, but it still remains unclear whether *premium chocolate* is what chocolate should be. Therefore, it has been decided to treat it as a non-standard LF. Some of the phrases with *chocolate* are highly lexicalized and are direct hyponyms of the LU *chocolate* in the Wordnet, i.e. they are treated as adjective-noun compound phrases, for instance *white chocolate*. For each of them, a non-standard LF had to be introduced that gives a precise definition of each meaning.

[8] `https://www.dwds.de` (*last accessed on 21.07.2021*)
[9] `https://www.duden.de` (*last accessed on 21.07.2021*)

Schuh 'shoe'
`Bon`: *bequem* 'comfortable', *elegant* 'elegant'
`high in price`: *teuer* 'expensive'
`with high heels`: *hochhackig* 'high-heeled'
`Anti + with high heels`: *flach* 'flat'
`of certain color`: *braun* 'brown', *schwarz* 'black', *rot* 'red', *weiß* 'white', *rosa* 'pink', *blau* 'blue'
`made of silk`: *seiden* 'silk'
`soiled`: *schmutzig* 'dirty'
`becoming narrower and ending in a point`: *spitz* 'pointy'
`covered or soaked with a liquid`: *naß* 'wet'

Fels 'cliff'
`with nothing growing on it`: *nackt* 'bare', *blank* 'bare'
`of certain color`: *schwarz* 'black', *rötlich* 'reddish', *grau* 'grey'
`covered or soaked with a liquid`: *naß* 'wet'
`having a surface without roughness`: *glatt* 'smooth'
`abruptly sloping and fissured`: *schroff* 'craggy'
`rising up or sloping downwards almost vertically`: *steil* 'steep'
`firm, not yielding to pressure`: *hart* 'hard'

Only the standard LF `Bon` could be applied in modeling of *Schuh* 'shoe' for the adjectives *bequem* and *elegant*. The remaining co-occurrences with the base *Schuh* can only be described in terms of non-standard LFs. Similarly to the examples with *chocolate*, two phrases with *shoe* are highly lexicalized (*hochhackig* and *flach*) and denote *a type of* shoe, i.e. they are in a hyponymic *is_a* relation to the more generic term *shoes*.[10] The modifiers of the noun *Fels* 'cliff' could not be accommodated by any standard LF.

[10]Note that, in English, it is common to omit the noun *shoes* and use a metonymy or a deajdectival noun in plural to express the same meaning – *high heels* and *flats*.

Hund 'dog'
Bon: *süß* 'sweet', *niedlich* 'cute'
Magn^{size}: *riesig* 'huge'
AntiMagn^{size}: *klein* 'small'
having certain character, temper: *scharf* 'vicious', *bissig* 'biting', *treu* 'loyal', *aggressiv* 'agressive'
intelligent: *klug* 'smart'
intelligent + Anti: *dumm* 'stupid'
of certain age: *jung* 'young'
having too much flesh: *dick* 'fat'
infected with rabies: *tollwütig* 'rabid'
pitiful: *arm* 'poor'

Wald 'forest'
Magn^{loc}: *tief* 'deep'
Bon: *herrlich* 'magnificent'
Magn^{size}: *riesig* 'huge', *endlos* 'endless'
sparse and bright: *licht* 'sparse'
characterized by a large quantity of water vapor: *feucht* 'humid'
characterized by a large quantity of water vapor + Anti: *trocken* 'dry'
not shedding foliage throughout the year: *immergrün* 'evergreen'
in a natural state: *wild* 'wild', *unberührt* 'untouched'
at a short distance: *nah* 'close'
without gaps, crowding together: *dicht* 'dense'

Baum 'tree'
Magn^{size}: *mächtig* 'mighty, massive', *riesig* 'huge'
Magn^{height}: *hoch* 'high'
Bon: *prächtig* 'magnificent'
not shedding foliage throughout the year: *immergrün* 'evergreen'
having existed for a long time: *alt* 'old'
having existed for a long time + Anti: *jung* 'young'
impaired in physical well-being: *krank* 'sick'
impaired in physical well-being + Anti: *gesund* 'healthy'
in a state in which the vital functions have ceased: *tot* 'dead'
brittle due to age or decay: *morsch* 'rotten'
distinguished by special consecration: *heilig* 'sacred'
having no leaves: *kahl* 'leafless'

Bein 'leg'
Magn^{length}: *lang* 'long'
AntiMagn^{length}: *kurz* 'short'
of certain physical strength: *stämmig* 'sturdy', *schwach* 'weak'
having too much flesh: *dick* 'fat'
having too much flesh + Anti: *schlank* 'slim', *dünn* 'thin'
completely unclothed: *nackt* 'naked'
impaired in physical well-being + Anti: *gesund* 'healthy'
trembling, without a firm hold: *wacklig* 'shaky'
long and lean: *staksig* 'spindly'
in need of rest, without energy: *schwer* 'heavy', *müde* 'tired'

Regen 'rain'
$\texttt{Magn}_{[precipitation]}$: *sintflutartig* 'torrential', *stark* 'heavy', *kräftig* 'heavy', *heftig* 'heavy'
$\texttt{AntiMagn}_{[precipitation]}$: *leicht* 'light'
$\texttt{Magn}_{[drops]}$: *fein* 'fine'
$\texttt{Magn}^{quant}_{[drops]}$: *spärlich* 'sparse'
$\texttt{Magn}^{temp}$: *ewig* 'endless', *endlos* 'endless'
`of certain temperature`: *warm* 'warm', *kalt* 'cold', *kühl* 'cool'
`flowing strongly and in large quantities`: *strömend* 'pouring'
`with acidic components`: *sauer* 'acid'

For the five nouns, *Hund* 'dog', *Wald* 'forest', *Baum* 'tree', *Bein* 'leg', and *Regen* 'rain', there is a mixture of available LFs. On the one hand, the non-standard LFs are still necessary. On the other hand, standard LFs can be also applied to a few phrases. This is especially true for the noun *rain*: it is a process and not a physical entity like a *shoe*, therefore, intensification of its meaning is possible (expressed by the LF `Magn`). However, a specification is required of what aspect of *rain* is being intensified. This approach follows the modeling of the noun *dozhd'* 'rain' in the *Explanatory combinatorial dictionary of modern Russian* compiled by Igor Mel'čuk and his colleagues (Mel'čuk and Zholkovsky, 1984, p. 267). In the dictionary entry for *dozhd'* 'rain', the following specifications of the LF `Magn` are listed as subscripts and superscripts:

(14) a. $\texttt{Magn}_{[precipitation]}$ = *sil'niy* 'heavy'

b. $\texttt{Magn}_{[drops]}$ = *krupnie* 'large'

c. $\texttt{Magn}^{quant}_{[drops]}$ = *chastiy* 'thick' (a lot of drops)

d. $\texttt{Magn}^{temp}$ = *zatyazhnoy* 'continuous'

These examples illustrate how specifications of LFs are used to model different aspects of a very general meaning of intensification. A subscript is a "semantic component of the key word's definition" that indicates "that the meaning of this LF interacts exactly with this component" (Mel'čuk and Zholkovsky, 1984, p. 60). In this example, the words subscripted in square brackets, *precipitation* and *drops* are components of the definition of the LU *rain* given in (Mel'čuk and Zholkovsky, 1984, p. 267): "**precipitation** of liquid X, usually water, falling from celestial objects Y onto Z in the form of **drops** or streams (or the weather when such precipitation occurs)". As for the superscripts, they are semantic labels that specify the meaning of an LF.

It is, however, not entirely clear, how the decision is made as to when to specify a certain aspect and when not. For instance, in example (14b), the LF

Magn is only specified in terms of what it refers to – the drops, where clearly, the size of the drops is meant. In example (14c), two specifications are made: quantity and drops. It suggests, that size is considered the default attribute of *drops* that is automatically selected by an intensifying adjective, whereas quantity has to be additionally indicated.

Freund 'friend'
Magn^{temp}: *alt* 'old', *lebenslang* 'lifelong', *langjährig* 'longtime'
AntiMagn^{temp}: *neu* 'new'
$\text{Magn}_{[understand]}$: *eng* 'close', *dick* 'close', *nah* 'close', *gut* 'good'
$\text{Magn}_{[loyal]}$: *treu* 'faithful', *verlässlich* 'reliable'
Ver: *wahr* 'true', *echt* 'real'
AntiVer: *falsch* 'fake'
having great influence: *mächtig* 'powerful'
possessing material wealth: *reich* 'rich'

The modeling of the AN phrases with the noun *Freund* 'friend' as their base follows to a great extent the dictionary entry in Mel'čuk and Zholkovsky (1984, p. 293). Almost all of the collocates could be assigned to standard LFs. As in the previous examples, the LF Magn is specified: in *alter Freund*, the temporal aspect is intensified, whereas in *enger Freund* the understanding is. For the second meaning of the phrase *alter Freund* where the adjective refers to age, most likely, the LF-constant having lived for a long time should be applied.

Winter 'winter'
Magn$_{[cold]}$: *streng* 'severe', *bitterkalt* 'bitter cold', *eiskalt* 'ice cold'
AntiMagn$_{[cold]}$: *warm* 'warm'
Magn^{temp}: *lang* 'long', *ewig* 'endless'
difficult to bear + Magn$_{[cold]}$: *hart* 'harsh'
remote + Magn^{temp}: *tief* 'deep'
temperate: *mild* 'mild'
having moderately cool temperature: *kalt* 'cold', *kühl* 'cool'
characterized by a large quantity of water vapor: *feucht* 'humid'
characterized by a large quantity of water vapor + Anti: *trocken* 'dry'
causing sad feelings: *trist* 'dull, depressing'
previous: *letzt* 'last'

Three phrases with the base *Winter* 'winter' are modelled with the standard LF Magn specified with a subscript or a superscript. Several phrases were modeled by combining a non-standard LF with a standard one. For instance, the adjective *hart* 'harsh' expresses, in a way, a negative characteristic of winter. However, it is not a subjective evaluation, but rather an objective reflection of multiple aspects of winter such as very low temperature, weather in general, duration. Thus, the standard complex LF AntiBon cannot be used in this case and a non-standard LF difficult to bear combined with the LF Magn$_{[cold]}$ is used.

Droge 'drug'
dangerous and likely to cause addiction: *hart* 'hard'
dangerous and likely to cause addiction + Anti: *weich* 'soft', *leicht* 'soft', *sanft* 'soft'
allowed by official rules: *legal* 'legal'
allowed by official rules + Anti: *illegal* 'illegal'
produced internally: *körpereigen* 'endogenous'
causing danger or risk: *gefährlich* 'dangerous', *schädlich* 'harmful'
causing danger or risk + Anti: *harmlos* 'harmless'
causing danger or risk + Magn: *tödlich* 'lethal'
high in price + Anti: *billig* 'cheap'
capable of flowing: *flüssig* 'liquid'

All the phrases with the noun *Droge* 'drug' could only be modelled with non-standard LFs. In some cases they were mixed with standard ones `Anti` and `Magn`. Similarly to the phrases with *chocolate* and *shoes*, some of the pairs are highly lexicalized and are commonly accepted terms, for instance *harte/weiche Drogen*. The relation between such adjectives and their base is that of hyponymy: a hard drug *is_a_type_of* drug.

Team 'team'
`Bon`: *toll* 'excellent', *gut* 'good'
$\texttt{Magn}_{[performance]}$: *stark* 'strong', *schlagkräftig* 'powerful'
$\texttt{AntiMagn}_{[performance]}$: *schwach* 'weak'
`having knowledge or skill`: *erfahren* 'experienced'
`experienced through a lot of training together`: *eingespielt* 'coordinated, experienced'
`having existed for a long time` + `Anti`: *neu* 'new'
$\texttt{having existed for a long time}_{[members]}$ + `Anti`: *jung* 'young'
$\texttt{of certain origin}_{[members]}$: *spanisch* 'Spanish', *deutsch* 'German'
$\texttt{of the same or similar kind or nature}_{[members]}$: *homogen* 'homogenous'
$\texttt{encompassing everything belonging to it}_{[members]}$: *komplett* 'entire'
`consisting of many members` + `Anti`: *klein* 'small'

Three collocates of the noun *Team* 'team' are used as intensifiers (two positive and one negative) of the semantic component of a team, namely, its performance. Hence, they can be described with standard LFs `Magn` and `AntiMagn` with subscripts referring to that semantic component. The generic evaluative adjectives *toll* and *gut* can be modeled by a standard Lf `Bon`. The rest of the phrases are covered by non-standard LFs, sometimes combined with the standard one `Anti`.

Kurve 'curve'
`Magn`: *scharf* 'sharp', *steil* 'steep', *eng* 'narrow'
`AntiMagn`: *sanft* 'gentle', *weich* 'gentle'
`of certain color`: *blau* 'blue'
`with a sharply uneven outline`: *zackig* 'jagged'
`bulging outward`: *konvex* 'convex'

The collocates of the noun *Kurve* 'curve' that could be modelled with standard LFs refer to the angle at which a line curves. They are modelled as intensifiers by means of the standard LFs `Magn` and `AntiMagn`.

Objekt 'object'
Magn^{quant}: *zahlreich* 'numerous'
of certain worth + Magn: *wertvoll* 'valuable'
full of light: *hell* 'bright'
full of light + Anti: *obskur* 'obscure'
arousing or holding someone's attention: *interessant* 'interesting'
high in price: *teuer* 'expensive'
distinguished by special consecration: *sakral* 'sacred'
of high quality or Bon: *hochwertig* 'of high quality'
having three dimensions: *dreidimensional* 'three-dimensional'
serving an aesthetic purpose: *dekorativ* 'decorative'
located far away, remote: *entfernt* 'distant'
relatively small, taking up little space: *kompakt* 'compact'

Only one collocate of the noun *Objekt* 'object' could be modeled with a standard LF with a superscript. As in the case of *chocolate*, the adjective denoting high quality could either be modeled with a non-standard LF or with the LF Bon.

Geruch 'smell'
Magn: *streng* 'strong', *intensiv* 'intense'
Magn+AntiBon: *beißend* 'acrid', *stechend* 'pungent', *penetrant* 'pungent'
AntiBon: *unangenehm* 'unpleasant', *übel* 'foul', *widerlich* 'disgusting', *ekelhaft* 'disgusting'
fragrant + Bon: *süß* 'sweet'
smelling of mold + AntiBon: *modrig* 'musty', *muffig* 'musty'
smelling burnt: *brenzlig* 'burnt'
gone bad + AntiBon: *ranzig* 'rotten'

The noun *Geruch* 'smell', similarly to several examples above, can be intensified. Moreover, the collocates *beißend* 'acrid', *penetrant* 'pungent', and *stechend* 'pungent' express an intensified negative evaluation which can be modeled by means of a complex LF: combining Magn and AntiBon. A general unpleasant smell is covered by the complex standard LF AntiBon. The remaining collocates rather express a particular kind of smell, therefore, they can only be accommodated by a non-standard LF.

Preis 'price'
`Magn`: *hoch* 'high', *horrend* 'horrendous', *stolz* 'stiff'
`AntiMagn`: *niedrig* 'low', *erschwinglich* 'affordable', *vernünftig* 'reasonable', *moderat* 'moderate', *tief* 'low'
`becoming` + `AntiMagn`: *rückläufig* 'descreasing'
`Bon`: *toll* 'great', *attraktiv* 'attractive'
`that changes often` + `Anti`: *fest* 'fixed', *stabil* 'stable', *konstant* 'konstant'
`existing in the present time`: *heutig* 'of today'

Strafe 'punishment'
`Magn`: *drakonisch* 'draconian', *unmenschlich* 'inhumane', *scharf* 'harsh', *streng* 'strict', *hart* 'harsh' , *empfindlich* 'severe'*saftig* 'stiff', *deftig* 'severe', *hoch* 'high', *schwer* 'heavy', *schlimm* 'severe'
`AntiMagn`: *niedrig* 'weak', *mild* 'mild'
$\texttt{Magn}^{temp}$: *lebenslang* 'lifelong'
`representing something else`: *symbolisch* 'symbolic'

Both nouns, *Preis* 'price' and *Strafe* 'punishment', are mainly described in terms of their intensity or rather scale, which is expressed as the simple LF `Magn` or the complex one `AntiMagn`.

Thema 'issue/topic'
$\texttt{Magn}^{importance}$: *groß* 'big/major', *heiß* 'hot', *beherrschend* 'dominant', *wichtig* 'important', *zentral* 'central'
$\texttt{Magn}^{temp}$: *ewig* 'eternal'
$\texttt{AntiBon}_{[discussion]}$: *unbequem* 'uncomfortable', *leidig* 'vexed', *heikel* 'delicate', *sensibel* 'sensitive', *brisant* 'controversial'
`very complex`: *sperrig* 'unwieldy'
`requiring immediate action`: *drängend* 'pressing', *brennend* 'urgent'
`existing in the present time`: *aktuell* 'current'
`arousing or holding someone's attention`: *interessant* 'interesting'
`restricted to a particular locality or neighborhood`: *lokal* 'local'

The nominal base *Thema* 'issue/topic' and most of its modifiers could be successfully modeled with standard LFs. However, specifications had to be made for each relation. In contrast to the previous two nouns denoting scales of some sort, the meaning of *topic* does not have a salient component that is automatically enhanced in combination with an intensifier. Similarly, the subjective evaluation, `Bon`, had to be specified as $\texttt{AntiBon}_{[discussion]}$: it is the discussion of a topic that may make someone feel 'uncomfortable' or speak 'delicately', not the *topic* itself.

Ähnlichkeit 'similarity'
`Magn`: *groß* 'strong', *stark* 'strong'
`AntiMagn`: *leise* 'slight'
$\texttt{AntiMagn}^{temp}$: *flüchtig* 'fleeting'
`highly unusual or remarkable` + `Magn`: *unheimlich* 'uncanny'
`causing surprise` + `Magn`: *frappierend* 'striking', *frappant* 'striking', *erstaunlich* 'surprising'
`seen from the outside`: *äußerlich* 'external'
`in an accurate manner`: *exakt* 'precise'
`in an accurate manner` + `Anti`: *vage* 'vague'
`determined at random`: *beliebig* 'arbitrary'
`that changes often`: *variabel* 'variable'

Anzahl 'amount'
`Magn`: *stattlich* 'stately', *groß* 'large', *hoch* 'high', *schier* 'sheer', *beträchtlich* 'large', *riesig* 'huge', *erklecklich* 'considerable', *bedeutend* 'substantial'
`AntiMagn`: *bescheiden* 'modest', *gering* 'small', *klein* 'small', *überschaubar* 'manageable'

Both *Ähnlichkeit* 'similarity' and *Anzahl* 'amount' have a dominant *scale* meaning component. This component is targeted by most of their collocates and is described in terms of the LFs `Magn` and `AntiMagn`.

Ziel 'goal'
Magn: *ober* 'ultimate', *groß* 'great,big'
that can be achieved with a lot of ambition: *hochgesteckt* 'ambitious', *hoch* 'high', *ehrgeizig* 'ambitious'
chosen by someone themselves: *selbstgesteckt* 'self-imposed'
accurately stated or described: *klar* 'clear'
characterized by determination: *fest* 'firm'
of foremost importance or value: *primär* 'primary'
serving as an essential component: *zentral* 'central', *wichtig* 'important'
extending over a relatively long time: *langfristig* 'long-term'
extending over a relatively long time + Anti: *kurzfristig* 'short-term'
sublime, awe-inspiring: *hehr* 'noble'
pursued by two or more persons: *gemeinsam* 'joint'

The noun *Ziel* 'goal' has a large number of adjectival modifiers that cannot be modeled by means of standard LFs. The main reason for that is the weakness of scale component in the meaning if the noun - it is still present and thus can be selected by a few collocates such as *ober* 'ultimate' and *groß* 'great, big'. However, there is a large number of other meaning aspects expressed by adjectival collocates that are described by non-standard LFs.

Angst 'fear'
$\text{Magn}_{[scared]}$: *groß* 'strong/big', *tief* 'deep', *höllisch* 'hellish', *furchtbar* 'tremendous'
Magn^{temp}: *ständig* 'constant', *ewig* 'eternal', *stet* 'constant'
nothing but: *nackt* 'pure', *pur* 'pure', *blank* 'sheer'
such that its reason is not clear: *diffus* 'vague', *dumpf* 'vague'
increased to abnormal: *krankhaft* 'pathological'

The modeling of the phrases with the nominal base *Angst* 'fear' partially follows the dictionary entry in Mel'čuk and Zholkovsky (1984, p.806). This concerns mainly the standard LFs. Additionally, a number of non-standard LF has to be introduced to describe the aspects of the meaning of *fear* not addressed in the dictionary entry. It indicates that such a complex emotion as *fear* has different meaning aspects often modified by adjectives that cannot

be simply reduced to a scale or an evaluation. Thus only standard LFs do not suffice.

Interesse 'interest'
`Magn`: *hoch* 'strong', *groß* 'strong/big', *stark* 'strong', *massiv* 'massive'
`AntiMagn`: *gering* 'little'
$\texttt{Magn}_{[involvement]}$: *lebhaft* 'lively', *rege* 'active'
$\texttt{Magn}_{[importance]}$: *vital* 'vital', *elementar* 'fundamental'
$\texttt{Magn}_{[domain]}$: *breit* 'broad'
`Ver`: *echt* 'genuine', *wahr* 'genuine'
`belonging or peculiar to a specified person`: *ureigen* 'own'
`contradictory, in disagreement`: *widerstreitend* 'conflicting'
`characterized by a material value`: *handfest* 'tangible'

For the noun *Interesse* 'interest', about half of the modifiers could be modelled with standard LFs, most of them being an intensifier of some sort. The non-specified `Magn` LF describes the intensity of *interest* as an emotion, whereas the specified ones refer to the degree of other aspects. For instance, *a vital interest* is an essential interest, therefore the degree of `importance` is high – hence the specified LF $\texttt{Magn}_{[importance]}$. However, for a number of meanings no standard LFs could be applied and very specific non-standard LFs had to be introduced.

Conclusion

This section presented an overview of the Meaning-Text Theory and its key concepts relevant for modeling the relations between the constituents of AN phrases. Relying on a selected sample of semantically diverse data, the applicability of LFs has been empirically investigated.

First of all, the study has shown that the degree of concreteness of the base noun is important. For the more abstract nouns the collocate can be successfully modeled in terms of Lexical Functions. Modeling of AN phrases with concrete nouns by means of standard LFs is challenging. The suggested non-standard LFs have a huge disadvantage of being extremely fine-grained. Thus, as Jousse (2007) suggests, a way of generalizing the non-standard LFs is required.

Secondly, standard LFs require further specifications in many examples. This approach has been adopted from Mel'čuk and his colleagues as presented in the *Explanatory Combinatorial Dictionary of Modern Russian* (Mel'čuk and Zholkovsky, 1984). The assignment of the specifications such as $\texttt{Magn}_{[understand]}$ or $\texttt{Magn}_{[drops]}$ is an arbitrary decision of the lexicographer that requires further formalization.

Finally, even the most abstract nouns have aspects of meaning that are not reflected in the inventory of standard adjectival LFs. Moreover, the covered aspects (intensification, objective and subjective evaluation) are semantically very similar to each other and it is not yet clear whether it is feasible to keep them as separate categories in the classification scheme.

The next section is devoted to the theory of Generative Lexicon as it takes special interest in the semantics of nouns and in the compositional processes that are activated when a nominal base is modified.

4.3 The Generative Lexicon

According to the theory of the Generative Lexicon (Pustejovsky, 1991), a lot of the generative power of natural languages stems from the lexicon. The assumptions about generativity of lexicon within the framework of the Generative Lexicon are the following: (1) there is a finite number of meaning elements that create an infinite number of word senses via specific lexical mechanisms; (2) syntax and semantics are incorporated in the lexicon (Pustejovsky and Batiukova, 2019). Thus, the *decomposition* of meaning is performed generatively: there is a fixed number of "*generative devices* that construct semantic expressions" (Pustejovsky, 1991, p.417).

It is important to keep in mind that "the lexicon is not just verbs" (Pustejovsky, 1991, p.410). As this thesis is primarily interested in the semantics of nominals, this sections turns to the theory of the Generative Lexicon for possible solutions to the problems reported in the previous section. More specifically, the unresolved issues are concerned with outlining the essential aspects of nominal meaning at an optimal level of granularity. Such meaning aspects should be defined for nouns from different semantic classes: both concrete and abstract ones.

The Generative Lexicon (GL) distinguishes four aspects of meaning representation: Argument Structure, Event Structure, Qualia Structure, Inheritance Structure.

Argument Structure maps a word's meaning to syntax by specifying the arguments of a predicate. Basically, this is the information about 'who does what to whom with what?' For instance, the verb *love* requires both a subject and an object, whereas the verb *drink* requires only one argument – a subject.

Event Structure represents the temporal and structural aspects of events. Pustejovsky (1991) distinguishes three types of events in this regard: states, processes, and transitions.

Qualia Structure. As Pustejovsky (1991) emphasizes, the lexicon does not amount only to verbs, and it is equally important to consider the semantic structure of their arguments. The aspects that characterize the meaning of nominals, similarly to the argument structure of verbs, are called in GL *qualia structure* and are made up of four *qualia roles*. Qualia structure will be discussed in detail in the next subsection (4.3.1).

Inheritance structure represents the relations between lexical structures globally within the lexicon.

Based on these levels of representation, it is possible to account for the flexibility of meaning in different contexts via a set of generative devices. They include *type coercion, selective binding*, and *co-composition*. In **type coercion**, the type of the argument is changed depending on the function. Consider the following example:

(15) The children enjoy the ice-cream.

The verb *enjoy* typically selects events or processes as its arguments, however, a *ice-cream* is an artifact. In this case, the verb coerces the noun into a different semantic type, and thus example (15) can be interpreted as "The children enjoyed *eating* the ice-cream". Typical verbs for this coercion type are aspectual verbs such as *begin* and *finish*, as in *begin the cake* and *finish the coffee* that can be interpreted as 'begin eating/baking a cake', 'finish drinking the coffee' (Pustejovsky and Batiukova, 2019).

The second generative mechanism, **selective binding**, is best illustrated with polysemous adjectives: their interpretation is dependent on the internal semantic structure of the noun they modify (Pustejovsky, 1995). For example:

(16) a. a fast boat
b. a fast typist

In example (16a), the adjective is used in the sense "moving quickly", whereas in example (16b), it does not mean "a typist who moves quickly". Rather, the adjective refers to the manner of performing an action – typing – "a typist who can type quickly" (Pustejovsky, 1991, p. 411). Similarly, evaluative adjectives such as *good* or *bad* receive a different interpretation depending on their head. For instance, *a good TV* is a TV that it functions well, whereas in *a good meal*, the adjective could be equaled to *tasty*. In all the above given examples, the modifier selectively binds one of the aspects of the head, i.e. one of the qualia roles.

Finally, **co-composition** is the mechanism that is activated when a predicate's semantic type and the qualia structure of its argument are identical (Pustejovsky and Batiukova, 2019). In a way, the basic meaning of the predicate is part of the semantics of the argument. Consider the two examples from (Pustejovsky, 1991):

(17) a. John baked the cake.
b. John baked the potato.

According to GL, in the qualia structure of the noun *cake* there is a reference to the verb *bake*: it is reflected in the AGENTIVE quale that provides information about the origin or bringing about of the object. Thus, in example (17a), the verb *bake* acquires a creation reading. A *potato*, on the other hand, is not an artifact and its internal semantic structure does not have a reference to being created through any action. In example (17b), the verb *bake* only expresses a change of state. In that way, the verb *bake* is not polysemous, but rather acquires different interpretation based on its object: its complement *cake* specifies its meaning as creation via qualia structure. In the next subsection, the theory of qualia is presented in detail.

4.3.1 Qualia Roles

The idea of qualia was inspired by Aristotle's theory of *aitia*: to understand a thing or to have a knowledge about something in this world is only possible when one grasps the *aitia* of this thing (Moravcsik, 1991), or, in other words, its *cause*. According to Aristotle's theory, there are four *aitiai*:

- **Material/constituency**: a classic example for this *aitia* is *the bronze of a statue*. However, Moravcsik (1991, p.44) notes that Aristotelian notion of *matter* is wider than the modern interpretation of the notion *material*. It can also refer to abstract concepts, for instance, premises in an argument.
- **Form**: for instance, the shape of a statue.
- **Function/goal**: for instance, the intended purpose of a statue.
- **Efficient cause**: for instance, the creator of the statue or the act of creating a statue.

Similarly to *aitiai*, qualia structure in GL has four components representing the fundamental aspects of a word's meaning. All categories have a qualia structure, but not all qualia roles are necessarily realized for a give word (Pustejovsky, 1991). The four qualia roles are the following:

- FORMAL: distinguishes an entity within a larger domain. It may express properties such as shape, dimensionality, color, position, orientation, magnitude. It is is defined through the *is_a* relation: for example, a *knife* is a *tool*.
- CONSTITUTIVE: constitution of an object, the relation between an object and its parts. It includes properties such as weight, material, parts and components of an object. For example, *handle* is a part of a *knife*.
- AGENTIVE: the origin of an object, including information about the 'bringing about', for instance artifact, natural kind, creator, causal chain. For example, a *knife* is an artifact.
- TELIC: purpose or use of an object. It expresses the built-in function or aim. For example, *cutting* is the purpose of a *knife*.

Pustejovsky (1995) emphasizes that qualia roles are not simply a number of features of lexical items. Qualia structure does not just describe the meaning of a word, it provides information about its use in language and about semantic constraints applied to a particular lexical unit. The template of qualia structure and the generative mechanisms such as coercion, selective binding, and co-composition are means of modeling the meanings of words. As previously mentioned, all lexical items have a qualia structure, but not necesserily all four qualia roles are expressed. The FORMAL role is expressed for any word because any lexical item can be defined in terms

of IS_A relation. This role cannot have "none" value or, in other words, be *underspecified*.

Compare the qualia structures (QS) of two nouns: *knife* – an artifact, and *horse* – a natural kind noun. F stands for FORMAL, C stands for CONSTITUTIVE, T stands for TELIC, and A stands for AGENTIVE:

knife

$$\text{QS} = \begin{bmatrix} F = physical\_object \\ C = blade \\ T = cut \\ A = create \end{bmatrix}$$

tree

$$\text{QS} = \begin{bmatrix} F = physical\_object \\ C = roots \\ T = none \\ A = none \end{bmatrix}$$

In the case of the noun *tree*, the FORMAL and CONSTITUTIVE roles are present in the qualia structure, whereas the AGENTIVE and TELIC roles are underspecified (Pustejovsky and Batiukova, 2019). This behaviour is typical for the nouns of natural kind in contrast to artifacts because artifacts are created and they have a function. This difference is manifested in the combinatorial behaviour of artifacts and natural types. For instance, combination with evaluative adjectives such as *good* or *efficient* is more easily interpretable for artifacts because these adjectives specify the TELIC role: *a good knife* is a knife that cuts well. Interpreting a phrase such as *a good tree* out of context is not intuitive as it is not clear what aspect of the tree is evaluated because *a tree* does not have a TELIC dimension in its qualia structure.

There are also similar combinability patterns that hold true for the nouns expressing professions: nouns such as *doctor* or *teacher* have an inherent functional dimension, a reference to a typical activity, such as *to heal* or *to teach*. This information is captured in the qualia structure where the TELIC role is always specified (Lenci et al., 2000):

doctor

$$\text{QS} = \begin{bmatrix} F = human\_being \\ C = none \\ T = heal \\ A = none \end{bmatrix}$$

teacher

$$\text{QS} = \begin{bmatrix} F = human\_being \\ C = none \\ T = teach \\ A = none \end{bmatrix}$$

Thus evaluative adjectives address the TELIC role when combined with nouns denoting professions: *a good teacher* is a teacher who is good at teaching and *a good doctor* is a doctor who is good at healing.

For the same noun, different adjectives modify various aspects by selecting different qualia roles from the noun's qualia structure. Consider the example from Pustejovsky (1991):

(18) a. a bright bulb
b. an opaque bulb

The two adjectives describe a certain aspect of the noun *bulb*: *bright* in example (18a) refers to a qualia-derived event addressing its TELIC role - *illumination*. The second adjective refers to the FORMAL role of the noun (Pustejovsky, 1991).

Extended Qualia in the SIMPLE model

Pustejovsky (1991) emphasizes that the theory of Generative Lexicon can be used not only in theoretical linguistics, but in computational linguistics as well. A GL-based framework has been implemented in a large-scale project in computational lexical semantics SIMPLE (Semantic Information for Multipurpose PLurilingual Lexica) (Bel et al., 2000). It comprises semantic lexicons for twelve languages and provides a model that encodes uniform and consistent semantic types, relations between semantic units, semantic features, and domain features (Lenci et al., 2000).

The fundamental aspects of meaning are represented in SIMPLE in terms of the *extended qualia roles* as relations that hold between semantic units (Ruimy et al., 2001). For instance, the FORMAL role as the relation *is_a* and CONSTITUTIVE role as the relation *is_a_part* or *has_a_part*. For each of the four qualia a subset of relations has been specified which offers a more fine-grained approach to a word's meaning. Consider the examples from Lenci et al. (2000):

(19) a. an old swimmer
b. an old fish

There are two possible interpretations of sentence (19a): (1) a person who is old and who swims; (2) a person who has been swimming for a long time. In contrast, example (19b) can have only one interpretation: a fish which is old. The two nouns thus exhibit a different linguistic behaviour.

As discussed above, the nouns denoting professions and activities have an inherent TELIC role that is expressed as the function of the person. In the case of *swimmer*, the function is *to swim* which makes the second interpretation plausible. Thus, in *old swimmer*, the adjective either selects the FORMAL or the TELIC role in the qualia structure of the noun. In the SIMPLE model, the TELIC for the noun *swimmer* is expressed as the relation *Is_the_activity_of*: to swim is the activity of swimmer. Note that swimming is also a typical activity of a *fish*. However, the noun *fish* exhibits a different linguistic behaviour: there is only one interpretation for example (19b). It suggests that the swimming activity for a fish belongs to a different qualia aspect as it is obviously not within the TELIC dimension. The activity of swimming rather has to do with the way a fish is built and thus with its constitution (Lenci et al., 2000). This information is encoded in SIMPLE in the CONSTITUTIVE quale as the relation *Constitutive_activity*.

(20) *fish*
F: *Isa*(animal)
C: *Constitutive_activity* (swim)

swimmer
F = *Isa*(person)
T = *Is_the_activity_of*(swim)

Formal	Constitutive	Agentive	Telic
Isa	*Has_as_member*	*Result_of*	*Purpose*
Antonym	*Has_as_part*	*Agentive_prog*	*Object_of_the_activity*
	Location	*Artifactual_agentive*	*Is_the_activity_of*
	Property	*Agentive_cause*	*Instrumental*
	Instrument	*Agentive_experience*	*Is_the_ability_of*
	Relates	*Caused_by*	*Is_the_habit_of*
	Resulting_state	*Source*	*Used_for*
	Is_a_follower_of	*Derived_from*	*Used_by*
	Made_of	*Created_by*	*Used_against*
	Is_in		*Used_as*
	Lives_in		
	Constitutive_activity		
	Produces		
	Produced_by		
	Property_of		
	Concerns		
	Contains		
	Quantifies		
	Measured_by		
	Related_to		
	Successor_of		
	Has_as_effect		
	Typical_of		
	Causes		

TABLE 4.2: Qualia relations used in the SIMPLE model (Lenci et al., 2000)

The difference between the two nouns as encoded in SIMPLE in terms of qualia roles is shown in example (20). Table 4.2 lists all the relations adopted in SIMPLE that represent the extended qualia structure.

Depending on the semantic type of the noun, the interpretation of qualia roles may differ. Ruimy et al. (2001) provide an overview of qualia shifts for nouns of types `human` and `abstract`:

qualia	human	abstract
CONSTITUTIVE	membership	connotation, domain
AGENTIVE	kind of agent	social and conventional source
TELIC	social roles	frequently underspecified function

Moreover, Ruimy et al. (2001) emphasize that specifying the qualia structure of non-concrete nouns poses a challenge.

Qualia for abstract nouns

An observation similar to that made in Ruimy et al. (2001) is also found in Warren (2003): namely, that qualia roles reflect the semantics of *first-order entities*, but can only indirectly be applied to non-concrete nouns or,

following the terminology adopted from Lyons (1977), *second-, and third-order entities*. Consider the definitions of the three categories given in Lyons (1977, pp.442–447):

- **First-order entities**: physical objects (persons, animals, things). "They are relatively constant as to their perceptual properties; they are located in a three-dimensional space; they are publicly observable"
- **Second-order entities**: "events, processes, states-of-affairs, etc., which are located in time" and which "are said to occur or to take place rather than to exist."
- **Third-order entities**: "abstract entities such as propositions which are outside space and time".

Indeed, the examples given in most works concerned with the qualia structure belong to the category of first-order entities. In order to fill that gap, Ruimy et al. (2001) investigated the qualia structure of nouns that belong to the following semantic types: **abstract entities** such as *spring, medicine, charity*; **properties** such as *intelligence, power, color, shape*; **representations** such as *language, number*. The analysis shows that the meaning of such non-concrete nouns is in general more complex than that of the concrete ones, but, at the same time, there is little information about their semantics. Modeling their meaning poses a challenge for the theory of GL, but it proved to be feasible. Ruimy et al. (2001)'s work on normalizing the meaning of abstract entities provides a number of qualia relations presented in Table 4.3. Apart from relations, there is a number of features that are expressed in terms of "yes", "no", and "underspecified" values. For instance, for the noun *season*, the value of the feature *iterative* is "yes". The authors point out that many of the introduced relations are more general than the ones for concrete nouns. However, they come to the conclusion that qualia roles are adequate means of representing the semantics of abstract entities. Moreover, qualia roles can be successfully applied in the task of specifying the genus term to abstract entities. For instance, it is very hard to find a genus term for such nouns as *aim*, *origin*, and *part*, but they are basically instantiations of qualia roles: *aim* for TELIC, *origin* for AGENTIVE, *part* for CONSTITUTIVE (Ruimy et al., 2001).

As discussed previously, abstract nouns tend to form more collocations than concrete ones and, and thus the findings by Ruimy et al. (2001) are promising. However, the question still remains open whether qualia roles can be implemented to describe the relations in collocations at an optimal level of granularity. The next subsection presents an empirical study where this issue is investigated in detail in the same fashion as in section 4.2.2.

4.3.2 Case study: Qualia Roles

The idea of relying on qualia structure to determine semantic relations between the constituents of multi-word expressions is not new. This approach has been explored in research on nominal compounds, more specifically,

Quale	Relations (r) and features (f)
CONSTITUTIVE	r: *concerns, has_as_member, is_in, has_as_part, is_a_part_of, measured_by, property_of, quantifies, related_to, successor_of* f: *connotation (pos, neg, underspec), dimension (1, 2, 3), iterative (yes, no, underspec), scalar (yes, no), punctual (yes, no, underspec)*
AGENTIVE	r: *agentive, result_of, source*
TELIC	r: *telic, indirect_telic, direct_telic, used_by, used_for*

TABLE 4.3: Qualia relations and features for non-concrete nouns from Ruimy et al. (2001, pp. 360-361).

in classifying the relations that hold between the head and its modifier in a compound. Sorokin et al. (2015) present a hybrid annotation scheme to model such relations in compounds. The scheme combines paraphrases with prepositions with semantic properties inspired by *qualia roles*. The coarse-grained qualia roles are further differentiated into more specific relations. For instance, the CONSTITUTIVE role is represented by following properties (Sorokin et al., 2015):

- `component` as in *Chlorwasser* 'chlorine water'
- `ingredient` as in *Gurkensalat* 'cucumber salad'
- `material` as in *Holzhaus* 'wooden house'
- `part` as in *Siegelring* 'seal ring'
- `part`$^{-1}$ as in *Kinderhand* 'child's hand'

The results of the IAA study indicate that their approach is reliable with the Kappa score (Cohen, 1960) of 0.74. However, Sorokin et al. (2015) study only German compounds with concrete nouns as their head. Therefore, it remains unclear whether the same approach would work for a more semantically diverse sample of data including abstract nouns. Moreover, relations in nominal compounds are different from those in adjective-noun phrases with qualitative adjectives: in noun-noun compounds, the relationship is between two objects, while a qualitative adjective describes a property of the noun it modifies.

This subsection presents a case study on modeling the previously described sample of adjective-noun phrases (section 4.1) using *qualia roles*. In order to facilitate the modeling, the more specific qualia relations proposed in Lenci et al. (2000) (table 4.2), Ruimy et al. (2001) (table 4.3), and in Sorokin et al. (2015) and those mentioned in the definitions of each role will be utilized when applicable. If the specifications belong to the same qualia dimension, they will be listed for it in round brackets and separated from each

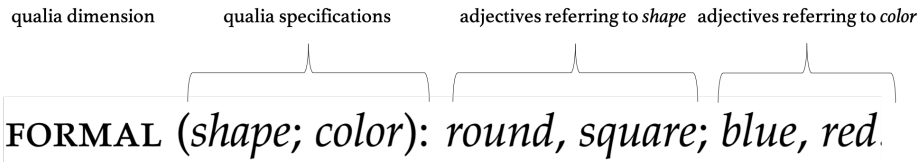


FIGURE 4.2: Notation used in this chapter for modeling adjectival collocates with qualia roles and their specifications.

other by a semicolon. The corresponding adjectives will also be separated by a semicolon (see figure 4.2 for illustration).

The collocates that could not be assigned to a qualia role are defined in the same manner as in the previous case study in section 4.2.2.

Schokolade 'chocolate'
FORMAL (*is_a*; *taste*; *consistency*): *schwarz* 'dark', *dunkel* 'dark', *weiß* 'white', *heiß* 'hot', *flüssig* 'liquid'; *süß* 'sweet'; *knackig* 'crunchy'
`of high quality`: *edel* 'premium, fine', *hochwertig* 'of high quality'
`having a pleasant taste`: *lecker* 'tasty'
`high in price`: *teuer* 'expensive'

Schuh 'shoe'
FORMAL (*is_a*; *color*; *shape*): *hochhackig* 'high-heeled', *flach* 'flat'; *spitz* 'pointy'
CONSTITUTIVE (*material*; *has_as_color*): *seiden* 'silk'; *braun* 'brown', *schwarz* 'black', *rot* 'red', *weiß* 'white', *rosa* 'pink', *blau* 'blue'
TELIC: *bequem* 'comfortable'
`soiled`: *schmutzig* 'dirty'
`covered or soaked with a liquid`: *naß* 'wet'
`high in price`: *teuer* 'expensive'
`refined and tasteful in style`: *elegant* 'elegant'

Many of the co-occurrences with *chocolate* denote a *type of* chocolate, i.e. the modifier does not refer to a particular property of chocolate, but rather distinguishes it within the domain of different kinds of solid chocolate. Thus, these phrases are in an *is_as* relation to *chocolate*. However, these modifiers could also be interpreted as referring to the CONSTITUTIVE role as they indicate what ingredients the chocolate has: *white chocolate* is made of cocoa butter, milk, and sugar, whereas *dark chocolate* contains cocoa solids and cocoa butter and no milk. As the second interpretation is less straightforward than

the first one, these phrases have been assigned to the FORMAL role. Apart from that, there are two other specifications of the FORMAL dimension: *taste* and *consistency* that distinguish chocolate within the domain.

In the same manner, two collocates of the noun *shoe* have been assigned to the specification *is_a* of the qualia role FORMAL. Both *flat* and *high-heeled* distinguish the corresponding types of shoes within the domain shoes.

A few rather abstract modifiers of both *Schokolade* and *Schuh* could not be assigned to any qualia dimension.

Fels 'cliff'
FORMAL (*surface*; *shape and surface*, ?): *glatt* 'smooth'; *schroff* 'craggy'; *steil* 'steep', *hart* 'hard'
CONSTITUTIVE (*has_as_part*; *has_as_color*): *nackt* 'bare', *blank* 'bare'; *schwarz* 'black', *rötlich* 'reddish', *grau* 'grey'
`covered or soaked with a liquid`: *naß* 'wet'

For the noun *Fels*, some specifications could not be made at this stage. In such cases, the adjectives are grouped together and the specification is expressed as a question mark.

Hund 'dog'
FORMAL (*size*; *appearance*; *age*): *riesig* 'huge', *klein* 'small'; *dick* 'fat'; *jung* 'young'
`pitiful`: *arm* 'poor'
`evoking a pleasant feeling`: *süß* 'sweet', *niedlich* 'cute'
`having certain character, temper`: *scharf* 'vicious', *bissig* 'biting', *treu* 'loyal', *aggressiv* 'agressive'
`(un)intelligent`: *klug* 'smart', *dumm* 'stupid'
`infected with rabies`: *tollwütig* 'rabid'

Although the noun *Hund* 'dog' is very concrete, it is a living being and has abstract semantic components such as character and intelligence. A dog is capable of causing emotions (*arm* 'poor') and can be described in terms of a subjective evaluation (*süß* 'sweet'). It is unclear, which qualia dimensions can accommodate those meanings.

Wald 'forest'
FORMAL (*is_a*; *size*): *immergrün* 'evergreen'; *riesig* 'huge', *endlos* 'endless'
CONSTITUTIVE (*location*): *tief* 'deep', *nah* 'close'
CONSTITUTIVE/FORMAL (*has_as_part*/*is_a*): *licht* 'sparse', *dicht* 'dense', *feucht* 'humid', *trocken* 'dry'
`in a natural state`: *wild* 'wild', *unberührt* 'untouched'
`characterized by grandeur`: *herrlich* 'magnificent'

Relying on the qualia relations listed in table 4.2, some adjectives have been classified as referring to the CONSTITUTIVE role of the noun *forest*. The adjective *tief* 'deep' indeed points at a location in the forest, thus, in a way, it reflects the relation between a forest and its parts, in this case a location within a forest. The interpretation of the adjectives *licht* and *dicht* is less straightforward. For *lichter forest,* possible interpretation is *consisting of scattered elements*, thus the adjective *licht* selects the CONSTITUTIVE role. However, based on the definition of the FORMAL role in Pustejovsky (1991) (see section 4.3.1) , the adjective may as well be assigned to the FORMAL role, as it distinguishes a *sparse forest* from a *thick forest*.

Baum 'tree'
FORMAL (*is_a*; *size*; *height*; *age*; ?): *immergrün* 'evergreen'; *mächtig* 'mighty, massive', *riesig* 'huge'; *hoch* 'high'; *alt* 'old', *jung* 'young'; *tot* 'dead', *krank* 'sick', *gesund* 'healthy'
CONSTITUTIVE (*has_as_part*): *kahl* 'leafless'
`distinguished by special consecration`: *heilig* 'sacred'

Bein 'leg'
FORMAL (*appearance*; *length*; ?): *staksig* 'spindly', *dick* 'fat', *schlank* 'slim', *dünn* 'thin'; *kurz* 'short', *lang* 'long'; *gesund* 'healthy', *wacklig* 'shaky', *stämmig* 'sturdy', *schwach* 'weak'
`in need of rest, without energy`: *müde* 'tired', *schwer* 'heavy'
`completely unclothed`: *nackt* 'naked'

Most of the collocates of the nouns *Baum* 'tree' and *Bein* 'leg' belong to the FORMAL qualia dimension. A number of adjectives could not be assigned to specifications of the FORMAL role. For some, no qualia dimension was

found fitting. For instance, *schwer(e) Bein(e)* 'heavy leg(s)' could not be accommodated by any qualia roles. It might be interpreted as belonging to the TELIC dimension - *heavy* is synonymous to *tired* in this case. The tired state is the result of some *activity_of* legs, which is a TELIC relation in the SIMPLE framework (Lenci et al., 2000).

Regen 'rain'
FORMAL (*size (of drops)*; *temperature*): *fein* 'fine'; *warm* 'warm', *kalt* 'cold', *kühl* 'cool'
FORMAL/CONSTITUTIVE (*is_a/part_of*): *sauer* 'acid', *spärlich* 'sparse'
CONSTITUTIVE (*degree;*): *sintflutartig* 'torrential', *stark* 'heavy', *kräftig* 'heavy', *heftig* 'heavy', *leicht* 'light'
`flowing strongly and in large quantities`: *strömend* 'pouring'
`tiresomely long, without a foreseeable end`: *ewig* 'endless', *endlos* 'endless'

The phrase *saurer Regen* 'acid rain' has two possible interpretations:(1) *acid rain* is a type of rain/precipitation – FORMAL role; (2) *acid* is a part of *rain* – CONSTITUTIVE role.

There is no explicit qualia relation or specification for the meaning of *degree* or *intensification*. The semantically closest relation from the inventory of the extended qualia roles (Lenci et al., 2000) is the relation *quantifies* the belongs to the CONSTITUTIVE role. In the SIMPLE ontology, this relation is used, for instance, in the entry of the word *grado* 'level': "*quantifies* (<grado_1>, <Usem>:[Entity]) //except for concrete entities//". Another possible interpretation is listing *degree* under the *property* relation, analogously to *has_as_color – has_as_degree*, which is also a specification of the CONSTITUTIVE role. Therefore, here and further, this relation will be listed as an extension of the CONSTITUTIVE role.

Freund 'friend'
TELIC (`for a certain amount of time`): *alt* 'old', *lebenslang* 'lifelong', *langjährig* 'longtime', *neu* 'new'
TELIC (`which you can(not) rely on`): *wahr* 'true', *echt* 'real', *falsch* 'fake', *treu* 'faithful', *verlässlich* 'reliable'
TELIC (`intimate, dear`): *eng* 'close', *dick* 'close', *nah* 'close', *gut* 'good'
`having great influence`: *mächtig* 'powerful'
`possessing material wealth`: *reich* 'rich'

According to Pustejovsky and Batiukova (2019, p.369), the adjective *old* in combination with the noun *friend* addresses either the FORMAL – the age of

the friend, or the TELIC role – the duration of friendship. In the same manner, the other groups of adjectives (except *mächtig* and *reich*) do not describe the referent of *friend*, but rather *friendship*. Therefore, they are assigned to the TELIC dimension of *friend*. A further specification is, however, needed, as the three groups are semantically different. Since none of the available qualia extensions cover the three required meanings, definitions are provided instead.

Winter 'winter'
FORMAL (*temperature*; ?): *streng* 'severe', *bitterkalt* 'bitter cold', *eiskalt* 'ice cold', *warm* 'warm', *kalt* 'cold', *kühl* 'cool'; *feucht* 'humid', *trocken* 'dry'
`difficult/easy to bear`: *hart* 'harsh', *mild* 'mild'
`in the middle of` (winter): *tief* 'deep'
`causing sad feelings`: *trist* 'dull, depressing'
`previous`: *letzt* 'last'
`long, without a foreseeable end`: *lang* 'long', *ewig* 'endless'

Droge 'drug'
FORMAL (*is_a*): *hart* 'hard', *weich* 'soft', *leicht* 'soft', *sanft* 'soft', *flüssig* 'liquid'
AGENTIVE (*origin*): *körpereigen* 'endogenous'
CONSTITUTIVE (*causes*) : *gefährlich* 'dangerous', *schädlich* 'harmful', *harmlos* 'harmless', *tödlich* 'lethal'
`(not) allowed by official rules`: *legal* 'legal', *illegal* 'illegal'
`low in price`: *billig* 'cheap'

Half of the listed collocates of *Winter* 'winter' could be assigned to qualia dimensions. A few pairs with *Droge* denote a subgroup of *drugs* and were, therefore, assigned to the FORMAL role. A drug has either a natural or an artificial origin, therefore, its AGENTIVE dimension can be targeted. Apart from that, a drug has an effect which according to the SIMPLE model (Lenci et al., 2000) belongs to the CONSTITUTIVE qualia dimension.

Team 'team'
FORMAL (*size*; *age*): *klein* 'small'; *neu* 'new'
TELIC (*function*): *stark* 'strong', *schlagkräftig* 'powerful', *schwach* 'weak'
CONSTITUTIVE (*has_as_members*): *homogen* 'homogenous'
CONSTITUTIVE (*has_as_members* (FORMAL (*age*; *nationality*)): *jung* 'young'; *spanisch* 'Spanish', *deutsch* 'German'
`experienced through a lot of training together`: *eingespielt* 'coordinated, experienced'
`having knowledge or skill`: *erfahren* 'experienced'
`constituting all of its members`: *komplett* 'entire'
`evaluating in a positive manner`: *toll* 'excellent', *gut* 'good'

A team is a group people who work on some task together, therefore, the performance of a team is its *function* which belongs to the TELIC qualia dimension. Members of a team belong to its CONSTITUTIVE dimension, but their characteristics such as *age* – to the FORMAL dimension of each member. In a way, the FORMAL qualia is embedded into the CONSTITUTIVE one. It has not been reported before whether it is possible to embed one qualia dimension into another.

Kurve 'curve'
FORMAL (*shape*): *scharf* 'sharp', *sanft* 'gentle', *steil* 'steep', *eng* 'narrow', *weich* 'gentle', *zackig* 'jagged', *konvex* 'convex'
CONSTITUTIVE (*has_as_color*): *blau* 'blue'

Objekt 'object'
CONSTITUTIVE (?): *zahlreich* 'numerous'
TELIC (*purpose*; ?): *dekorativ* 'decorative'; *wertvoll* 'valuable'
FORMAL (*is_a*; *size*; *brightness*): *dreidimensional* 'three-dimensional'; *kompakt* 'compact'; *hell* 'bright', *obskur* 'obscure'
`arousing or holding someone's attention`: *interessant* 'interesting'
`high in price`: *teuer* 'expensive'
`distinguished by special consecration`: *sakral* 'sacred'
`of high quality`: *hochwertig* 'of high quality'
`located far away, remote`: *entfernt* 'distant'

Geruch 'smell'
FORMAL: *modrig* 'musty', *muffig* 'musty', *brenzlig* 'burnt', *ranzig* 'rotten'
CONSTITUTIVE (*degree*): *streng* 'strong', *intensiv* 'intense'
`strong and sharp`: *beißend* 'acrid', *stechend* 'pungent', *penetrant* 'pungent'
`evaluating in a negative/positive manner`: *unangenehm* 'unpleasant', *übel* 'foul', *widerlich* 'disgusting', *ekelhaft* 'disgusting', *süß* 'sweet'

The three nouns *Kurve* 'curve', *Objekt* 'object', and *Geruch* 'smell' have various qualia dimensions selected by several adjectival collocates: FORMAL, CONSTITUTIVE, and TELIC. A number of collocates were not accommodated by any qualia role.

Preis 'price'
CONSTITUTIVE (*degree*): *hoch* 'high', *horrend* 'horrendous', *stolz* 'stiff', *niedrig* 'low', *erschwinglich* 'affordable', *vernünftig* 'reasonable', *moderat* 'moderate', *tief* 'low', *rückläufig* 'descreasing'
`that does not change often`: *fest* 'fixed', *stabil* 'stable', *konstant* 'konstant'
`positive evaluation`: *toll* 'great', *attraktiv* 'attractive'
`existing in the present time`: *heutig* 'of today'

Strafe 'punishment'
CONSTITUTIVE (*degree*): *drakonisch* 'draconian', *unmenschlich* 'inhumane', *scharf* 'harsh', *streng* 'strict', *hart* 'harsh' , *empfindlich* 'severe'*saftig* 'stiff', *deftig* 'severe', *hoch* 'high', *schwer* 'heavy', *schlimm* 'severe', *niedrig* 'low', *mild* 'mild'
TELIC : *symbolisch* 'symbolic'
`of long duration` : *lebenslang* 'lifelong'

Anzahl 'amount'
CONSTITUTIVE (*degree*): *stattlich* 'stately', *groß* 'large', *hoch* 'high', *schier* 'sheer', *beträchtlich* 'large', *riesig* 'huge', *erklecklich* 'considerable', *bedeutend* 'substantial', *bescheiden* 'modest', *gering* 'small', *klein* 'small', *überschaubar* 'manageable'

As discussed in section 4.2.2, the nouns *price*, *amount*, and *punishment* are mostly described in terms of their degree. All the corresponding collocates have been thus assigned to the CONSTITUTIVE role with the relation *degree*.

Thema 'issue/topic'
`important`: *groß* 'big/major', *heiß* 'hot', *beherrschend* 'dominant', *wichtig* 'important', *zentral* 'central'
`urgent`: *drängend* 'pressing', *brennend* 'urgent'
`unpleasant`: *unbequem* 'uncomfortable', *leidig* 'vexed', *heikel* 'delicate', *sensibel* 'sensitive', *brisant* 'controversial'
`very complex`: *sperrig* 'unwieldy'
`of long duration`: *ewig* 'eternal'
`existing in the present time`: *aktuell* 'current'
`arousing or holding someone's attention`: *interessant* 'interesting'
`restricted to a particular locality or neighborhood`: *lokal* 'local'

Ähnlichkeit 'similarity'
FORMAL (*is_a*): *äußerlich* 'external'
CONSTITUTIVE (*degree*): *groß* 'strong', *stark* 'strong', *unheimlich* 'uncanny', *leise* 'slight'
`causing surprise`: *frappierend* 'striking', *frappant* 'striking', *erstaunlich* 'surprising'
`in an accurate manner`: *exakt* 'precise'
`in an accurate manner` + `Anti`: *vage* 'vague'
`determined at random`: *beliebig* 'arbitrary'
`that changes often`: *variabel* 'variable'
`lasting for a short time`: *flüchtig* 'fleeting'

Ziel 'goal'
CONSTITUTIVE (*degree*): *ober* 'ultimate', *groß* 'great'
`that can be achieved with a lot of ambition`: *hochgesteckt* 'ambitious', *hoch* 'high', *ehrgeizig* 'ambitious'
`chosen by someone themselves`: *selbstgesteckt* 'self-imposed'
`accurately stated or described`: *klar* 'clear'
`characterized by determination`: *fest* 'firm'
`of foremost importance or value`: *primär* 'primary'
`serving as an essential component`: *zentral* 'central', *wichtig* 'important'
`extending over a certain period`: *langfristig* 'long-term', *kurzfristig* 'short-term'
`sublime, awe-inspiring`: *hehr* 'noble'
`pursued by two or more persons`: *gemeinsam* 'joint'

Angst 'fear'
CONSTITUTIVE (*degree*): *groß* 'strong/big', *tief* 'deep', *höllisch* 'hellish', *furchtbar* 'tremendous'
`nothing but`: *nackt* 'pure', *pur* 'pure', *blank* 'sheer'
`of long duration`: *ewig* 'eternal'
`of long duration` + *uninterrupted in time*: *ständig* 'constant', *stet* 'constant'
`such that its reason is not clear`: *diffus* 'vague', *dumpf* 'vague'
`increased to abnormal`: *krankhaft* 'pathological'

Interesse 'interest'
CONSTITUTIVE (*degree*): *hoch* 'strong', *groß* 'strong/big', *stark* 'strong', *massiv* 'massive', *gering* 'little'
`lively`: *lebhaft* 'lively', *rege* 'active'
`vital`: *vital* 'vital', *elementar* 'fundamental'
`broad`: *breit* 'broad'
`genuine`: *echt* 'genuine', *wahr* 'genuine'
`belonging or peculiar to a specified person`: *ureigen* 'own'
`contradictory, in disagreement`: *widerstreitend* 'conflicting'
`characterized by a material value`: *handfest* 'tangible'

The abstract nouns *issue/topic*, *resemblance*, *goal*, *fear* and *punishment* could not be successfully modeled with the qualia roles. Only the relation *degree*, the extension of the CONSTITUTIVE role could be applied. Other meanings have not been assigned to any qualia roles as it is not clear which dimensions they belong to. Consider the AN phrase *unbequemes Thema* 'uncomfortable issue/topic': the interpretation of this expression is *an issue that people feel uncomfortable to discuss or to deal with*. Basically, in this case, *Thema* is an abstract entity that *causes* discomfort. Sorokin et al. (2015) do not introduce new relations for such cases, but instead use the already existing ones and just mark the direction of the relation. For instance:

(21) a. AGENTIVE (*cause*): *Vulkaninsel* 'volcanic island' – the modifier *Vulkan* 'volcano' is the cause of *Insel* 'island'

b. AGENTIVE ($cause^{-1}$): *Regenwolke* 'rain cloud' – the head *Wolke* 'cloud' is the cause of *Regen* 'rain'

Thus, the relation $cause^{-1}$ belonging to the AGENTIVE qualia dimension is a possible relation for expressions such as *uncomfortable issue*. Similarly, the adjective in the phrases *frappierende/frappante Ähnlichkeit* 'striking similarity' expresses the caused effect of the base noun: similarity that *causes* surprise or amazement.

Conclusion

Coverage. The case study supports the hypothesis that the qualia roles are well-suited for modeling the semantics of concrete nouns. It shows that there is a noticeable tendency for adjectives in AN combinations to select either the FORMAL or the CONSTITUTIVE dimension of concrete nouns. There are also a few examples assigned to the TELIC role, whereas the AGENTIVE dimension has only been selected once. The AGENTIVE role encodes the information about the origin of an entity. Its absence in the modeling of the sample might indicate that this information is rarely expressed by descriptive adjectives. The semantic structure of abstract nouns, or in Lyons (1977)'s

terminology, second- and third-order entities, is not completely captured by the available qualia roles or by their extended version. It is at least true for the aspects of nominal meaning that are typically described by adjectives. Most of the extended qualia roles suggested in Lenci et al. (2000) and Ruimy et al. (2001) are listed under the CONSTITUTIVE role. The possible solution would be to add new relations to that dimension. However, it is not yet clear whether the aspects of meaning such as *degree*, or, say, *manner* belong to the CONSTITUTIVE dimension of abstract entities or any other qualia dimension at all.

Granularity. In contrast to the fine-grained Lexical Functions, the qualia roles are very general. The resulting modeling based only on the qualia dimensions is thus very coarse-grained and further refinements for each role are required. A more specific classification should be similar to the annotation scheme for the relations in nominal compounds described in Sorokin et al. (2015), but tailored for adjective-noun phrases.

4.4 Frame Semantics

The main idea of Frame Semantics is that word meanings are defined against background knowledge about situations in the real world (Fillmore, 1982). The reason for words to exist is grounded in human experiences and institutions. For instance, the existence of lexical units denoting days of the week and other temporal markers is motivated by understanding that there is a natural cycle of one day starting and ending, there is a commonly accepted concept of a *week* that consists of seven days, and the tradition to work on some days and rest on others – a *weekend* (Fillmore, 1985). In a similar manner, words like *novel*, *monograph*, and *limerick* exist to refer to texts and imply understanding a situation where a TEXT is created by an AUTHOR and has a TOPIC.

Frames are a way of describing such situations in a structured way. Fillmore (1982, p.119) defines frames as "systems of categories structured in accordance with some motivating context". Words associated with certain situations or scenes *evoke* the corresponding frames with all the other concepts that belong to that frame. In this manner, the noun *limerick* evokes the frame "Text" presented in Figure 4.3.[11]

Other lexical items that evoke the frame "Text" are *brochure*, *diary*, *essay*, *fiction*, *manuscript*, *etc.* Think of how L2 learners are usually taught vocabulary: the words they learn in one or several consequent lessons usually belong to the same domain, they are not random. Such word groups are held together by the underlying frameworks of knowledge (Fillmore, 1985).

Based on the theory of Frame Semantics, a large lexical database for English has been developed at the International Computer Science Institute in Berkeley, California. Inspired by the Princeton WordNet (Fellbaum, 1998), it was named *FrameNet* (Fillmore et al., 2003). FrameNet describes meanings

[11] https://framenet.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Text, *last accessed on 07.10.2022*

Frame "Text"

Definition: A Text is an entity that contains linguistic, symbolic information on a Topic, created by an Author at the Time_of_creation. It may be a physical entity that is made of a certain Material etc. It may be constructed for a Beneficiary.

Core FEs:

- **Text** - The Text is the entity that contains linguistic information. This an incorporated FE for each LU in this frame.

Non-Core:

- **Author** - The Author is the person who produced a particular Text.
- **Beneficiary** - The person for whom the Text is produced.
- **Containing_text** - This is a text in which the Text is located.
- **Descriptor** - A characteristic of the Text
- **Genre** - The Genre or sub-Genre of the Text.
- **Material** - What a Text is made of (if applicable).
- **Medium** - The Medium is the language in which a text is composed, or the physical medium (radio, ink) in which it is distributed.
- **Place** - The region where the Text is sold, circulated, or otherwise popular.
- **Subpart** - The smaller texts that make up the Text.
- **Time_of_creation** - The Time at which the Text was created.
- **Title** - The Title of a Text.
- **Topic** - Topic is what the Text is about.
- **Use** - The use for which the text was produced; its purpose.

LUs: *account.n, article.n, autobiography.n, ballad.n, benediction.n, biography.n, book.n, booklet.n, brochure.n, bulletin.n, chronicle.n, comedy.n, diary.n, drama.n, edition.n, editorial.n, elegy.n, epic.a, epic.n, epigram.n, epilogue.n, epistle.n, essay.n, eulogy.n, exemplum.n, fable.n, fanzine.n, festschrift.n, fiction.n, grimoire.n, hagiography.n, haiku.n, handbook.n, hardback.n, history.n, imprecation.n, issue.n, journal.n, lay.n, letter.n, life.n,* ***limerick.n****, list.n, literature.n, lyric.n, mag.n, magazine.n, manual.n, manuscript.n, material.n, meditation.n, memoir.n, missive.n, monograph.n, mystery.n, newspaper.n, novel.n, novelette.n, novella.n, obituary.n, ode.n, pamphlet.n, paper [article].n, paper [newspaper].n, paperback.n, periodical.n, poem.n, poetry.n, publication.n, record.n, report.n, rhyme.n, saga.n, screenplay.n, script.n, sermon.n, song.n, sonnet.n, speech.n, spellbook.n, tetralogy.n, textbook.n, thriller.n, tome.n, tract.n, tractate.n, tragedy.n, treatise.n, trilogy.n, volume.n, whodunit.n, writings.n*

FIGURE 4.3: Frame "Text": Definition, Frame Elements and evoking Lexical Units as listed in the Berkeley FrameNet.

of words in terms of semantic frames reflecting the words' semantic and syntactic valences. A sense of a word is represented in FrameNet as a lexical unit (LU). A semantic frame is a structured description of a situation or event evoked by an LU. The main building blocks of the frames are semantic roles called *Frame Elements*. FrameNet contains manual annotations of more than 13,000 word senses in context (>200,000 sentences) with respect to the frames they evoke (1,224 frames). Table 4.4 presents the number of LUs available in FrameNet at the moment.[12]

LUs	amount in FN
Nouns	5,575
Verbs	5,214
Adjectives	2,407
Other POS	490
Total	13,686

TABLE 4.4: Statistics for the Berkeley FrameNet: number of Lexical Units by their part-of-speech.

As illustrated in Figure 4.3, a frame entry in FrameNet contains the following information:

- the definition of the frame;
- the list of Frame Elements (FEs) with definitions and example sentences;
- the list of all the LUs that evoke that frame.

The highlighted words in the definition are Frame Elements (FEs). For many FEs, apart from a definition, one or few example sentences are listed. For instance, the FE GENRE in the frame "Text" is exemplified by the following sentences:

(22) a. She writes only historical$_{[genre]}$ fiction$_{[text]}$.

b. I've got three comedic$_{[genre]}$ screenplays$_{[text]}$ on my desk right now!

Apart from that, for each frame, relations to other frames are listed. All the possible relations are listed leaving some fields left underspecified if they cannot be applied. For the frame "Text", the **frame-frame relations** are as follows:

Inherits from: Artifact

Is Inherited by: Documents, Records

[12] https://framenet.icsi.berkeley.edu/fndrupal/current_status, *last accessed on 09.10.2022*

Perspective on:

Is Perspectivized in:

Uses: Text_creation

Is Used by: Law, Medium, Reference_text, Topic

Subframe of:

Has Subframe(s):

Precedes:

Is Preceded by:

Is Inchoative of:

Is Causative of:

See also:

The interest of this thesis lies in describing nominal meaning in terms of properties that are selected by adjectives. There are two possible approaches as to how FrameNet can be used for this purpose: to consider frames evoked by nominal bases or the ones evoked by adjectival collocates. Depending on the approach, different information from FrameNet would be used.

Noun. Starting with the noun as the evoking element is an intuitive approach as the noun is the base of co-occurrence both syntactically and semantically, and the collocate is contingent on the noun. In this approach, the relevant information is available in the structure of each frame, namely, in the Frame Elements. For instance, in the AN phrase *old friend*, the noun *friend* evokes the frame "Personal relationship".[13] To describe the relation between the noun and its collocate, one needs to select the suitable Frame Element from the list of Frame Elements available within the frame. In this case, it is the FE DURATION referring to the duration of friendship.

Adjective. Starting with the dependent collocate is less intuitive, but also possible within Frame Semantics as adjectives are capable of evoking frames as well. This is a one-step approach where Frame Elements are of no importance - it is only the title of the evoked frame that provides the required information. Frame Elements cannot be taken into consideration in this approach because in an adjectival frame, the noun is an entity that is being modified and that is the only information that will be reflected in a frame's structure.

For instance, given the phrase *old friend*, one would search for the frame evoked by the adjective *old*. Already at this stage, a first challenge occurs: the LU *old* in the sense "long-time" is not present in FrameNet. This can be resolved by searching for a synonym such as *long-time*, *lasting* or just *long*.

[13] All the frames mentioned in this section are listed in alphabetic order in Appendix F.

The last two adjectives are indeed present in FrameNet and evoke the frame "Duration description" (presented in Figure 4.4).

Frame "Duration description"

Definition: An event or state (here termed an Eventuality) or time-Period is described in terms of the Degree to which its duration diverges from a standard. As is usual for gradable attributes, the standard of comparison is normally based on the expected duration for the type of Eventuality or Period.

Core FEs:

- **Eventuality** - A situation, action or event whose duration is being described. The Eventuality occurs as the External Argument if the adjective is used predicatively, or as the modified noun if the adjective is used attributively.
- **Period** - Sometimes these adjectives describe simply a period of time, rather than an event or state. The Period almost invariably occurs as the modified noun with an attributive adjective.

Non-Core:

- **Degree** - Degree to which event occurs.

LUs: *abiding.a, brief.a, chronic.a, enduring.a, ephemeral.a, eternal.a, extended.a, interim.a,* ***lasting.a****, lengthy.a,* ***long.a****, momentary.a, perpetual.a, persistent.a, rapid.a, short.a, sustained.a*

FIGURE 4.4: Frame "Duration description": Definition, Frame Elements and evoking LUs as listed in the Berkeley FrameNet.

The name and the definition of this frame tell us that the adjective *old* expresses the *duration* of the friendship. What about the Frame Elements available for this frame? To describe the relation between the two constituents of the phrase *old friend* from the point of view of the adjective, only the FE EVENTUALITY can be applied which does not provide any new semantic information apart from generalizing from *friendship* to *eventuality*. Thus, relevant semantic information is reflected only in the name of the frame - "Duration".

As mentioned above, frames have a hierarchical structure in terms of inheritance. The frame "Duration description" inherits from a more general frame "Measurable attributes", which in its turn inherits from the frame "Gradable attributes". The latter does not contain any LUs but serves as a parent frame for a large number of attribute-frames: Abundance, Addiction, Artificiality, Being_active, Being_at_risk, Being_dry, Being_rotted, Being_wet, Biological_urge, **Body_description_holistic**, Body_description_part, Capability, Certainty, Change_post-state, Chemical_potency, Chemical-sense_description,

Color, Color_qualities, Compatibility, Completeness, Correctness, Desirability, Difficulty, Dynamism, Familiarity, Frequency, Fullness, Gradable_artistic_quality, Inclination, Intoxication, **Judgment_of_intensity**, Level_of_force_exertion, Level_of_force_resistance, Level_of_light, Likelihood, Lively_place, Luck, Margin_of_resolution, Measurable_attributes, Medical_conditions, Notability, Obscurity, **Obviousness**, Personal_success, Position_on_a_scale, Prominence, Risky_situation, Sharpness, Similarity, **Social_behavior_evaluation**, Suitability, System_complexity, Taking_time, Typicality, Usefulness, Wealthiness.

Many of the listed attribute-frames look like a reasonable way of expressing the adjectival meaning in an adjective-noun phrase. Consider a few examples from the sample:

- *spindly legs*: the adjective *spindly* evokes the frame "Body description holistic" → possible relation Body Description
- *strong smell*: the adjective *strong* evokes the frame "Judgment of intensity" → possible relation Intensity
- *clear goal*: the adjective *clear* evokes the frame "Obviousness" → possible relation Obviousness
- *sweet dog*: the adjective *sweet* evokes the frame "Social interaction evaluation" → possible relation Evaluation

In general, the approach with adjective as target, though less intuitive, is promising. However, adopting this approach using FrameNet is rejected due to several reasons. First of all, the coverage of adjectives is small - remember that it was necessary to search for synonyms of the adjective *old* in the example above. As presented in table 4.4, FrameNet covers only 2,407 adjectival LUs - this amount is quite low. Secondly, this approach does not take advantage of the main idea of Frame Semantics as it ignores the structure of frames and does not make use of Frame Elements. Thus, combining the two previous points: for such an approach, it would be more reasonable to use another lexical resource that provides information about semantics of adjectives in a more structured way and offers a larger coverage (e.g. Princeton Wordnet covers over 31 thousand adjectival LUs)[14] - this will be tested in section 4.5.

This section will focus on the nouns and their frame structure in terms of Frame Elements - the approach was taken in the study by Strakatova and Hinrichs (2019) and proved to be reasonable. The next section will elaborate on the notion of a Frame Element and will be followed by the case study in the subsequent section 4.4.2. In the case study, the frames evoked by the head nouns in the phrases will provide a list of Frame Elements. These FEs will be used to describe the relation between the constituents of the phrases.

[14] https://wordnet.princeton.edu/documentation/21-wnstats7wn, *last accessed on 13.10.2022*

4.4.1 Frame Elements

Frame elements describe *semantic valences* of LUs and resemble in a way argument variables in logic (Fillmore et al., 2003). Valence of a word reflects its required combinability with other words or phrases. A typical example is a transitive verb such as *to borrow* that requires a direct object. Using such a verb without an object would result in both syntactic and semantic unacceptability:

(23) a. Alex *borrowed* the book.
b. *Alex *borrowed*.

The sentence in example (23b) is ungrammatical in English and semantically not acceptable. In FrameNet, the verb's requirement of a subject and an object is expressed in terms of Frame Elements Borrower (the subject) and Theme (the object).
Four groups of Frame Elements are distinguished in FrameNet: **core**, **core-unexpressed**, **peripheral**, and **extra-thematic**.

Core elements are conceptually essential components of a frame that distinguish it from other frames. In the case of the frame "Text", there is only one core element – Text. For comparison, the frame "Duration description" has two core elements Eventuality and Period. An element is considered core if it is always overtly realised or, in case it is implicit, it always has to be understood from the context (Ruppenhofer et al., 2016).

Core-unexpressed FEs behave like core FEs, but they are not necessarily expressed in the descendants of the given frame. The main reason for having this category of FEs is that those FEs are not expressed separately for the LUs of the children frames.

Peripheral FEs are not unique to a particular frame. They specify additional characteristics of a situation such as Manner, Time, Place, Degree, Status, etc. In the frame "Text", the peripheral FEs are Author, Genre, Material, Title, Topic, etc., whereas in "Duration description" there is only one - Degree.

Extra-thematic FEs do not conceptually belong in the given frame, they provide details about the participants or the setting. They combine with a large number of other frames. In some sense, they are not proper frame elements because they "explicitly introduce an independent scene" (Ruppenhofer et al., 2016, p.103), i.e. there are separate elements evoking separate frames. However, these elements are so closely connected to the the target elements that they are annotated from the perspective of the target elements. FEs that are classified as extra-thematic in some frames can be considered peripheral or core in other frames (Ruppenhofer et al., 2016). The major types of extra-thematic frames are those related to:

- temporal structure;
- places;
- additional participants;

- describing participants;
- describing events;
- co-occurring events and circumstances;
- causal chain;
- degree.

The FrameNet book by Ruppenhofer et al. (2016) gives a detailed overview of those groups of FEs and provides some comparisons between some typical peripheral FEs such as DURATION and extra-thematic ones such as PERIOD_OF_ITERATIONS and DURATION_OF_FINAL_STATE.

All the four groups of FEs are distinguished in abstract frames such as "State" or "Process". In concrete frames, such as "Vehicle" or "Food", two categories of FEs are distinguished: core and non-core (peripheral), where in most cases there is one core frame element, e.g. VEHICLE in the frame "Vehicle" or FOOD in the frame "Food", and the rest are peripheral – CONSTITUENT_PARTS, TYPE, DESCRIPTOR (Ruppenhofer et al., 2016).

Granularity and coverage

The topic of granularity is central in the context of FrameNet. As Ruppenhofer et al. (2016) note, the approach to the level of granularity may differ depending on research questions. Frames can be constructed for any domain and thus be very fine-grained. For instance, there is a multilingual FrameNet only for the football domain.[15] The more specific the frames are, the more manual work is required. Therefore, the developers of FrameNet aim at keeping the level of granularity at an optimal level based on utility: the major goal is to reflect the most fundamental differences between different word senses (Ruppenhofer et al., 2016).

In terms of coverage, FrameNet provides descriptions for a wide semantic variety of nouns, verbs, and adjectives. Natural kind nouns and artifacts are represented less than event nouns or relational nouns. The latter two categories are more frame-evoking, whereas artifacts and natural kind nouns usually take up the role of dependents. For instance, the noun *soda* is not listed in FrameNet as a frame-evoking LU. However, as example (24a) illustrates, it is annotated as dependent in a sentence with the frame-evoking verb *to buy* within the frame "Commerce buy".

Consider two other nouns: *lioness* and *zebra*. None of them is listed in FrameNet as a frame-evoking LU although there is a frame that can accommodate them, namely "Animals". However, as dependents, both nouns are annotated with Frame Elements in example (24b). Within the frame "Hunting" evoked by the verb *to hunt*, they are assigned the roles of a HUNTER and FOOD.[16]

(24) a. Now I can BUY a soda$_{[goods]}$ and spend money.
b. I once watched a lioness$_{[hunter]}$ HUNTING a zebra$_{[food]}$.

[15] www.kicktionary.de, *last accessed on 17.07.2021*

[16] Both example sentences and their annotations are taken from Berkeley FrameNet.

Nevertheless, artifacts and natural kind nouns are capable of evoking frames of their own and they do possess a basic frame structure (Ruppenhofer et al., 2016, p.8). For instance, natural features are described by the following FEs: CONSTITUENT PARTS, CONTAINER-POSSESSOR, DESCRIPTOR, FORMATIONAL_CAUSE, NAME, LOCATION, TYPE. Artifacts are described in terms of the FEs ARTIFACT, NAME, CREATOR, MATERIAL, TYPE, USE, DESCRIPTOR,TIME_OF_CREATION.

The above listed FEs for physical entities bear a striking resemblance to the *qualia roles* from the Generative Lexicon (see section 4.3) indicating that those roles are well-suited for describing the semantics of physical entities. Abstract concepts such as processes are described in terms of the FEs MANNER, DURATION, TIME, PLACE, RESULT, EXPLANATION, where the latter two resemble telic and agentive roles in a way, but the rest are not reflected in the qualia structure.

The idea that qualia roles are relevant for FrameNet has been recently tested in the FrameNet Brasil.[17] Belcavello et al. (2020) establish qualia-inspired relations between LUs. As qualia roles are way too generic, they are specified via frames in the same database as *mediators*. For instance, the LU *pizza* is connected to the LU *flour* via the relation `is_made_of` (constitutive quale) mediated by the frame "Ingredients" in terms of FEs MATERIAL (*flour*) and PRODUCT (*pizza*). In the same manner, the telic role for *pizza* is reflected in the relation `meant_to` to the LU *eat* and is specified by the frame "Tool_purpose" (Belcavello et al., 2020).

Employing the extended qualia roles for modeling the relations in adjective-noun phrases proved to be a reasonable approach that exhibited substantial gaps in terms of coverage (see section 4.3.2). The similarity of FEs to the extended qualia roles allow for an assumption that FEs can offer an equivalently suitable inventory of relations with a broader coverage and a more refined granularity. This hypothesis will be tested in the next subsection.

4.4.2 Case study: Frame Elements

The case study presented in this section investigates whether Frame Elements can be used to model the relations in AN combinations. It is based on the investigation by Strakatova and Hinrichs (2019). The original data sample of 11 nouns has been extended to 23 nouns as discussed in section 4.1. Apart from that, new adjectives have been added to the sample.

FrameNet for German data

FrameNets exist for several languages and German is not an exception. The ambition to create a Framenet-like resource for German has been successfully realized in the SALSA and SALSA II projects.[18] In the SALSA project,

[17] `http://www.framenetbr.ufjf.br`, *last accessed on 10.01.2022*

[18] `http://www.coli.uni-saarland.de/projects/salsa/`, *last accessed on 10.01.2022*

a large German corpus TIGER (Brants et al., 2004) was manually annotated with Frame Elements. The project assumes that frames can be applied cross-linguistically, but for some frames, adaptation and modification are necessary (Burchardt et al., 2009). SALSA II provides a semantic lexicon for German that contains frame descriptions for German lexical units. However, the number of frame-evoking nouns is much lower than in the Berkeley FrameNet for English: release 2.0 of SALSA contains 477 nouns.

The phenomenon which Padó (2007) refers to as *concept-level cross-lingual parallelism* implies that the linguistic categories introduced by a linguistic theory can be used to describe linguistic units in different languages. It presupposes *language independence of the conceptual inventories* of the theories. The degree of parallelism depends to a large extent on the granularity of the categories: the more coarse-grained they are, the higher is the degree of cross-lingual parallelism. Frames in FrameNet exhibit a high degree of concept-level parallelism (Padó, 2007). Burchardt et al. (2009) have shown that Frame Semantics offers a robust framework for modeling word meanings in different languages.

Based on this evidence, the case study has been conducted relying on the English FrameNet as it provides a broader coverage than the German one. In collocations (as defined in this thesis) and in free phrases, the nominal base retains its literal meaning, thus, translating the sample nouns from German into English does not pose any problems. The translated lexical unit is then searched in the Berkeley FrameNet database and the corresponding frame is selected. The adjectival collocates are classified based on the FEs belonging to that frame. Following the procedure described in the previous case studies, the discussion of the annotation starts with the most concrete nouns followed by the more abstract ones. All the definitions of the FEs are taken from the entries in FrameNet.

Modeling with Frame Elements

The LU *chocolate* evokes the frame "Food" which is described in terms of the following FEs:

- Food (core)
- Constituent_parts (peripheral) – A part of the Food.
- Descriptor (peripheral) – This FE indicates a characteristic or description of the food.
- Type (peripheral) – This FE identifies a particular Type of the food item.

The suitable FEs from this list will be assigned to the collocates of the noun *chocolate*, the unused FEs will be listed for each noun as well. The full definitions and structure of frames for the following nouns in the sample will be presented in Appendix F. In the main body of the thesis, only the FEs will be listed without their definitions.

Schokolade 'chocolate' Frame "Food"
FOOD (C): *heiß* 'hot' DESCRIPTOR: *edel* 'premium, fine', *hochwertig* 'of high quality', *lecker* 'tasty', *süß* 'sweet', *knackig* 'crunchy', *flüssig* 'liquid', *teuer* 'expensive' TYPE: *schwarz* 'dark', *dunkel* 'dark', *weiß* 'white'
Other FEs within the frame: CONSTITUENT_PARTS

Some of the AN combinations with *chocolate* as their base are highly lexicalized. As discussed in subsection 4.3.2, the pairs *dark/white chocolate* are in an *is_a* relation to *chocolate*. FrameNet offers a perfect FE for this type of relation – TYPE. The available FEs in this frame allow a fine-grained differentiation for the relation *a type of* suggested in section 4.3.2: *hot chocolate* is not just a type of chocolate, it is a separate kind of food. For this case, FrameNet's notion of a *core element* is highly suitable, as combining *chocolate* with *hot* does not just introduce an additional/specificational meaning, it changes the very nature of the nominal base from a solid food to a beverage. Thus, the collocate *heiß* 'hot' has been accommodated by the core FE FOOD.

Schuh 'shoe' Frame "Clothing"
STYLE: *hochhackig* 'high-heeled', *flach* 'flat', *spitz* 'pointy' DESCRIPTOR: *bequem* 'comfortable', *elegant* 'elegant', *braun* 'brown', *schwarz* 'black', *rot* 'red', *weiß* 'white', *rosa* 'pink', *blau* 'blue', *schmutzig* 'dirty', *naß* 'wet' MATERIAL: *seiden* 'silk'
Other FEs within the frame: GARMENT (core), BODY LOCATION, CREATOR, NAME, SUBREGION, TIME OF CREATION, USE, WEARER.

In contrast to the previous example, the frame "Clothing" evoked by the LU *shoe* has a very rich structure in terms of FEs (1 core and 10 peripheral FEs). Three collocates have been labeled with the FE STYLE ("This FE identifies the design, or cut, of the clothing"); one with MATERIAL. The remaining adjectives are accommodated by a very FE DESCRIPTOR that expresses a very general meaning (this issue will be discussed at the end of this section).

Hund 'dog' Frame "Animals"
DESCRIPTOR: *süß* 'sweet', *arm* 'poor', *niedlich* 'cute', *tollwütig* 'rabid' CHARACTERISTIC: *bissig* 'biting', *treu* 'loyal', *aggressiv* 'agressive', *scharf* 'vicious', *klug* 'smart', *riesig* 'huge', *dumm* 'stupid', *jung* 'young', *dick* 'fat', *klein* 'small'
Other FEs within the frame: ANIMAL (core), ORIGIN.

The frame "Animals" does not have a very rich structure. However, the available FEs could accommodate the few collocates of the base *Hund*. However, physical and mental characteristics were conflated together under the FE CHARACTERISTIC.

Fels 'cliff' Frame "Natural features"
CONSTITUENT_PARTS: *nackt* 'bare', *blank* 'bare' DESCRIPTOR: *schwarz* 'black', *rötlich* 'reddish', *grau* 'grey', *naß* 'wet', *glatt* 'smooth', *schroff* 'craggy', *steil* 'steep', *hart* 'hard'
Other FEs within the frame: LOCALE (core), CONTAINER POSSESSOR, FORMATIONAL CAUSE, NAME, RELATIVE LOCATION, TYPE.

Two collocates of the noun *Fels* 'cliff' were accommodated by an FE CONSTITUENT_PARTS within the frame "Natural features". All the other ones are labeled as generic DESCRIPTOR.

Wald 'forest' Frame "Biological area"
DESCRIPTOR: *tief* 'deep', *herrlich* 'magnificent', *riesig* 'huge', *endlos* 'endless', *feucht* 'humid', *trocken* 'dry', *immergrün* 'evergreen', *wild* 'wild', *unberührt* 'untouched', *nah* 'close' CONSTITUENT_PARTS: *licht* 'sparse', *dicht* 'dense'
Other FEs within the frame: LOCALE (core), CONTAINER POSSESSOR, FORMATIONAL CAUSE, NAME, RELATIVE LOCATION.

The frame "Biological_area" includes seven FEs that are enough for accommodating the modifiers of the base *Wald*. A *sparse forest* is a forest with

less trees and more open areas, therefore, it was assigned to the FE CONSTITUENT_PARTS. For the adjective *deep* in the context of describing a forest, FrameNet provides the following example sentence: "Apparently they found him at the bottom of a really deep bog." The phrase *really deep* is annotated as DESCRIPTOR. Thus, the role of *deep* in *deep forest* is also annotated as DESCRIPTOR. All the remaining adjective were also labeled as DESCRIPTOR because none of the other FEs available are suitable.

Baum 'tree' Frame "Plants"
CHARACTERISTIC: *mächtig* 'mighty,massive', *riesig* 'huge', *hoch* 'high', *kahl* 'leafless', *immergrün* 'evergreen', *alt* 'old', *jung* 'young', *krank* 'sick', *gesund* 'healthy', *morsch* 'rotten' DESCRIPTOR: *prächtig* 'magnificent', *heilig* 'sacred'
Other FEs within the frame: PLANT (core), ORIGIN.

The frame "Plants" evoked by the noun *Baum* is similar to the frame "Animals". Its structure is highly minimalistic which is typical for concrete nouns of natural kind. The two collocates are accommodated by the available FEs. The general FE CHARACTERISTIC accommodated most of the adjective as there are no other suitable options. It is also not entirely clear what the difference between CHARACTERISTIC and DESCRITOR is. In the present modeling, the former is considered the innate property of a tree whereas the latter refers to its external perception or evaluation.

Bein 'leg' Frame "Body parts"
DESCRIPTOR: *staksig* 'spindly', *schwer* 'heavy', *lang* 'long', *kurz* 'short', *stämmig* 'sturdy', *schwach* 'weak', *dick* 'fat', *schlank* 'slim', *dünn* 'thin', *nackt* 'naked', *gesund* 'healthy', *wacklig* 'shaky', *müde* 'tired'
Other FEs within the frame: BODY PART (core), POSSESSOR (core), ATTACHMENT, ORIENTATIONAL LOCATION, SUBREGION.

All the collocates of the noun *Bein* 'leg' were modeled within this approach by means of a highly general FE DESCRIPTOR.

Regen 'rain' Frame "Precipitation"
PRECIPITATION (C): *sauer* 'acid'
QUANTITY: *sintflutartig* 'torrential', *stark* 'heavy', *kräftig* 'heavy', *heftig* 'heavy', *leicht* 'light'
MANNER: *strömend* 'pouring', *spärlich* 'sparse', *fein* 'fine'
DURATION: *ewig* 'endless', *endlos* 'endless'
TEMPERATURE: *warm* 'warm', *kalt* 'cold', *kühl* 'cool'
Other FEs within the frame: PLACE (core), TIME (core), EXPLANATION, FREQUENCY, RATE.

The LU *rain* evokes the frame "Precipitation". Unlike the above discussed LUs, it is not an entity or a locale, but a process. Along with events and relations, processes have a rich frame structure. It is a typical characteristic of more abstract frames to include more than one core frame element. In the case of "Precipitation", there are three core FEs (PRECIPITATION, PLACE, TIME) and seven peripheral ones (DURATION, EXPLANATION, FREQUENCY, MANNER, QUANTITY, RATE, TEMPERATURE). A number of collocates describe the intensity of the rain which implies the amount of precipitation, thus, they are accommodated by the FE "Quantity". The phrase *acid rain* is a term that presupposes a certain type of precipitation – the one containing acidic components. Therefore, it has been annotated as with the FE PRECIPITATION.

Freund 'friend' Frame "Personal relationship"
DURATION: *alt* 'old', *lebenslang* 'lifelong', *langjährig* 'longtime', *neu* 'new'
DEGREE: *eng* 'close', *dick* 'close', *nah* 'close', *gut* 'good'
MANNER: *wahr* 'true', *echt* 'real', *falsch* 'fake', *treu* 'faithful', *verlässlich* 'reliable'
DEPICTIVE: *mächtig* 'powerful', *reich* 'rich'
Other FEs within the frame: PARTNER 1 (core), PARTNER 2 (core), PARTNERS (core), MEANS, RELATIONSHIP, SOURCE OF RELATIONSHIP.

The frame "Personal_relationship" associated with the noun *friend* is quite broad and covers not only different types of relationships such as *friendship* or *marriage* and people engaged in them, but also events linked to relationships, e.g. *marriage, adultery, divorce* . Thus, this frame has social components

to it and requires an understanding of the interactions between people in the society and the associated events. This is reflected in the collocates of the noun *friend*: they actually refer to *friendship* and not to a *friend* as a person. For instance, *old friend* describes the DURATION of friendship, not the AGE of the friend.[19] Other collocates describe the DEGREE of friendship (*eng, dick* 'close) and its MANNER (*wahr* 'true', *echt* 'real', *falsch* 'fake'). In the cases when the person themselves is described, the FE DESCRIPTOR was chosen.

Winter 'winter' Frame "Calendric unit"
RELATIVE TIME (C): *letzt* 'last'
`having certain temperature`: *streng* 'severe', *bitterkalt* 'bitter cold', *eiskalt* 'ice cold', *kalt* 'cold', *kühl* 'cool', *warm* 'warm',
`difficult/easy to bear`: *hart* 'harsh', *mild* 'mild'
`in the middle of` (winter): *tief* 'deep'
`characterized by a large/little quantity of water vapor`: *feucht* 'humid', *trocken* 'dry'
`causing sad feelings`: *trist* 'dull, depressing'
`of certain duration`: *lang* 'long', *ewig* 'endless'
Other FEs within the frame: UNIT (core), WHOLE (core), COUNT, LANDMARK PERIOD, NAME, SALIENT EVENT, TRAJECTOR EVENT.

Only one of the collocates of the noun *Winter* 'winter' could be described in terms of FEs within the frame "Calendric unit": *letzt* 'last'. This issue has already been addressed in Strakatova and Hinrichs (2019): the FEs within this frame refer solely to the temporal aspects of the LUs evoking it, and do not reflect other aspects of their meaning such as climate and weather.

[19] This example is of course ambiguous. There can be a context in which *old* would refer to the AGE of a person. The frame "Personal relationship" does not offer a suitable FE for this reading of the adjective. It seems that not only the meaning of the adjective shifts, but also of the noun: in the "age" sense of *old friend*, the noun does not evoke the frame "Personal relationship" but rather the frame "People". In that frame, one of the core FEs is AGE which is suitable for modeling the non-collocational meaning of *old*.

Droge 'drug' Frame "Intoxicants"
INTOXICANT (C): *körpereigen* 'endogenous'
TYPE: *hart* 'hard', *weich* 'soft', *leicht* 'soft', *sanft* 'soft'
DESCRIPTOR: *gefährlich* 'dangerous', *schädlich* 'harmful', *harmlos* 'harmless', *tödlich* 'lethal', *billig* 'cheap', *flüssig* 'liquid'
`(not) allowed by official rules`: *legal* 'legal', *illegal* 'illegal'
Other FEs within the frame: COUNTRY OF ORIGIN.

Similarly to the frame "Food", the structure of the frame "Intoxicants" evoked by the noun *drug* is quite basic with only four FEs available. Not all the adjectives could be accommodated by those FEs.

Team 'team' Frame "Aggregate"
INDIVIDUALS (C): *spanisch* 'Spanish', *deutsch* 'German', *jung* 'young'
AGGREGATE_PROPERTY: *stark* 'strong', *schwach* 'weak', *eingespielt* 'coordinated, experienced', *toll* 'excellent', *gut* 'good', *schlagkräftig* 'powerful', *erfahren* 'experienced', *neu* 'new', *homogen* 'homogenous', *klein* 'small'
`constituting all of its members`: *komplett* 'entire'
Other FEs within the frame: AGGREGATE (core), CONTAINER POSSESSOR, DOMAIN, NAME.

The modeling of the noun *Team* 'team' by means of FEs was only possible on a very general level. This follows the trend observed for the above described LUs: the FE AGGREGATE_PROPERTY in that sense is identical to the FE DESCRIPTOR. It targets nearly any adjectival modifier of the base noun. The adjectives that refer to the members of a team and not the team as a whole were labeled as the core FE INDIVIDUALS.

Kurve 'curve' Frame "Shapes"
SHAPE_PROP: *scharf* 'sharp', *sanft* 'gentle', *steil* 'steep', *eng* 'narrow', *weich* 'gentle', *blau* 'blue', *zackig* 'jagged', *konvex* 'convex'
Other FEs within the frame: SHAPE (core), SUBSTANCE (core), COUNT.

Geruch 'smell' Frame "Sensation"
PERCEPT (C): *modrig* 'musty', *muffig* 'musty', *brenzlig* 'burnt', *ranzig* 'rotten'
DESCRIPTOR: *beißend* 'acrid', *stechend* 'pungent'
DEGREE: *streng* 'strong', *intensiv* 'intense', *penetrant* 'pungent', *unangenehm* 'unpleasant', *übel* 'foul', *widerlich* 'disgusting', *ekelhaft* 'disgusting', *süß* 'sweet'
Other FEs within the frame: BODY PART (core), GROUND (core), PERCEIVER PASSIVE (core), SOURCE (core), TIME.

The nouns *Kurve* 'curve' and *Geruch* 'smell' were successfully modeled with FEs available for their corresponding frames. Four collocates of the noun *smell* are a rare example of using a core FE to accommodate an adjectival modifier. The FE PERCEPT refers to "the entity or phenomenon that produces the sensation".[20] This FE should not be confused with the SOURCE - which is the actualy source, whereas PERCEPT can express a comparison. In the case of *smell*, the collocates define the type of smell, but they do not imply that there is actual mold or mildew or something burning or rotten present.

Objekt 'object' Frame "Entity"
TYPE: *dreidimensional* 'three-dimensional', *dekorativ* 'decorative', *sakral* 'sacred'
`of certain worth`: *wertvoll* 'valuable'
`of certain brightness`: *hell* 'bright', *obskur* 'obscure'
`arousing or holding someone's attention`: *interessant* 'interesting'
`high in price`: *teuer* 'expensive'
`of high quality`: *hochwertig* 'of high quality'
`located far away, remote`: *entfernt* 'distant'
`relatively small, taking up little space`: *kompakt* 'compact'
`amounting to a large number`: *zahlreich* 'numerous'
Other FEs within the frame: ENTITY (core), CONSTITUENT_PARTS, FORMATIONAL_CAUSE, NAME, TYPE.

The frame evoked by the noun *object* has a very basic structure and only a few adjectives could be assigned to a suitable FE.

[20] Definition from the entry for the frame "Sensation" in the Berkeley FrameNet.

Preis 'price' Frame "Commerce scenario"
RATE: *hoch* 'high', *horrend* 'horrendous', *stolz* 'stiff', *niedrig* 'low', *tief* 'low', *moderat* 'moderate', *erschwinglich* 'affordable', *vernünftig* 'reasonable', *fest* 'fixed', *stabil* 'stable', *rückläufig* 'descreasing', *konstant* 'konstant' MANNER: *toll* 'great', *attraktiv* 'attractive' `existing in the present time`: *heutig* 'of today'
Other FEs within the frame: BUYER (core), GOODS (core), MONEY (core), SELLER (core), MEANS, PURPOSE, UNIT.

Strafe 'punishment' Frame "Rewards and punishments"
DEGREE: *drakonisch* 'draconian', *unmenschlich* 'inhumane', *scharf* 'harsh', *streng* 'strict', *hart* 'harsh' , *empfindlich* 'severe'*saftig* 'stiff', *deftig* 'severe', *hoch* 'high', *schwer* 'heavy', *niedrig* 'weak', *mild* 'mild', *schlimm* 'severe' INSTRUMENT: *symbolisch* 'symbolic' OF CERTAIN DURATION: *lebenslang* 'lifelong'
Other FEs within the frame: AGENT (core), EVALUEE (core), REASON (core), RESPONSE ACTION (core unexpressed), DEPICTIVE, INSTRUMENT, MANNER, MEANS, PLACE, PURPOSE, RESULT, TIME.

The LUs *price* and *punishment* are similar in terms of their collocates. In both cases, there are a lot of options for lexicalizing the meaning of *intensity* or *degree* or *scale* of the noun in the sense of the Lexical Function `Magn` (see the discussion in subsection 4.2.2). The frame "Rewards_and_punishments" has a corresponding FE DEGREE that accommodates those collocates. In contrast, the frame "Commerce_scenario" evoked by the LU *price* does not have a direct FE corresponding to this meaning. The only FE that describes price is RATE, and thus the collocates were assigned to it, however, it is important to bear in mind, that this option is not optimal and, perhaps, an additional FE should be introduced.

Thema 'issue/topic' Frame "Point of dispute"
STATUS: *groß* 'big/major', *heiß* 'hot', *beherrschend* 'dominant', *wichtig* 'important', *zentral* 'central'
TIME: *drängend* 'pressing', *brennend* 'urgent', *aktuell* 'current', *ewig* 'eternal'
GROUP: *unbequem* 'uncomfortable', *leidig* 'vexed', *heikel* 'delicate', *sensibel* 'sensitive', *brisant* 'controversial'
DOMAIN: *sperrig* 'unwieldy'
DESCRIPTOR: : *interessant* 'interesting'
`restricted to a particular locality or neighborhood`: *lokal* 'local'
Other FEs within the frame: QUESTION (core), CONTEXT, VIEWPOINT.

The noun *issue* remains a challenge for this framework as well. The evoked frame "Point_of_dispute" offers suitable FEs, however, the interpretation may seem slightly far-fetched. It is unavoidable due to the abstractness level of the frame and the lexical units involved. Consider the following collocates: *unbequem* 'uncomfortable', *leidig* 'vexed', *heikel* 'delicate', *sensibel* 'sensitive', *brisant* 'controversial'. These adjectives describe an issue/topic from someone's point of view: a *sensitive topic* implies that someone is sensitive about the topic, a *controversial issue/topic* is an issue on which different opinions are expressed by different people. The frame element GROUP is defined as "the persons or organizations who have different points of view on the Question" and thus is a suitable FE for annotating this group of collocates. Another group of collocates includes the adjectives *groß* 'big/major', *heiß* 'hot', *beherrschend* 'dominant'. The FE STATUS "used for expressions that indicate the relative importance of a Question with respect to the other topics that are disputed" describes the semantics of this group of collocates very precisely.

Ähnlichkeit 'similarity'
Frame "Similarity"

DIMENSION (C): *äußerlich* 'external', *beliebig* 'arbitrary'

DEGREE: *groß* 'strong', 'slight', *stark* 'strong', *leise* 'slight'

MANNER: *frappierend* 'striking', *frappant* 'striking', *erstaunlich* 'surprising', *unheimlich* 'uncanny', *exakt* 'precise', *vage* 'vague'

TIME: *flüchtig* 'fleeting'

`that changes often`: *variabel* 'variable'

Other FEs within the frame: DIFFERENTIATING FACT (core), ENTITIES (core), ENTITY 1 (core), ENTITY 2 (core), CIRCUMSTANCES, DEPICTIVE, EXPLANATION, PLACE.

Anzahl 'amount'
Frame "Quantity"

DESCRIPTOR: *stattlich* 'stately', *groß* 'large', *hoch* 'high', *schier* 'sheer', *beträchtlich* 'large', *riesig* 'huge', *erklecklich* 'considerable', *bedeutend* 'substantial', *bescheiden* 'modest', *gering* 'small', *klein* 'small', *überschaubar* 'manageable'

Other FEs within the frame: ENTITY (core), QUANTITY (core), VALUE (core).

Ziel 'goal' Frame "Purpose"
AGENT(C): *selbstgesteckt* 'self-imposed', *gemeinsam* 'joint'
TIME: *kurzfristig* 'short-term', *langfristig* 'long-term'
`that can be achieved with a lot of ambition`: *hochgesteckt* 'ambitious', *hoch* 'high', *ehrgeizig* 'ambitious'
`highest in degree or order`: *ober* 'ultimate', *groß* 'great'
`accurately stated or described`: *klar* 'clear'
`characterized by determination`: *fest* 'firm'
`of foremost importance or value`: *primär* 'primary'
`serving as an essential component`: *zentral* 'central', *wichtig* 'important'
`sublime, awe-inspiring`: *hehr* 'noble'
Other FEs within the frame: ATTRIBUTE (core), GOAL (core), MEANS (core), VALUE (core), DOMAIN, RESTRICTOR.

The collocates of both *Ähnlichkeit* 'similarity' and *Anzahl* 'amount' could be accommodated by the FEs in the corresponding frames. The noun *Ziel* 'goal', on the other hand, poses a challenge. The evoked frame "Purpose" does not have any FEs suitable for adjectival modifiers, and in that sense is similar to the previously discussed frame "Calendric unit".

Angst 'fear' Frame "Fear"
DEGREE: *groß* 'strong/big', *tief* 'deep', *höllisch* 'hellish', *furchtbar* 'tremendous'
CIRCUMSTANCES:*krankhaft* 'pathological', *unterschwellig* 'subconscious'
MANNER: *nackt* 'pure', *pur* 'pure', *blank* 'sheer', *diffus* 'vague', *dumpf* 'vague'
TIME: *ständig* 'constant', *ewig* 'eternal', *stet* 'constant'
Other FEs within the frame: EXPERIENCER (core), EXPRESSOR (core), STATE (core), STIMULUS (core), EMPATHY TARGET, EXPLANATION, PARAMETER.

Interesse 'interest' Frame "Emotion directed"
EXPERIENCER(C): *ureigen* 'own', *widerstreitend* 'conflicting'
PARAMETER: *breit* 'broad', *handfest* 'tangible', *vital* 'vital', *elementar* 'fundamental'
MANNER: *lebhaft* 'lively', *rege* 'active', *echt* 'genuine', *wahr* 'genuine'
DEGREE: *hoch* 'strong', *groß* 'strong/big', *stark* 'strong', *massiv* 'massive', *gering* 'little'
Other FEs within the frame: EVENT (core), EXPRESSOR (core), REASON (core), STATE (core), STIMULUS (core), TOPIC (core), CIRCUMSTANCES, EMPATHY TARGET, FREQUENCY.

The last two LUs evoke frames of emotions: *interest* belongs to "Emotion_directed", and *fear* – to the frame "Fear". They both follow a similar pattern: the FEs DEGREE, MANNER, TOPIC, CIRCUMSTANCES. The interpretation of almost all the collocates is rather straightforward. However, there is a group of modifiers of *fear* that pose a challenge. Those are *krankhaft* 'pathological' and *unterschwellig* 'subconscious'. Strakatova and Hinrichs (2019) argue that the most suitable FE to accommodate these collocates is CIRCUMSTANCES. It is defined in FrameNet as follows:

> "The Circumstances is the condition(s) under which the Stimulus evokes its response. In some cases it may appear without an explicit Stimulus. Quite often in such cases, the Stimulus can be inferred from the Circumstances."

The choice of this FE is motivated by the fact that these kinds of fear are triggered by certain stimuli under certain circumstances which are only elaborated in the context. However, this additional meaning is very subtle. A possible solution could be annotating *krankhaft* 'pathological' and *unterschwellig* 'subconscious' with a more general FE such as MANNER thus combining the two collocates with the adjectives *blank* 'sheer', *pur* 'pure', *nackt* 'pure', *diffus* 'vague', *dumpf* 'vague'. The Frame Element Manner is ubiquitous throughout FrameNet and requires a more thorough discussion.

FEs Descriptor and Manner

All the FEs discussed so far have very specific definitions as they describe particular aspects of the situations. However, there are two peripheral FEs that have very generic definitions: MANNER and DESCRIPTOR. Note that they basically have the same role, the only difference is that DESCRIPTOR belongs to frames evoked by entities, whereas MANNER is found in processes, events, and states. In both cases, the FE describes some quality of an entity or some way of performing an action. Consider the definitions of the two FEs found in the frames evoked by the sample nouns discussed above.

Descriptor

(a) “Animals": A characterization of the Animal not covered by another frame element.

(b) “Biological area": Any description, characteristic, or property of the Locale which is not covered by more specific FEs.

(c) “Body parts": This Frame Element is used for a characteristic or description of the Body part. Note that Descriptor is not an FE of the same character as other FEs. Specifically, rather than filling a slot, it identifies a modifier of its head noun.

(d) “Clothing": This FE identifes any description of the clothing besides Style and Material.

(e) “Food": This FE indicates a characteristic or description of the food.

(f) “Intoxicants": Descriptor describes a quality of the Intoxicant.

(g) “Natural features": Any description, characteristic, or property of the Locale which is not covered by more specific FEs.

(h) “Plants": A characterization of the Plant not covered by another frame element.

(i) “Point of dispute": A characterization of the Question, often with regard to its complexity or whether it has been resolved or not.

(j) “Quantity": A characterization of the Quantity.

(k) “Sensation": A characterization of the sensation not covered by a more specific FE.

Manner

(l) “Commerce scenario": Manner of performing an action.

(m) “Emotion directed": The Manner is the way in which the Experiencer experiences the Stimulus.

(n) “Fear": Any description of the way in which the Experiencer experiences the Stimulus which is not covered by more specific FEs, including secondary effects (quietly, loudly), and general descriptions comparing events (the same way). Manner may also describe a state of the Experiencer that affects the details of the emotional experience.

(o) “Personal relationship": Manner of performing an action.

(p) “Precipitation": The way in which the precipitation falls.

(q) “Rewards_and_punishments": Manner of performing the reward or punishment.

(r) “Similarity": This FE identifies the Manner in which Dimension is similar or different.

A recurring pattern of defining this peripheral FE is "any description (quality, characteristic) not covered by other FEs in the frame". This implies that assigning a predicate to this role is only possible through an elimination process: when none of the other available FEs applies. On the one hand, the danger of having such a label in an inventory is the semantic heterogeneity of the predicates that may be assigned to that role. On the other hand, splitting this property into a number of more specific ones may lead to creating very fine-grained categories that can accommodate very few semantically related collocates – this will reduce the generalization power of the inventory. For the small data sample presented in this section, it does not seem to be a problem: the adjectives annotated as MANNER or DESCRIPTOR are rather homogeneous.

Conclusion

The case study shows that Frame Elements provided by FrameNet are rich enough to model relations between the constituents of adjective-noun phrases. They are especially suitable for describing the semantics of abstract nouns. However, there are some challenges concerned with using FrameNet database for annotation.

First of all, as Strakatova and Hinrichs (2019) point out, there are meaning aspects of some frames not reflected in the inventory of frame elements available for that frame. For instance, the LU *future* that evokes the frame "Alternatives" has a temporal property as in *near/distant future* and an evaluative property as in *bright/dark future*. However, none of the FEs reflect these two aspects of the noun's meaning. Another example is the frame "Calendric_unit" where several LUs such as *winter*, *day*, *month*, *weekend* have a climate-related aspect of meaning as in *a cold winter* or *a sunny day*. This aspect is not described by any of the FEs available for the frame.

Secondly, annotating a large amount of data using the English FrameNet as it has been done for this case study is costly and time-consuming. Finding the correct translations, identifying the corresponding frames, and searching for suitable frame elements within each frame requires a lot of effort. Moreover, there is always a possibility that some LUs are not present in the database. Such cases should either be excluded from the study or substituted by synonyms or semantically related words that are frame evoking. Therefore, although FrameNet is a conceptually suitable resource for modeling the semantic relations in AN phrases, it requires an adaptation for this task.

A possible solution would be creating a stable inventory of relations based on the most frequent frame elements from FrameNet. This approach implies skipping the step of matching the LUs with the frames they evoke and choosing a Frame Element directly from the list of all the available relations. The success of using this method depends on the amount of such relations. There are 1,285 Frame Elements in total available in FrameNet (release 1.5). As expected, they have a Zipfian distribution with a long tail: more than half of them are encountered only in one frame. In table 4.5,

Frame Element	N frames	in case study
Time	818	4
Place	685	-
Manner	663	6
Degree	479	6
Means	450	-
Purpose	402	-
Explanation	380	-
Duration	273	2
Depictive	266	1
Circumstances	263	1
Agent	209	1
Frequency	185	-
Descriptor	161	11
Result	157	-
Instrument	132	1
Entity	112	-
Source	102	-
Goal	92	-
Path	92	-
Theme	90	-

TABLE 4.5: Twenty most frequent Frame Elements in FrameNet.

twenty most frequent FEs are presented,[21] almost half of them have been used in the case study which indicates that this approach does provide adequate relations. However, it cannot be ruled out that some relations do not have high frequency in FrameNet, but are relevant in the context of adjective-noun phrases. This issue could be resolved by categorizing the rare FEs and grouping them together based on the semantic similarity. In such a bottom-up fashion, more coarse-grained categories based on Frame Elements could be created. Finally, a solution might be provided by another frame-based approach which will be presented in the next section.

4.5 Concept frames

FrameNet, which was extensively discussed in the previous section, is not the only approach to describing semantic knowledge in terms of frames. FrameNet frames are what Gamerschlag et al. (2014) call **predicative frames** following the terminology in Busse (2012). Predicative frames are flat and non-recursive, they describe the semantic valence of LUs in terms of participants of the evoked situation - Frame Elements. The second type of frames is termed **concept frames**. Gamerschlag et al. (2014, p.5) state that the main purpose of concept frames is to "represent attributes and properties of an entity", thus associating this type of representation mainly with nouns, whereas the predicative frames are mostly associated with verbs. This section will investigate the appropriateness of using concept frames, and more specifically attributes, for modeling the relations in adjective-noun collocations and free phrases.

Concept frames originate from the work by Minsky (1974) who proposed structures called "Frames" for representing situations. Minsky's frames have nodes and relations as their elements with "slots" (≈ attributes) and "fillers" (≈ values) on lower levels. Minsky's work on frames was influential in the development of artificial intelligence and cognitive science. The idea that knowledge in human cognition is represented in terms of concept frames was then advocated in Barsalou (1992). This claim was later adopted by Löbner (2014, 2021) as the "Frame Hypothesis". The Frame Hypothesis makes two strong assumptions: 1) "The human cognitive system operates with a single general format of representations." 2) "...this format is essentially Barsalou frames." (Löbner, 2014, pp.23-24).

In Barsalou's framework, concept frames are recursive structures consisting of three elements: *attribute-value sets*, *structural invariants*, and *constraints*. An **attribute** "describes an aspect of at least some category members" (Barsalou, 1992, p.30). For instance, the members of the category CLOTHING have the attributes `size, color, style, etc.` Attributes are further specified in terms of their **values**: for example, possible values for the attribute `color` are *yellow, blue, black, green, red, etc.* Structural invariants and constraints are relations between attributes of a frame. Miller (1998, p. 48) illustrates the concepts of attribute and value as follows:

[21] Appendix A presents 50 most frequent FEs in FrameNet.

> <...> to say *x is Adj* presupposes that there is an attribute A such that A(*x*)=*Adj*. To say *The package is heavy* presupposes that there is an attribute WEIGHT such that WEIGHT(*package*)=*heavy*. Thus, *heavy* and *light* are values for the attribute WEIGHT.

Discussing what properties can be viewed as attributes, Barsalou (1992) mentions that in many cases attributes denote *parts*. He also points out that other properties that are not parts of category members can be considered attributes as well, for instance `evaluation`, `quantity`, `costs`, `benefits`, etc. Figure 4.5 illustrates how a concept frame of attributes and their values may look for the category CLOTHING. In this representation, the category CLOTHING, depicted as the central node, has five attributes: `sleeve`, `size`, `style`, `material`, and `purpose`. The first attribute `sleeve` is a *part of* attribute which, in turn, has two attributes `color` and `length` that take up the values *blue* and *short* respectively. The next two attributes of the central node, `size` and `style`, also take up adjectives as their values: *large* and *casual* respectively. The attribute `material` takes a noun as its value – *wool*. Finally, the attribute `purpose` is what Löbner (2021) calls a FOR attribute – an *affordance* attribute. The term "affordance" comes from the Theory of affordances developed by Gibson (1979) in the field of visual perception. An affordance is what an environment *provides* for an animal, for instance, a horizontal, flat and large enough surface *affords support* (Gibson, 1979, p.127). In the same way, clothing affords staying warm. Integration of affordances into frames is discussed in Attardo (2005) who emphasizes that affordances usually refer to the purpose of an object. Thus, artifacts are a very prominent semantic category when it comes to affordances because of their inherent quality of being created for a purpose.[22]

In cognitive science, the idea of representing conceptual knowledge as frames competes with feature list representations (Barsalou, 1992, 2017). For illustration, see Figure 4.6 that depicts a simplified feature representation of the category CLOTHING. In this case, there is no implicit frame structure containing abstract attributes, only a list of independent features such as *blue sleeves*, *keeps warm*, *large*, and so on.

[22] Werning (2008) also suggests introduction of affordances into frame theory in the context of the theory of neuro-frames.

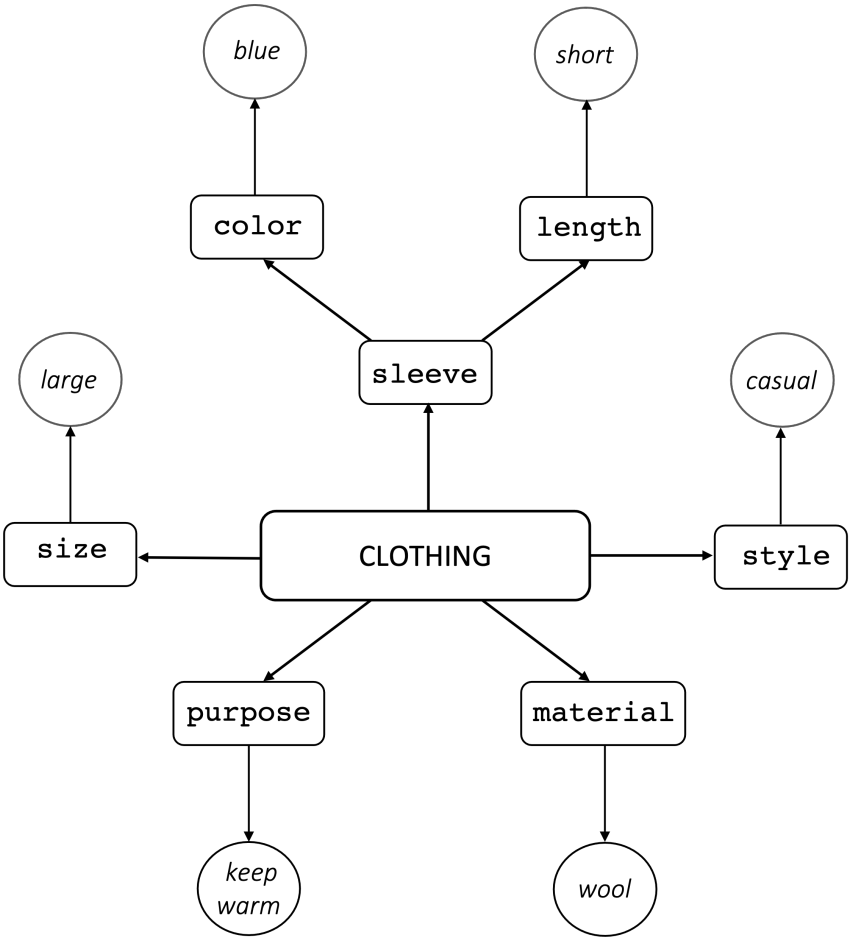


FIGURE 4.5: Example of a simplified frame representation of the category CLOTHING in terms of attribute-value sets.

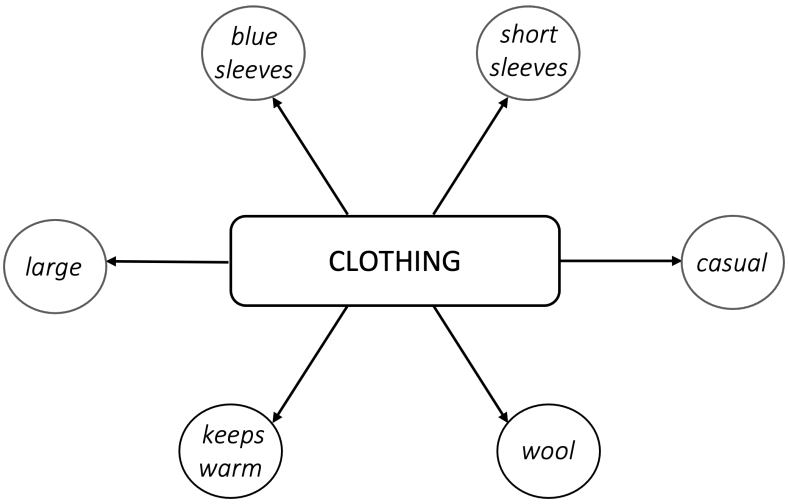


FIGURE 4.6: Example of a simplified feature representation of the category CLOTHING.

Barsalou (1992) presents several findings from cognitive science supporting the Frame Hypothesis. He discusses several studies that suggest that people and animals discriminate between different categories on the basis of attributes rather than on the basis of independent features. For instance, animals learn that an attribute is a signal for receiving an award. When the value of this attribute changes, animals would still pay attention to the learnt attribute despite the change of its value. However, if the new stimulus is a value of a different attribute, the learning process is longer. People also tend to make generalizations over specific values to discriminate between different categories.

Löbner (2014) provides an extensive overview of evidence for frames found in natural language. For instance, he states that syntactic structures, both dependency or constituency, are basically frames. Grammatical features such as case, number and tense can be seen as attributes with unique values – the approach taken in Lexical-Functional Grammar (LFG) (Bresnan, 2001) and Head-Driven Phrase Structure Grammar (HPSG) (Pollard and Sag, 1994). Apart from those two aspects, Löbner (2014, p.42) discusses verbs being "the most *framish* elements of of language" in the context of semantic roles. In a later publication, Löbner (2021) brings into discussion examples of using frame structures in every day life by pointing out that various information in the society is organized and stored in databases in the form of attribute-value sets: names, surnames, dates and places of birth, etc.

One of the major characteristics of concept frames is their recursiveness, which implies that any attribute can exhibit a frame structure of its own and can be further specified in terms of other attributes. This is illustrated in Figure 4.7. The frame for the members of the category CLOTHING has attributes such as `size, color, style` and each of them can be specified by the corresponding values: e.g. *large, blue, tailored* accordingly. However, the attribute `color` has properties typical of colors: `brightness, hue, saturation`, and the latter one can be further described in terms of its `intensity`.

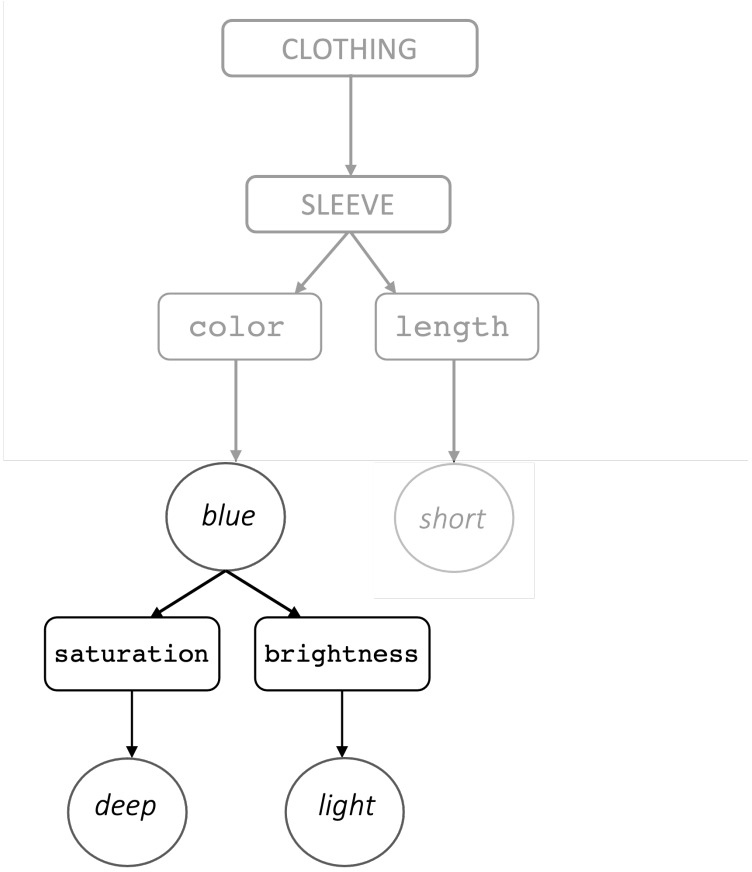


FIGURE 4.7: Frame structure for *blue* embedded in the frame for SLEEVE that is an attribute of CLOTHING .

Naturally, frame recursiveness raises the question of whether it is possible to arrive at a terminal level of frame representation. According to Barsalou (1992), this is not the case: attribute construction is a highly productive and creative process.

Depending on the context, people may come up with new attributes suitable for that particular context and, theoretically, there is an endless amount of possible attributes. However, this process has constraints such as people's experience, goals, and intuition. Moreover, similarly to Fillmore's theory of core and peripheral Frame Elements, Barsalou (1992) acknowledges the existence of *core attributes*. Those are attributes that are defining for a particular frame: without them the understanding of the frame would not be possible. For instance, to understand the concept of *buying*, one needs the attributes *buyer*, *seller*, *merchandise*, and *payment*.[23] As for the concepts that can be represented in terms of frames, there are no limitations: according to Barsalou (1992), frames represent all kinds of concepts including physical objects and events and attributes can represent any property of those concepts.

Hartung (2015, p.17) states that adjective-noun phrases are "*ubiquitous sources of attribute knowledge in natural language*". Therefore, concept frames,

[23]Compare the core FEs in FrameNet for the frame "Commerce_scenario": BUYER, GOODS, MONEY, SELLER.

and more specifically attributes, seem to be a suitable approach for modeling adjective-noun collocations and free phrases.

4.5.1 Attributes

Attributes in adjective-noun phrases can be compared to the *qualia roles* from the Generative Lexicon and be seen as fine-grained specifications of the qualia roles. The process that Pustejovsky (1995) refers to as *selective binding* (see section 4.3) involves an interaction between an adjective and a noun, where the adjective selects a certain property (=attribute) of the noun. Hartung (2015) terms that process *attribute selection* and investigates this topic within the framework of distributional semantics. He emphasizes that attributes in adjective-noun phrases are implicit, since they are not overtly realized on the textual surface.

(25) a. *white dress*: the **color** of the **dress** is **white**
b. *small dress*: the **size** of the **dress** is **small**
c. *beautiful dress*: *the **evaluation** of the **dress** is **beautiful**

Nevertheless, there are examples where the attribute is explicitly mentioned, and thus it is possible to encounter triples of the type "attribute noun-head noun-adjective", as in examples (25a) and (25b). However, a much more common linguistic pattern is just the adjective-noun pair without mentioning the attribute explicitly. Thus in most cases, the knowledge that *white* refers to the `color` of the *dress* and that *small* refers to its `size` remains implicit. There are attributes that cannot be expressed by adjective-noun-attribute triples in that way. Any attempt to do so results in an ill-formed phrase as illustrated by the example (25c), where the well-formed paraphrase should rather be "the **evaluation** of the **dress** is **positive**". Nevertheless, in spite of the absence of the corresponding paraphrase in the natural language, it is clear that the implicit attribute in the phrase *beautiful dress* is `evaluation`.

The inventory of attributes used in Hartung (2015) to create the gold standard dataset HeiPLAS[24] stems from the Princeton WordNet (Fellbaum, 1998). The attribute information about the adjectives is represented in WordNet via the *attribute* relation. The *attribute* pointer links the adjectives to the synsets of the corresponding attribute nouns. Figure 4.8 illustrates this approach: the adjective *high* in its first sense is linked to the synset containing the attribute nouns *degree, grade, level*. In its second sense, it is connected to the nouns *height, tallness*. Most entries of adjectives in WordNet contain a few examples of adjective-noun pairs illustrating each sense of the adjective as illustrated in Figure 4.8.

Retrieving the attribute information from WordNet is not always straightforward because of the way the adjectives are organized there. Adjective meaning is represented in a bipolar or dumbbell structure based on antonymy (Miller, 1998) as depicted in figure 4.9. For sense 2 of *high*, the

[24]Heidelberg Princeton Large-scale Attribute Selection

- S: (adj) **high#1** (greater than normal in degree or intensity or amount) *"a high temperature"; "a high price"; "the high point of his career"; "high risks"; "has high hopes"; "the river is high"; "he has a high opinion of himself"*
  - see also
  - similar to
  - **attribute**
    - S: (n) degree, grade, level (a position on a scale of intensity or amount or quality)
- S: (adj) **high#2** ((literal meaning) being at or having a relatively great or specific elevation or upward extension) *"a high mountain"; "high ceilings"; "high buildings"; "a high forehead"; "a high incline"; "a foot high"*
  - see also
  - similar to
  - **attribute**
    - S: (n) height, tallness (the vertical dimension of extension; distance from the base of something to the top)

FIGURE 4.8: WordNet entries for the first two senses of the adjective *high* with the attribute pointers to the corresponding noun synset

antonym is *low* – these adjective are the *anchors*. Adjectives that are semantically similar to the anchors are linked to them via the *similar to* relation: *steep, upper, lofty, deep, nether, underslung*. Direct attribute pointers are only available for the anchors. Therefore, for a large number of adjectives, the corresponding attributes can only be determined indirectly: through the *similar to* relation to the closest anchor adjective (Hartung, 2015).

Relying on this information, Hartung (2015) compiles the HeiPLAS dataset: the adjective-noun pairs come from the examples in the adjectival entries and are labeled with the corresponding attributes directly or via the *similar to* relation. In order to validate the resulting triples, manual assessment has been undertaken. The annotators were asked to judge whether an attribute refers to the implicit meaning in a given adjective-noun pair (Hartung, 2015, p.101):

(26) Can you say that [*AN phrase*] makes a statement about the [`attribute`] of the [*head noun*]?

E.g.: Can you say that *white dress* makes a statement about the `color` of the *dress*?

The template from the annotation guidelines is presented in example (26) . The inter-annotator agreement for the task is quite low: Fleiss' kappa κ=0.28 (Fleiss, 1971) which indicates that the task is highly challenging. The final HeiPLAS dataset contains only the phrases with full agreement. The dataset is split into development and test set with 919 and 737 phrases accordingly. A major advantage of the HeiPLAS dataset is its lexical and semantic diversity. The dataset contains a large number of unique lexical units: both adjectives and nouns. This lexical diversity ensures a broad coverage of different semantic classes.

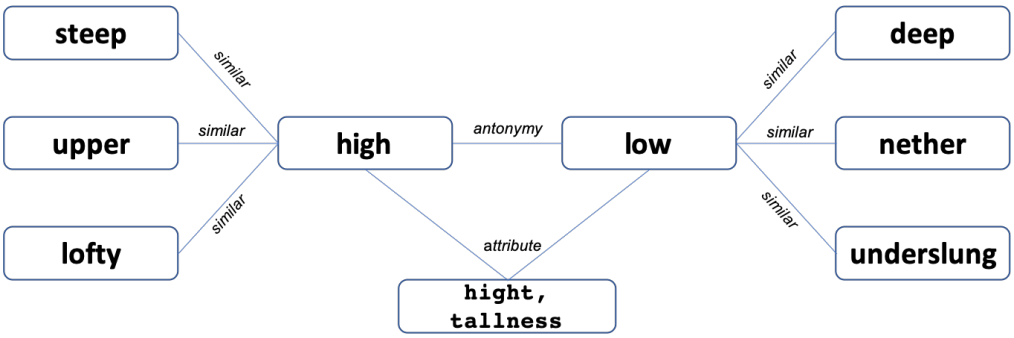


FIGURE 4.9: Modeling of the adjectives *high* and *low* in WordNet.

The total number of attributes in HeiPLAS is 253 – thus the inventory is very fine-grained. Table 4.6 provides some examples from the dataset with the corresponding attributes. The top-5 triples in the table are typical examples of attributes in adjective-noun phrases that are usually listed in literature on this topic along with `shape`, `temperature`, `speed` and `taste`. The rest of the examples in the table have been selected from the dataset to illustrate the sources of such fine granularity of the inventory:

1. There are a lot of attribute nouns derived from the adjectives in the phrases such as `commonness` in *common knowledge*, `niceness` in *nice dress*, `staleness` in *stale bread*, etc. In comparison to the more prototypical attributes such as `size` or `color`, they are less intuitive when referring to the properties of nouns.

2. The second issue is closely linked to the first one: the deadjectival attributes (or at least some of them) are way too specific – for instance, the attribute `staleness`. Another example is the attribute `gluttony` in *gluttonous husband*.

3. Another source of the fine granularity of the attributes is the polarity of adjectives. For instance, the phrase *easy job* is labeled with the attribute `ease` whereas *tough job* – with the attribute `difficulty`. The polarity

differentiation is not consistent: for example both *dirty overall* and *clean house* are annotated with the same attribute `cleanness`.

adj-nn pair	attribute
young people	`age`
white dress	`color`
bad headache	`degree`
short story	`duration`
low ceiling	`height`
small car	`size`
hot topic	`emotionality`
friendly host	`friendliness`
common knowledge	`commonness`
nice dress	`niceness`
stale bread	`staleness`
active interest	`activeness`
gluttonous husband	`gluttony`
clean house	`cleanness`
dirty overall	`cleanness`
easy job	`ease`
tough job	`difficulty`

TABLE 4.6: Examples of adjective-noun-attribute triples from the HeiPLAS dataset (Hartung, 2015)

The examples illustrate that the attributes in WordNet are extremely fine-grained. A logical solution to this problem would be merging the most similar categories into more coarse-grained ones, thus creating an inventory of relations in a bottom-up fashion. Another way of tackling the granularity issue is turning to other lexical resources that provide attribute information about lexical items. The data sample used for the case studies in this thesis has been compiled relying on the German wordnet GermaNet (Hamp and Feldweg, 1997; Henrich and Hinrichs, 2010). Thus in the next step, it is investigated whether and how GermaNet provides attribute information about adjectives.

Attributes in GermaNet

By and large, the structure of GermaNet is very similar to the Princeton WordNet, but a few aspects are different. One of the main differences between the two wordnets lies in the organization of the adjectives:[25] GermaNet does not use antonymy as the main relation between adjectival synsets, but structures the adjectives hierarchically. In that sense, the approach to modeling of adjectives follows the approach taken for the nouns and verbs: it is based on the *hyponymy* relation. The hierarchy of

[25] `https://uni-tuebingen.de/en/142852` (*last accessed on 14.12.2022*).

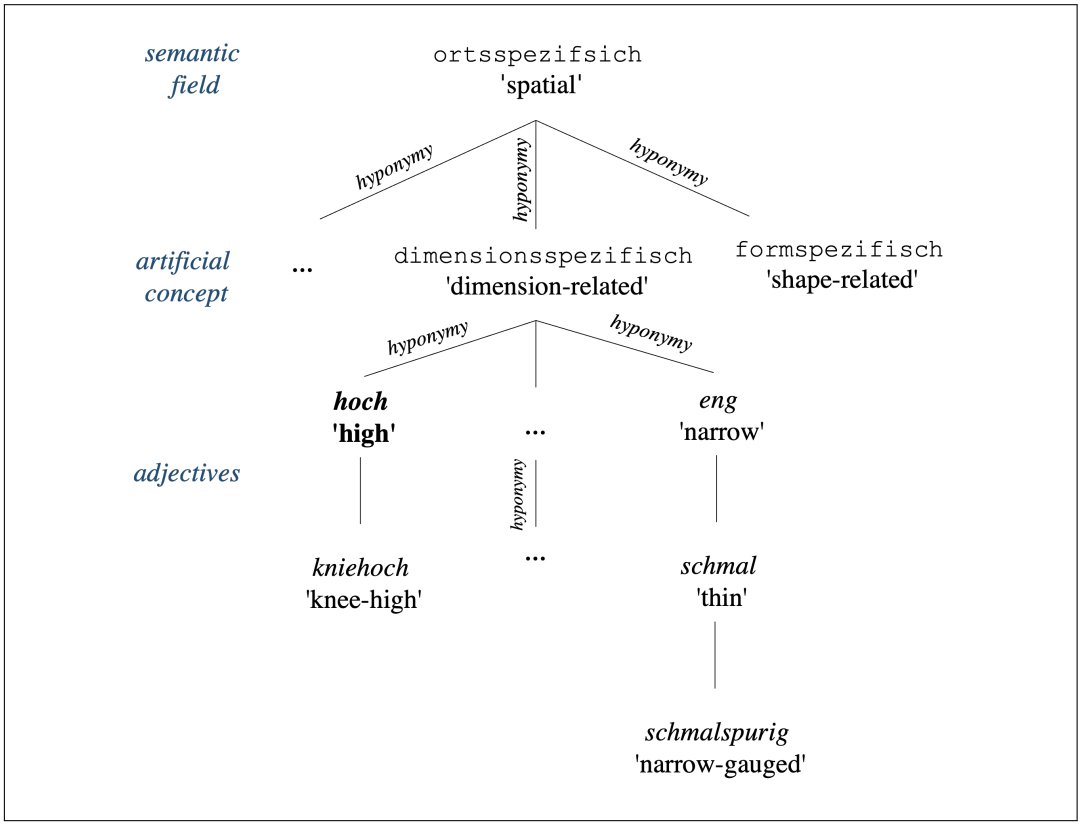


FIGURE 4.10: A part of the GermaNet taxonomy of spatial adjectives *hoch* 'high' and *eng* 'narrow'. The top node `spatial` defines the semantic class of all its hyponyms - *Ort* 'spatial'. The second from the top level is represented by the *artificial* concepts: `dimension-related, shape-related` and others.

many adjectives is quite flat (Figure 4.10), but there are also examples of deep hyponymy connections (Figure 4.11).

The information about attributes is not directly encoded in GermaNet as it has been done in the Princeton WordNet. There are no pointers to attribute nouns, but similar information is integrated directly into the hierarchical structure of adjectives. Consider two examples of adjectival modeling in GermaNet in Figures 4.10 and 4.11. Figure 4.10 depicts the modeling of the literal sense of the adjective *hoch* 'high'. Figure 4.11 illustrates the structure for the adjective *schlau* 'smart'. The top nodes `spatial` and `spirit-related` represent different semantic classes comprising in total 16 categories. Each of the 16 semantic classes is further specified in terms of more fine-grained categories referred to as *artificial concepts*. In the hierarchy, the artificial concepts are just under the top nodes: `direction`, `dimension`, `shape`, `knowledge`, `intelligence`. There are 80 artificial concepts in total in GermaNet. These semantic categories had been adopted in GermaNet from the semantic classification of German adjectives developed by Hundsnurscher and Splett (1982).

The artificial concepts in GermaNet bear a strong resemblance to attributes: *dimensionsspezifisch* 'dimension-related' – `dimension`, *intelligenzspezifisch*

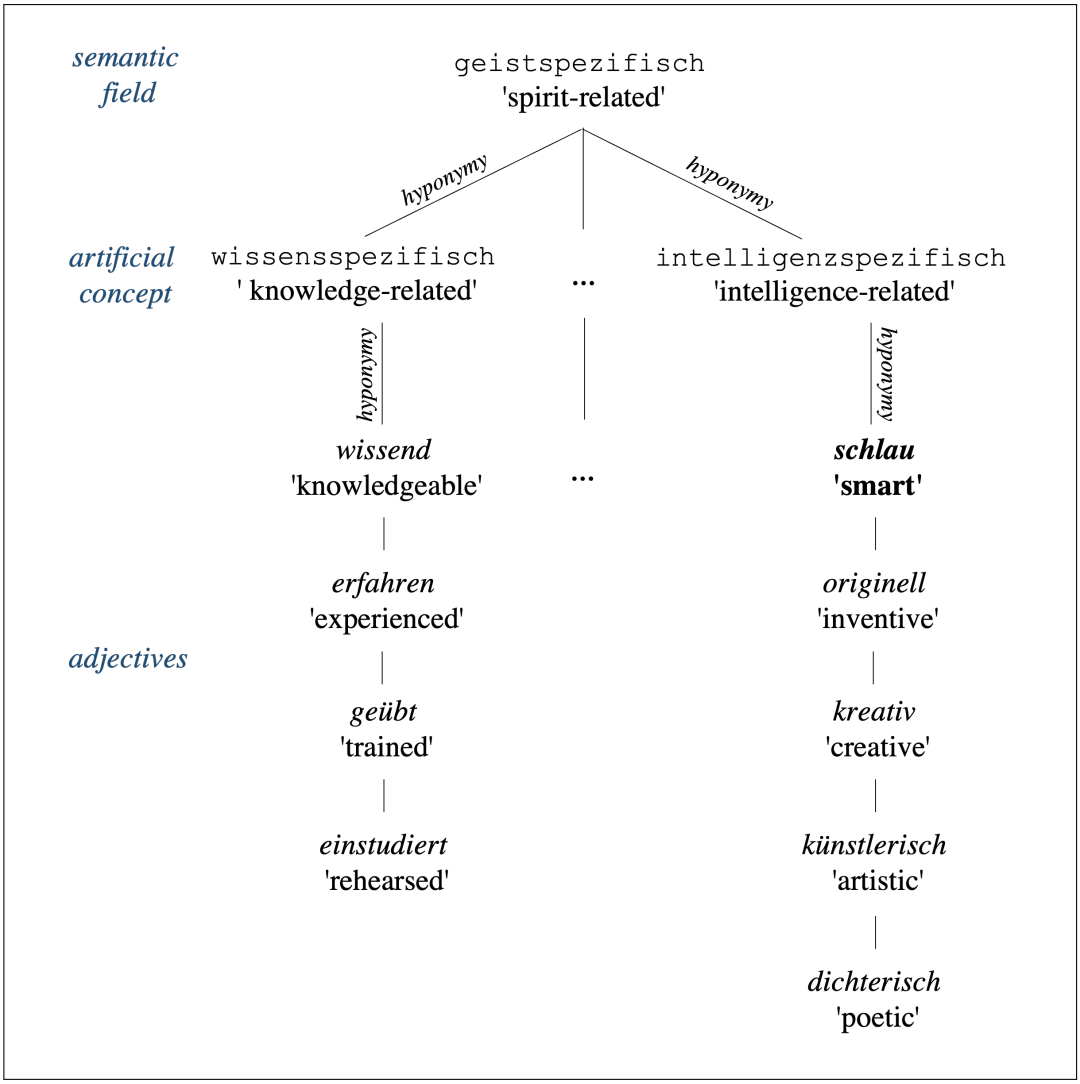


FIGURE 4.11: A part of the GermaNet taxonomy of spirit-related adjectives *wissend* 'knowledgeable' and *schlau* 'smart'. The top node `spirit-related` defines the semantic class of all its hyponyms - *Geist* 'spirit'. The second from the top level is represented by the *artificial* concepts: `knowledge-related` and `intelligence-related` and others.

'intelligence-related' – `intelligence`. An overview of the semantic classes and the more fine-grained artificial concepts is presented in table 4.7. Thus, the artificial concepts from the hierarchy of adjectives in GermaNet can be treated as attributes and their adjectival hyponyms are the possible values those attributes can take in German. The artificial concepts will further be referred to as attributes. The next section presents the results of modeling the sample of AN pairs with the attributes based on the artificial concepts in GermaNet combined with the relations from the other three inventories investigated in the previous sections of this chapter.

4.5.2 Case study: Attributes+

The procedure for modeling the data consists of the following steps:

1. look up the adjectival collocate in GermaNet - because attribute information in GermaNet is available for adjectives, but not for nouns;
2. select the corresponding sense;
3. extract the artificial concept from the hierarchy, turn it into attribute by removing *spezifisch* 'specific/related' from its name. The attribute names in the study will be given in English.

The case study in this section is slightly different from the three previous ones in sections 4.2.2, 4.3.2, and 4.4.2. The "plus" in the title of this section suggests that not only the attributes will be used for modeling the data. This section presents an attempt to combine the findings of the previous three case studies and to resolve the existing issues by integrating the categories from different inventories. Therefore, in all the cases that are not covered by the attributes inventory obtained from GermaNet, other inventories will be used. Each successfully modeled phrase will be marked accordingly:

(A) – Attribute;
(LF) – Lexical Function;
(Q) – Qualia Role;
(FE) – Frame Element

Schokolade 'chocolate'
(Q/FE) TYPE: *schwarz* 'dark', *dunkel* 'dark', *weiß* 'white', *heiß* 'hot'
(A) `evaluation`: *edel* 'premium, fine', *hochwertig* 'of high quality'
(A) `taste`: *lecker* 'tasty', *süß* 'sweet'
(A) `consistency`: *flüssig* 'liquid', *knackig* 'crunchy'

sem class	artificial concepts = attributes
Perceptional	Helligkeit 'brightness', Farbe 'color', Geräusch 'sound', Geschmack 'taste', Geruch 'smell', Oberfläche 'surface', Sicht 'sight'
Spatial	Dimension 'dimension', Richtung 'direction', Lokalisierung 'localisation', Herkunft 'origin', Raumverteilung 'spatial distribution', Form 'shape', Existenz 'existence'
Temporality-related	Zeit 'time', Geschwindigkeit 'speed', Alter 'age', Gewohnheit 'habit', *Dauer* 'duration'
Motion-related	Bewegung 'movement'
Material-related	Beschaffenheit 'texture', Zustand 'state', Stabilität 'stability', Konsistenz 'consistency', Reife 'ripeness', Feuchtigkeit 'dampness', Reinheit 'purity', Gewicht 'weight', Physik 'physics', Chemie 'chemistry', Temperatur 'temperature'
Weather-related	Klima 'climate', Landschaft 'landscape'
Body-related	Belebtheit 'life', Konstitution 'constitution', Krankheit 'affliction', Geschlecht 'sex', Körpergefühl 'desire/feeling', Aussehen 'appearance', Körperzustand 'bodily state'
Mood-related	Gefühl erweckend 'emotion-stimulus', Gefühl empfindend 'emotion-experienced'
Spirit-related	Intelligenz 'intelligence/attention', Wissen 'knowledge/experience'
Behaviour-related	Charakter 'behaviour/character', Tier 'animal specific', Geschicklichkeit 'skill', Beziehung 'relations', Sympathie 'sympathy', Neigung 'inclination'
Social-related	Schicht 'stratum', Institution 'institution', Religion 'religion,' Land 'state,' Region 'region,' Aufgabe 'role'
Quantity-related	Zahl 'number,' Menge 'quantity,' Kosten 'costs,' Ertrag 'return'
Relational	Geltung 'validity,' Gewissheit 'certainty,' Bedarf 'requirements,' Wirksamkeit 'effectiveness,' Aufwand 'difficulty/energy requirement',' Funktionalität 'functioning,' Sicherheit 'security,' Geordnetheit 'order,' Verknüpfung 'linking,' Übereinstimmung 'correspondence,' Genauigkeit 'accuracy,' Vollständigkeit 'completeness,' Ursache 'cause,' Bezug 'reference', Zuträglichkeit 'tolerability'
General	Steigerung 'intensity', Bewertung 'evaluation', Norm 'norm'
Pertainyms	'derived from'
Privative	defined by negation

Table 4.7: Semantic classification of adjectives in GermaNet: it consists of 16 semantic classes and 80 artificial concepts that are direct hyponyms of the semantic class nodes.

Schuh 'shoe'
(Q/FE) TYPE: *hochhackig* 'high-heeled', *flach* 'flat'
(A) `emotion-stimulus`: *bequem* 'comfortable'
(A) `appearance`: *elegant* 'elegant'
(A) `cost`: *teuer* 'expensive'
(A) `color`: *braun* 'brown', *schwarz* 'black', *rot* 'red', *weiß* 'white', *rosa* 'pink', *blau* 'blue'
(A) `composition`: *seiden* 'silk'
(A) `cleanliness`: *schmutzig* 'dirty'
(A) `shape`: *spitz* 'pointy'
(A) `wetness`: *naß* 'wet'

In order to model the meaning of *is_a* as in *dark chocolate is_a* (type of) *chocolate*, the label TYPE is used. The name is borrowed from the FrameNet, but conceptually it is closer to the FORMAL qualia role, as it is slightly more general than the FE TYPE.[26] The same holds true for two modifiers of the noun *Schuh hochhackig* and *flach*.

Fels 'cliff'
(FE/Q) CONSTITUENT_PARTS/CONSTITUTIVE: *nackt* 'bare', *blank* 'bare'
`color`: *schwarz* 'black', *rötlich* 'reddish', *grau* 'grey'
(A) `wetness`: *naß* 'wet'
(A) `surface`: *glatt* 'smooth'
(A) `dimension`: *schroff* 'craggy', *steil* 'steep'
(A) `consistency`: *hart* 'hard'

For the noun *Fels*, there is no suitable attribute only for the for the meaning of 'bare'. Instead, a corresponding FE or Qualia Role can be applied.

[26] Remember that for *hot chocolate* a different FE has been used in section 4.4.2.

Hund 'dog'
(A) `evaluation`: *süß* 'sweet', *niedlich* 'cute'
(A) `emotion`: *arm* 'poor'
(A) `character`: *scharf* 'vicious', *bissig* 'biting', *treu* 'loyal', *aggressiv* 'agressive'
(A) `dimension`: *riesig* 'huge', *klein* 'small'
(A) `intelligence`: *klug* 'smart', *dumm* 'stupid'
(A) `age`: *jung* 'young'
(A) `appearance`: *dick* 'fat'
(A) `health`: *tollwütig* 'rabid'

Wald 'forest'
(Q/FE) TYPE: *immergrün* 'evergreen'
(A) `dimension`: *tief* 'deep'
(A) `quantity`: *licht* 'sparse', *dicht* 'dense'
(A) `evaluation`: *herrlich* 'magnificent'
(A) `dimension`: *riesig* 'huge', *endlos* 'endless'
(A) `climate`: *feucht* 'humid', *trocken* 'dry'
(A) `location`: *nah* 'close'
`in a natural state`: *wild* 'wild', *unberührt* 'untouched'

Baum 'tree'
(Q) CONSTITUTIVE (*has_as_part*): *kahl* 'leafless'
(Q/FE) TYPE: *immergrün* 'evergreen'
(A) `dimension`: *mächtig* 'mighty, massive', *riesig* 'huge', *hoch* 'high'
(A) `evaluation`: *prächtig* 'magnificent'
(A) `age`: *alt* 'old', *jung* 'young'
(A) `health`: *krank* 'sick', *gesund* 'healthy'
(A) `life`: *tot* 'dead'
(A) `consistency`: *morsch* 'rotten'
(A) `religion`: *heilig* 'sacred'

Bein 'leg'
(FE) DESCRIPTOR: *staksig* 'spindly'
(A) `dimension`: *lang* 'long', *kurz* 'short'
(A) `appearance`: *stämmig* 'sturdy',*dick* 'fat', *schlank* 'slim', *dünn* 'thin'; *nackt* 'naked'
(A) `health`: *gesund* 'healthy'
(A) `constitution`: *schwach* 'weak', *wacklig* 'shaky'
(A) `sensation`: *müde* 'tired', *schwer* 'heavy'

Regen 'rain'
(Q/FE) TYPE: *sauer* 'acid'
(FE) MANNER: *strömend* 'pouring'
(A) `degree`: *stark* 'heavy', *kräftig* 'heavy', *heftig* 'heavy', *leicht* 'light', *sintflutartig* 'torrential'
(A) `consistency`: *fein* 'fine'
(A) `quantity`: *spärlich* 'sparse'
(A) `duration`: *ewig* 'endless', *endlos* 'endless'
(A) `temperature`: *warm* 'warm', *kalt* 'cold', *kühl* 'cool'

The phrases with the noun *Hund* 'dog' as their base have been successfully modeled with the attributes from GermaNet. In contrast, the modeling of the nouns *Wald* 'forest', *Baum* 'tree', *Bein* 'leg', and *Regen* 'rain' is more diverse.

For the noun *Regen*, a suitable attribute for the meaning of intensification is available in GermaNet under the artificial concept *steigerungsspezifisch* 'comparison-related'. It will be further referred to as attribute `degree`The label TYPE borrowed from the qualia roles and Frame Elements has been used for the phrases *sauerer Regen*, *immergrüner Wald/Baum*. Finally, there is no suitable attribute in GermaNet that denotes a *manner* of an action, and thus the FE MANNER has been used for the collocate *strömend* 'pouring'.

Freund 'friend'
(A) `duration`: *alt* 'old'
(A) `degree`: *eng* 'close', *dick* 'close', *nah* 'close', *gut* 'good'
(A) `character`: *treu* 'faithful', *verlässlich* 'reliable'
(A) `evaluation`: *wahr* 'true', *echt* 'real', *falsch* 'fake'
(A) `social class`: *mächtig* 'powerful', *reich* 'rich'

Attributes provide suitable labels for all the collocates of the noun *Freund* 'friend'. There is one exception: the adjectives *eng* and *dick* in the meaning 'close' are modeled in GermaNet as "relationship-related". That is correct, but it does not specify which aspect of the relationship is addressed by those adjectives. In combination with *friend*, both adjectives intensify the meaning of the noun and express the `degree` of friendship. For the adjective *alt* 'old', both its meanings can be modeled in this approach: both the sense of `duration` and the sense of `age` are reflected in GermaNet.

Winter 'winter'
(A) `temperature`: *streng* 'severe', *bitterkalt* 'bitter cold', *eiskalt* 'ice cold', *warm* 'warm', *kalt* 'cold', *kühl* 'cool'
(A) `difficulty/energy requirement`: *hart* 'harsh', *mild* 'mild'
(A) `time`: *tief* 'deep', *letzt* 'last'
(A) `climate`: *feucht* 'humid', *trocken* 'dry'
(A) `duration`: *lang* 'long', *ewig* 'endless'
(A) `emotion stimulus`: *trist* 'dull, depressing'

All the collocates of the noun *Winter* 'winter' were successfully accommodated by the attributes from GermaNet.

Droge 'drug'
(Q/FE) Type: *hart* 'hard', *weich* 'soft', *leicht* 'soft', *sanft* 'soft'
(Q) Agentive (*origin*): *körpereigen* 'endogenous'
(A) `effect` or `safety`: *schädlich* 'harmful', *harmlos* 'harmless', *gefährlich* 'dangerous', *tödlich* 'lethal'
(A) `consistency`: *flüssig* 'liquid'
(A) `cost`: *billig* 'cheap'
(A) `validity`: *illegal* 'illegal', *legal* 'legal'

GermaNet does not model the very specific meaning of five collocates available for the noun *Droge* 'drug'. For four of them, similarly to *chocolate* and *shoe*, the most suitable category is provided by the FORMAL Qualia Role extension or Frame Element Type. For the adjective *körpereigen*, GermaNet suggests the attribute `linking` (Ger. 'Verknüpfung'). However, the AGENTIVE qualia dimension specified as *origin* seems more suitable.

Team 'team'
(A) `functioning`: *stark* 'strong', *schlagkräftig* 'powerful', *schwach* 'weak'
(A) `social aspect`: *eingespielt* 'coordinated and experienced'
(A) `knowledge`: *erfahren* 'experienced'
(A) `evaluation`: *toll* 'excellent', *gut* 'good'
(A) `correspondence`: *homogen* 'homogenous'
(A) `country`: *spanisch* 'Spanish', *deutsch* 'German'
(A) `quantity`: *klein* 'small'
(A) `completeness`: *komplett* 'entire'
(A) `age`: *neu* 'new'; *jung* 'young'

The collocates of the nouns *Team* 'team' and *Kurve* 'curve' find corresponding attributes in the inventory provided by GermaNet. However, the adjectives *jung* and *neu* are quite problematic. Both address the `age` property of the noun *Team*. However, *jung* refers to the `age` of team members, whereas *neu* to the `age` of team as a whole, i.e. the amount of time a team has existed.

Kurve 'curve'
(A) `shape`: *scharf* 'sharp', *sanft* 'gentle', *weich* 'gentle', *zackig* 'jagged', *konvex* 'convex'
(A) `dimension`: *steil* 'steep', *eng* 'narrow'
(A) `color`: *blau* 'blue'

Objekt 'object'
(Q/FE) Type: *dreidimensional* 'three-dimensional'
(A) `functioning`: *dekorativ* 'decorative'
(A) `quantity`: *zahlreich* 'numerous'
(A) `cost`: *teuer* 'expensive', *wertvoll* 'valuable'
(A) `brightness`: *obskur* 'obscure', *hell* 'bright'
(A) `emotion-stimulus`: *interessant* 'interesting'
(A) `religion`: *sakral* 'sacred'
(A) `evaluation`: *hochwertig* 'of high quality'
(A) `location`: *entfernt* 'distant'
(A) `dimension`: *kompakt* 'compact'

Geruch 'smell'
(FE) Descriptor: *beißend* 'acrid', *stechend* 'pungent', *penetrant* 'pungent'
(FE) Percept: *modrig* 'musty', *muffig* 'musty', *brenzlig* 'burnt', *ranzig* 'rotten', *süß* 'sweet'
(A) `degree`: *streng* 'strong', *intensiv* 'intense'
(A) `evaluation`: *unangenehm* 'unpleasant', *übel* 'foul', *widerlich* 'disgusting', *ekelhaft* 'disgusting'

The noun *Geruch* 'smell' belongs to the semantic field "Attribute", and is by itself an attribute. For instance, `smell` can be considered an attribute of the noun *flower* as in *fragrant flower*, together with `color`, `size` and, say, `shape`. As discussed in section 4.5.1, attributes can have their own frame structure and can be further described by other attributes.

Preis 'price'
(A) `degree`: *hoch* 'high', *horrend* 'horrendous', *stolz* 'stiff', *niedrig* 'low', *moderat* 'moderate', *tief* 'low', *rückläufig* 'descreasing'
(A) `cost`: *erschwinglich* 'affordable'
(A) `evaluation`: *toll* 'great', *attraktiv* 'attractive', *vernünftig* 'reasonable'
(A) `certainty`: *stabil* 'stable', *fest* 'fixed', *konstant* 'konstant'
(A) `time`: *heutig* 'of today'

All the collocates of the noun *Preis* 'price' have been assigned to the corresponding attributes. In two cases, an adjective's meaning is modeled in GermaNet differently: the adjective *stolz* lit. 'proud' that serves as an intensifier of *price* is not modeled in that sense in GermaNet, therefore, it has been accommodated by the attribute `degree` used in the previous examples to label intensification. Similarly, the adjective *fest* 'fixed' has been listed under the attribute `certainty` due to the absence of a suitable attribute in its own hierarchy in GermaNet.

Strafe 'punishment'
(A) `degree`: *drakonisch* 'draconian', *unmenschlich* 'inhumane', *scharf* 'harsh', *streng* 'strict', *hart* 'harsh' , *empfindlich* 'severe'*saftig* 'stiff', *deftig* 'severe', *hoch* 'high', *schwer* 'heavy', *schlimm* 'severe', *niedrig* 'low', *mild* 'mild'
(A) `reference` : *symbolisch* 'symbolic'
(A) `duration`: *lebenslang* 'lifelong'

A similar approach has been taken to the modeling of the *Strafe* 'punishment'. More than half of the collocates labeled as `degree` do have have the corresponding artificial concept as their hypernym in GermaNet. In contrast, the adjectives *drakonisch* 'draconian', *streng* 'strict', *deftig* 'severe', and *mild* 'mild' do not have a corresponding sense in GermaNet.

Thema 'issue/topic'
(A) `requirement`: *groß* 'big/major', *heiß* 'hot', *beherrschend* 'dominant', *drängend* 'pressing', *brennend* 'urgent', *wichtig* 'important', *zentral* 'central'
(A) `duration`: *ewig* 'eternal'
(A) `emotion-stimulus`: *unbequem* 'uncomfortable', *leidig* 'vexed', *sensibel* 'sensitive', *interessant* 'interesting'
(A) `difficulty`: *heikel* 'delicate', *brisant* 'controversial'
(A) `time`: *aktuell* 'current'
(A) `region`: *lokal* 'local'
(A) `functioning`: *sperrig* 'unwieldy'

All the collocates could be accommodated by the attributes from GermaNet. In a number of cases, the corresponding sense of an adjective is not modeled in the database. Similarly to the previous examples, such examples were labeled with most suitable available attributes. For instance, the adjective

heiß 'hot' in combination with the noun *Thema* 'issue/topic' denotes 'importance' and thus is semantically similar to *beherrschend* 'dominant', *brennend* 'urgent', *groß* 'big/major'. The latter three adjectives are listed under the node *bedarfsspezifisch* 'requirement/demand/need-related', and, therefore, all the collocates with that meaning have been assigned to this category.

Ähnlichkeit 'similarity'
(A) `degree`: *groß* 'strong', *stark* 'strong', *leise* 'slight', *unheimlich* 'uncanny'
(A) `emotion-stimulus`: *frappierend* 'striking', *frappant* 'striking', *erstaunlich* 'surprising'
(A) `time`: *flüchtig* 'fleeting'
(A) `location`: *äußerlich* 'external'
(A) `precision`: *exakt* 'precise', *vage* 'vague'
(A) `correspondence`: *beliebig* 'arbitrary'
(A) `certainty`: *variabel* 'variable'

Anzahl 'amount'
(A) `degree`: *stattlich* 'stately', *groß* 'large', *hoch* 'high', *schier* 'sheer', *beträchtlich* 'large', *riesig* 'huge', *erklecklich* 'considerable', *bedeutend* 'substantial', *bescheiden* 'modest', *gering* 'small', *klein* 'small', *überschaubar* 'manageable'

Ziel 'goal'
(FE) DESCRIPTOR: *hochgesteckt* 'ambitious', *hoch* 'high', *ehrgeizig* 'ambitious'
(A) `degree`: *ober* 'ultimate', *groß* 'great'
(A) `requirements`: *primär* 'primary', *zentral* 'central', *wichtig* 'important'
(A) `linking`: *selbstgesteckt* 'self-imposed', *gemeinsam* 'joint'
(A) `accuracy`: *klar* 'clear'
(A) `certainty`: *fest* 'firm'
(A) `evaluation`: *hehr* 'noble'
(A) `time`: *langfristig* 'long-term', *kurzfristig* 'short-term'

The collocates of the nouns *Ähnlichkeit* 'similarity', *Anzahl* 'amount' and *Ziel* 'goal' were annotated with corresponding attributes. There was no suitable category only for three collocates of the noun *Ziel*, therefore, the Frame Element (FE) DESCRIPTOR was used to fill in this gap.

Angst 'fear'
(A) `degree`: *groß* 'strong/big', *tief* 'deep', *höllisch* 'hellish', *furchtbar* 'tremendous'
(A) `certainty`:: *nackt* 'pure', *pur* 'pure', *blank* 'sheer'
(A) `accuracy`: *diffus* 'vague', *dumpf* 'vague'
(A) `habit`: *krankhaft* 'pathological'

The modifiers of the noun *Angst* 'fear' have been successfully modeled with attributes. There is one adjective for which GermaNet does not have the corresponding senses: *nackt* 'pure' (lit. 'naked') as only its literal meaning is reflected in the database.An adjustment has been made for this adjective: it is annotated as `certainty`

Interesse 'interest'
(FE) `Descriptor`: *lebhaft* 'lively', *rege* 'active'
(A) `degree`: *hoch* 'strong', *groß* 'strong/big', *stark* 'strong', *massiv* 'massive', *gering* 'little'
(A) `requirement`: *vital* 'vital', *elementar* 'fundamental'
(A) `quantity`: *breit* 'broad'
(A) `evaluation`: *echt* 'genuine', *wahr* 'genuine', *handfest* 'tangible'
(A) `linking`: *ureigen* 'own', *widerstreitend* 'conflicting'

The inventory of attributes provides suitable categories for all the collocates of the noun *Interesse* 'interest'. The only exception is formed by the adjectives *lebhaft* 'lively' and *rege* 'active'. In GermaNet, both adjectives are listed under the category `character` which cannot be a property of the abstract noun *interest*. Thus a much more general category DESCRIPTOR is borrowed from FrameNet to accommodate the two collocates (see the discussion about this label in section 4.4.2).

The modeling of AN combinations presented in this section was to a large extent based on attributes. Additionally, it integrated categories from the other three inventories covered in the previous three sections of this chapter. This approach and the results are evaluated and discussed in the next section.

4.6 Summary

The main objective of this thesis is to establish an inventory of relations between the constituents of adjective-noun combinations. The theory underlying the inventory should offer a comprehensive description of the semantics of nouns and adjectives. Further criteria of the inventory's adequacy are granularity and coverage. An ideal inventory of relations is characterized by flexible granularity that can be adjusted depending on the task. This thesis aims at achieving a medium level of granularity: specific enough to distinguish multiple categories of collocates for both concrete and abstract nouns, and general enough to include various collocates in a category. Coverage can be measured by the number of semantic classes accommodated by the inventory.

In this chapter, four theoretical frameworks that can potentially provide inventories for such relations have been tested: the Meaning-Text Theory (MTT) with the inventory of Lexical Functions (LFs), the Generative Lexicon with qualia roles, Frame Semantics with Frame Elements (FEs), and Concept Frames with attributes. The study is based on a sample of German adjective-noun pairs that contains nominal bases from different semantic classes and of varying degrees of concreteness. The sample covers all 23 semantic classes available in the German wordnet GermaNet with the most concrete noun being *Schokolade* 'chocolate' and the most abstract being *Interesse* 'interest'. The broad semantic coverage of the data sample made it possible to reveal the advantages and limitations of each inventory.

The theories investigated in this chapter were chosen for analysis because they all provide an account of nominal and adjectival semantics. The Meaning-Text theory is very general; it covers all parts of speech. The Generative Lexicon theory pays close attention to the semantics of nouns and their behaviour in combination with other parts of speech. Frame Semantics covers all parts of speech, but focuses mainly on verbs or nouns denoting events and processes and provides only a very limited account of concrete nouns. Finally, concept frames model how conceptual knowledge is represented, and clearly the focus is shifted towards nominals. It represents nominal meaning in terms of attribute-value matrices where values are frequently expressed by adjectives.

In terms of coverage, each of the theories analysed in this chapter exhibits a different behaviour. Standard Lexical Functions do not cover the semantics of concrete nouns and only partially describe the semantics of abstract nouns. In contrast, qualia roles are most suitable for concrete nouns whereas abstract ones pose a challenge for this approach. Frame Elements take into account nouns from different semantic classes and thus are suitable for modeling the meaning of both concrete and abstract nouns. However, the frame structure of concrete nouns is very poor. Concept frames are suitable for describing all types of concepts in terms of attributes. The specific inventory of attributes investigated in this chapter was borrowed from GermaNet. Though this inventory exhibits some isolated gaps, they do not seem to form a pattern as in the case of the other three theories.

The final criterion is the granularity of the inventories. Among the frameworks investigated in this chapter, the most general inventory of relations is comprised by the qualia roles that include four broad categories. These four coarse categories require specifications. In contrast, Lexical Functions are extremely fine-grained, both standard and non-standard ones. Frame Elements have both aspects to them: some Frame Elements such as TYPE and DESCRIPTOR are very general, but there are also a lot of very specific ones applicable only to particular lexical units that belong to particular Frames. Attributes provide considerable flexibility in terms of granularity. For instance, WordNet is more granular than GermaNet as it distinguishes 254 attributes, whereas the German wordnet contains 80 attributes that are more general than those in WordNet.

Concept frames proved to be the most suitable approach for the given task and can serve as the basis of the required inventory. This is due to several factors:

1. Concept frames focus on the semantics of nouns. The structure of concept frames is an attribute-value matrix where the values are frequently expresses by adjectives.

2. Attributes have common ground with the other three theories and can address their shortcomings to a large extent.

   As already proposed by Jousse (2007), attributes serve as a way to generalize over the non-standard Lexical Functions (discussed in 4.2.1). Moreover, it has been observed in the first case study that standard Lexical Functions in most cases require specifications as to what aspect of the noun's meaning is evaluated or intensified, for example, $\texttt{Magn}^{temp}$ or $\texttt{Magn}_{[precipitation]}$. Attributes offer a consistent inventory of such specifications. Similarly, attributes can be considered to be refinements of the qualia roles. For example, the attributes `color`, `shape`, and `size` are specifications of the FORMAL role; `function` of the TELIC role; `quantity` of the CONSTITUTIVE role. As for the relation to Frame Semantics, it has already been advocated in Gamerschlag et al. (2014) that its combination with Concept Frames can provide much richer semantic representations: both the participants of situations and attributes are necessary for representing events, concepts, and states of affairs as frames.

3. There is a lexical-semantic resource, GermaNet, that provides a comprehensive and structured list of attributes with flexible granularity.

   The inventory of the categories in GermaNet that correspond to the notion of an attribute is based on a comprehensive semantic classification of adjectives by Hundsnurscher and Splett (1982).

4. Concept frames can provide an inventory of attributes at any required level of granularity.

   Barsalou (1992) emphasizes that almost any attribute can be further described in terms of other attributes. Hence, an important factor to bear in mind is that, for any degree of granularity, the relations' generalization degree should be consistent. For instance, it is undesirable to include in the same inventory both `dimension` and `height` as labels, because `height` is a subcategory of `dimension` along with the attributes `width` and `length`. So it should either be only `dimension`, or the three more specific attributes `height`, `width`, and `length`, depending on the required level of granularity. Furthermore, polarity should be treated consistently: either one attribute per property should be introduced or, for each property that supports polarity, two attributes are required. For example, an inventory that does not distinguish polarity may include the attributes `stability` and `accuracy`. An inventory that makes the polarity distinction would also include `instability` and `inaccuracy`. An inventory that only contains `stability`, `accuracy` and `instability` but not `inaccuracy` is inconsistent.

5. There are no limitations in terms of coverage: attributes are equally effective in modeling all semantic classes.

   The few limitations of the attribute approach encountered in the case study stem only from the specific inventory of attributes chosen in the case study. As mentioned above, those gaps can be filled by the categories from other investigated theories:

   - Only among Frame Elements is a label available that describes an entity or the way an action is performed in general. More specifically, there are two FEs: Descriptor – for entities and Manner – for actions and processes (see the discussion in section 4.4.2).
   - The Frame Element Type refers to a hyponymic relation, but not between the constituents of an adjective-noun phrase as in "*white* is the `color` of the *dress*", but rather between the phrase as whole and its head as in *white chocolate* is a Type of *chocolate* Such collocations behave as compound phrases and are more similar to nominal compounds than to adjective-noun pairs where the adjective typically selects a property of its head.
   - Only Lexical Functions offer a category for phrases where the modifier is semantically almost empty – the LF `Epith`.

Based on these observations, guidelines for annotating the relations between the constituents of adjective-noun collocations and free phrases can be compiled. The annotation scheme will to a large extent rely on the attributes provided in GermaNet complemented by three other inventories when needed. For instance, in the phrase *alter Freund* 'old friend', the adjective *alt* is the

value of the attribute DURATION. The attribute stems from the artificial concept DURATION-RELATED that is a hypernym of the adjective *alt* in GermaNet.

There is, however, one last step to make before the discussion turns to the guidelines and annotation. Namely, in order to test the new annotation scheme, a large collection of adjective-noun pairs is required. As discussed in section 3.6, none of the available collections of collocations fulfill the requirements of this investigation. More specifically, the required dataset should be large, semantically varied, and contain not only collocations, but also free phrases. In joint work with colleagues at the Brandenburg Academy of Sciences (BBAW), such data collection has been compiled and manually annotated. A detailed description of the process and the evaluation of the created dataset will be presented in the next chapter.

Chapter 5

Dataset of collocations and free phrases

This thesis studies lexical semantics of adjective-noun phrases focusing on the distinction between collocations and free phrases. The previous chapter ended with a proposal of a classification scheme for relations that hold between the constituents of adjective-noun phrases, namely attributes based on the semantic classification of adjectives in GermaNet. In order to ground this theoretical proposal in data, a large dataset is required that covers collocations and non-collocations from different semantic fields. As discussed in section 3.6, none of the existing resources for collocations is suitable for the undertaken study. Therefore, a major effort is devoted to constructing a large gold standard dataset of German adjective-noun phrases that contains both collocations and non-collocations. This chapter will be devoted to describing that work.

The dataset has been created in joint work with colleagues from the Brandenburg Academy of Sciences (BBAW): Isabel Fuhrmann, Dr. Alexander Geyken, and Dr. Lothar Lemnitzer, and shortly introduced in Strakatova et al. (2020). This chapter elaborates on the process of data selection and annotation. The chapter starts with a short discussion of different approaches to creating gold standards for collocations (section 5.1). The following sections will describe dataset construction that consists of four steps:

1. Selection of adjective-noun pairs (section 5.2) performed by Isabel Fuhrmann;
2. Creating guidelines for manual annotation (section 5.3.1) – joint work;
3. Annotation by independent annotators and adjudication of disagreement cases (section 5.3.2) – joint work;
4. Evaluation of the annotation scheme in terms of inter-annotator agreement – (section 5.3.3).

A thorough analysis of the created dataset will be presented in section 5.4. The robustness of the annotation scheme will be tested in a series of machine learning experiments on binary classification of adjective-noun phrases into collocations and free phrases (section 5.5), joint work with Neele Falk reported in Strakatova et al. (2020). Finally, the chapter ends with an a short summary (section 5.6).

5.1 Starting point for data selection

In adjective-noun phrases, the noun is the base chosen freely by the speaker, and the adjectival modifier is either chosen freely (in free phrases) or is contingent on the base – in collocations (Mel'čuk, 2012a). There are three approaches to creating a sample to study the semantics of adjective-noun phrases and collocations in particular: (1) extract phrases randomly from a corpus; (2) first choose the bases relying on certain criteria, then extract a number of collocates for the selected bases; or (3) first select the collocates based on certain criteria and then extract a number of bases that co-occur with the selected modifiers. All three approaches have been used in previous work on collocations.

Random. In research concerned with collocations and their extraction, a common approach to obtaining collocation candidates is automatic extraction from a corpus. The only criteria are the frequency, association strength, and the syntactic relation between the constituents (Evert, 2008; Pecina, 2008). Lists of co-occurrences extracted in this way are then manually classified into collocations and non-collocations. This approach lacks systematicity in terms of semantic coverage and is, therefore, not applicable in this thesis.

Starting with the base. According to Melcuk (Mel'čuk, 1996, 2012a) and Hausmann (1984), phrases are generated by the speaker in the following order: **base** → **collocate**. It is also the most logical and efficient way of representing collocations in a learners' dictionary: listing the collocates under their bases (Hausmann, 1984; Mel'čuk and Zholkovsky, 1984; Benson, 1989; Rundell, 2010; Quasthoff, 2011). For instance, phrases such as *heavy rain* or *black coffee* should be listed under *rain* and *coffee* respectively. This approach proved to be helpful for language learners as illustrated by Benson (1989) in the tasks focusing on text generation.

A systematic study of collocations with nominal bases as the starting point has been undertaken by Mel'čuk and Wanner (1996). They focus on verbal collocations from the domain of emotion lexemes in German. The choice of the domain is motivated by two factors: (1) emotion words are in general often collocational/idiomatic and (2) this domain has been extensively studied in linguistics and psychology.

The authors' assumption is that although the choice of collocates is arbitrary, regularities between the semantic class of the bases and the choice of their collocates can be established. For instance, German **reactive** lexemes such as *Zorn* 'wrath', *Staunen* 'astonishment', *Ekel* 'disgust' tends to co-occur with the verb *erregen* 'to excite', whereas **attitudinal** lexemes such as *Liebe* 'love' or *Achtung* 'respect' co-occur with the verb *entegenbringen* 'to show'.[1] Moreover, Mel'čuk and Wanner (1996) show that some verbs co-occur with (almost) all the lexemes from the emotion domain including the **generic** lexeme of the domain. In the emotion domain, the generic lexeme is the noun

[1] Reactivity, attitudinality are semantic dimensions of emotion lexemes. Mel'čuk and Wanner (1996) list 11 dimensions for the emotion lexemes they investigate.

Gefühl 'feeling'. For instance, the verb *empfinden* 'to feel/experience' that combines with the generic lexeme *Gefühl*, co-occurs with 35 out of 40 investigated emotion lexemes. Such collocates can be listed in a dictionary under the generic lexeme and there is no need to list them for each of the more specific lexemes. Relying on these findings, the authors conclude that the principle of lexical inheritance works to some extent for collocations, and thus it is possible to make certain generalizations over collocation bases and the collocates they combine with.

Starting with the collocate. An attempt to group collocation bases relying on semantic criteria is presented in Reuther (1996). In contrast to the study by Mel'čuk and Wanner (1996), the starting point are not nouns, but verbs. Reuther (1996) investigates inter alia whether and how the nominal collocates of each verb can be grouped semantically. The study focuses on three Russian support verbs that express the meaning "to conduct/carry out" when combined with certain nouns. For instance, in example (27a), all three support verbs are acceptable. In the other two examples, only one verb can be selected by the noun:

(27) a. *vesti/provodit'/proizvodit' dopros* 'to conduct an interrogation'
b. *vesti/*provodit'/*proizvodit' bor'bu* 'to wage a fight'
c. **vesti/*provodit'/proizvodit' aresty* 'to carry out arrests'

The analysis shows that generalizations are possible. For example, the nouns that select the verb *vesti* can be classified into five semantic groups: (1) activities of a group of people controlled by one person (*seminar, tutorial*); (2) multilateral communicative activities (*conversation, debate*); (3) texts (*diary, notes*); (4) military activities (*attack, battle*); (5) general activities (*activity, work*) (Reuther, 1996, p.203).

Approach in this thesis. The above described studies show that both approaches of starting with a base or with a collocate give insights into semantics of collocations. This thesis aims at investigating semantics of adjective-noun phrases and outlining differences between collocations and free phrases. In adjective-noun collocations, it is the adjective that changes its meaning depending on the noun it modifies. Therefore, the second approach seems more suitable for the undertaken investigation. Namely, the starting point should be the adjective which is polysemous and undergoes meaning shifts under the influence of its nominal head.

5.2 Data

The required data should include attributive adjectives from different semantic fields and nouns with which the selected adjectives typically co-occur. For German, rich linguistic resources have been developed in recent years that are suitable for collecting this kind of data. More specifically, the lexical-semantic net GermaNet (Hamp and Feldweg, 1997; Henrich and Hinrichs, 2010) provides required information about semantic fields for both

adjectives and nouns. In order to extract co-occurrences, the Digital Dictionary of the German language – the DWDS,[2] will be used. The DWDS has an integrated tool for creating word sketches, the DWDS-Wortprofil, introduced in previous chapter (see section 4.1). Both resources undergo regular updates, and with each new release, they cover more data based on high quality corpora.

5.2.1 Selecting adjectives

As described in Strakatova et al. (2020), the adjectives in the dataset are selected based on the lexical-semantic information provided by the German wordnet GermaNet. GermaNet (version 15.0)[3] contains 15,275 adjectives. Most of the adjectives (≈ 90%) have only one sense (see table 5.1).

Senses	Lexical units
10	2
9	1
8	1
7	1
6	6
5	17
4	57
3	247
2	1,189
1	13,767
∑	15,275

Table 5.1: Distribution of number of senses for the adjectives in GermaNet 15.0

Adjectives in GermaNet are modelled in a taxonomic manner in terms of hyponymy relations. The top nodes in the hierarchy are represented by 16 semantic classes following the classification of adjectives proposed in Hundsnurscher and Splett (1982) with an additional special class of *pertainyms* – adjectives that *pertain* to a noun. For example, the adjective *barock* 'baroque' pertains to the noun *Barock* 'baroque'. Another term for such adjectives common in linguistic literature is *relational*. For a more detailed description of adjectives in GermaNet see section 4.5.1.

From each of the 16 semantic classes in GermaNet, three adjectives have been selected in order to ensure that the dataset has a balanced semantic coverage. The majority of the selected adjectives are ambiguous. However, a few adjectives that are monosemous in GermaNet have been selected as

[2]DWDS – Digitales Wörterbuch der deutschen Sprache. Das Wortauskunftssystem zur deutschen Sprache in Geschichte und Gegenwart, hrsg. v. d. Berlin-Brandenburgischen Akademie der Wissenschaften, <https://www.dwds.de/>, *last accessed on 13.04.2023.*

[3]Note that GermaNet is updated on a yearly basis. The most recent version at the moment of writing this thesis is version 18.0 released in May 2023.

well. This is motivated by the fact that some senses may still be absent in GermaNet,[4] but still can be encountered in corpus data. All the adjectives with more than six senses in GermaNet have been included in the dataset. The rest of the adjectives were chosen from a set of adjectives with highest frequency from each semantic field.

The final selection of the adjectives is presented in table 5.2 including the number of senses in GermaNet and the semantic class the adjectives belong to. Note that most of the adjectives are polysemous and, therefore, they may belong to several different semantic classes. For instance, the adjective *bitter* 'bitter' is found in three semantic classes: ALLGEMEIN 'general', GEFÜHL 'mood-related', and PERZEPTION 'perceptional'.

The special class of adjectives, pertainyms, has also been considered. This class contains denominal adjectives that do not fit in any of the adjectival semantic classes because such adjectives retain the meaning of the nouns they *pertain to*. For instance, the adjective *barock* 'baroque' pertains to the noun *Barock* 'baroque' that belongs to the nominal semantic class EVENT. Therefore, the semantic class of the adjective *barock* is DENOMINAL EVENT. Relational adjectives are semantically similar to nouns and denote objects (Boleda, 2006). For instance, *barocke Musik* 'baroque music' means music *of* baroque period. However, pertainyms can develop other senses and be used as qualitative adjectives. For example, the adjective *barock* 'baroque' may express the meaning "strange, bizarre, weird" as in *barocke Fantasie* 'weird imagination'. Any adjective from the semantic class of pertainyms has a potential for developing novel, non-relational senses. Thus, three pertainyms are included in the dataset along with adjectives from other semantic classes.

Literal sense of adjectives

There are two assumptions made about collocations of type 1 (definition 6, repeated below for convenience as definition 8) in this thesis. Firstly, the modifier in collocations of this type does not retain its literal sense. The second assumption is that, for each adjective, there is only one literal sense (definition 3, repeated below for convenience as definition 9). Therefore, in the next step the literal sense for each of the selected 48 adjectives must be clearly defined. This has been done by an experienced lexicographer.

Definition 8
Adjective-noun collocations are recurrent binary word combinations consisting of a base (noun) and a collocate (adjective) where the base is used in its unrestricted sense and the collocate's meaning is restricted. There are two types of restricted collocates:

Type 1: the collocate's sense is not literal. Type 1 collocations have two subtypes:

Type 1.1: the collocate combines in this sense with a large number of nominal bases, for instance, *helle Aufregung* 'extreme excitement' (lit. 'bright').

[4]GermaNet is an ongoing project and new entries are added to the database every year.

Semantic class	Adjectives
ALLGEMEIN 'general'	*herrlich* 'wonderful' (1), *knapp* 'scarce' (4), *stark* 'strong' (5)
BEWEGUNG 'motion-related'	*starr* 'rigid' (2), *sanft* 'smooth' (3), *wild* 'wild' (6)
GEFÜHL 'mood-related'	*bitter* 'bitter' (4), *süß* 'sweet'(4), *zart* 'delicate' (5)
GEIST 'spirit-related'	*schlau* 'smart' (1), *dumm* 'stupid' (3), *hell* 'bright' (4)
GESELLSCHAFT 'social-related'	*deftig* 'solid' (2), *arm* 'poor' (3), *blank* 'broke' (4)
KÖRPER 'body-related'	*zäh* 'tough' (2), *blind* 'blind' (4), *dick* 'fat' (6)
MENGE 'quantity-related'	*prall* 'full' (2), *teuer* 'expensive' (2), *reich* 'rich' (3)
NATPHÄNOMEN 'weather-related'	*karg* 'bare' (2) , *mild* 'mild' (3), *stürmisch* 'stormy' (3)
ORT 'spatial'	*steil* 'steep' (2), *rund* 'round' (4), *tief* 'deep' (6)
PERTONYM 'pertainyms'	*steinig* 'stony' (1), *barock* 'baroque' (2), *historisch* 'historical' (2)
PERZEPTION 'perceptional'	*dunkel* 'dark' (5), *schwarz* 'black' (8), *scharf* 'sharp' (10)
PRIVATIV 'privative'	*windig* 'dubious' (2), *tot* 'dead' (3), *frei* 'free' (9)
RELATION 'relation'	*mächtig* 'mighty' (4), *sicher* 'safe' (4), *leicht* 'light' (7)
SUBSTANZ 'material-related'	*hölzern* 'wooden' (2), *grob* 'coarse' (4), *offen* 'open' (5)
VERHALTEN 'behaviour-related'	*frech* 'bold' (2), *rau* 'rough' (4), *hart* 'tough' (10)
ZEIT 'temporality-related'	*frischgebacken* 'recent' (1), *spät* 'late' (2), *alt* 'old' (5)

TABLE 5.2: 48 adjectives selected for the dataset and their semantic classes in GermaNet. The number of senses for each adjective is given in brackets.

Type 1.2: the collocate combines in this sense with very few nominal bases, for instance, *schwarzer Kaffee* 'black coffee'.

Definition 9
Literal meaning is the most basic meaning of a word (German: *Grundbedeutung*). Following the definition by Pragglejaz Group (2007, p.3): the most concrete, imaginable, historically older meaning, it is precise and is related to a bodily action.

For some adjectives, is it straightforward which sense is the literal one. For instance, perceptional adjectives are among the least problematic: for *schwarz* 'black', the color sense is the literal one, for *bitter* 'bitter' and *süß* 'sweet' – the taste sense. Spatial adjectives also have a distinct literal meaning:

- *steil* 'steep' – rising up vertically or sloping downwards;[5]
- *rund* 'round' – spherical, (semi)circular; without corners and edges, in the shape of a ring, arc or similar;
- *tief* 'deep' – reaching (far) down from above, extending far downwards.

In most cases, the literal sense is the sense listed first in the DWDS dictionary. However, not all cases are straightforward and clear. Additional heuristics have been adopted to help identify the most literal senses of adjectives.

Heuristic 1. Non-literal senses are marked in the DWDS entries as *übertragen* 'figurative' or *bildlich* 'metaphorical' which helps to identify the literal meaning in the cases, when it is not listed as the first sense. For instance, the entry of the adjective *frischgebacken* 'freshly baked' contains two senses. The first one is marked as figurative. The figurative sense is activated when the adjective is used with role nouns. For instance, nouns expressing profession such as *frischgebackener Präsident* 'newly elected president' or marital status such as *frischgebackene Ehepaar* 'newly married couple'. The second sense is unmarked and is the literal one: 'recently/freshly baked' – as in *frischgebackenes Brot/Kuchen* 'freshly baked bread/cake'.

Heuristic 2. There are adjectives for which none of the senses is marked as figurative, for example *sicher* 'safe/certain', *frech* 'cheeky/bold', *zart* 'delicate'. The reason is that all the senses of these adjectives are quite abstract. In such cases, the literal meaning is equated with the *salient meaning*. Philip (2011) defines *salient meaning* of a word or phrase as the most dominant one for a particular person, in this case, the lexicographer. For example, the adjective *frech* 'bold, cheeky' has one general entry and several subentries. The general entry defines the adjective as "lacking the necessary respect towards someone, (a little) disrespectful".[6] The definitions in the three subentries are

[5]Here and further: German definitions provided in the DWDS are translated into English.

[6]Here and further translated from German.

the following: (1) act unseemly toward somebody; (2) cheeky, rude, shameless; (3) bold, funny, provoking. The first one refers to the adverbial usage of *frech* with the verb *kommen* 'to come' and can be ignored.

The second and the third definitions differ in their connotations: the second one has a negative connotation, whereas the third one is positive. The selection of examples in the DWDS entries for each usage illustrates this difference: in the second subentry, *frech* modifies nouns such as *Lüge* 'lie', *Betrug* 'deception', whereas in the third one, it combines with artifacts such as *Karikatur* 'caricature', *Zeichnung* 'drawing'. It has been decided to treat the negative connotation as literal sense of *frech*. The positive use of the word is considered slightly figurative and thus collocational.

The definitions of literal senses for all the 48 adjectives are listed in the Appendix B. The second step in collecting collocation candidates is extracting the nominal bases that have a potential to form a collocation with the selected adjectives.

5.2.2 Selecting nouns

The choice of the adjectives described in the previous subsection was motivated purely semantically, the main factors being the semantic class and the level of polysemy of the adjectives. The selection of the nouns, on the other hand, is fully automatic and is based on the strength of statistical association between the adjective and the noun. In order to extract the candidates, the statistical tool of the DWDS, the Wortprofil, has been used (The Wortprofil was described in detail in section 4.1).

For extracting the corresponding nominal heads for the selected adjectives, only the relation "is an attribute of" in the Wortprofil has been considered. Figure 5.1 shows a part of the interface of the Wortprofil website: the top-20 noun bases for the adjective *hell* 'bright, light' sorted by the logDice values.[7] Already the top twenty collocates give an idea about the different senses of the polysemous adjective. With artifacts, *hell* 'bright' is used in its literal sense 'radiating light, not dark': *heller Raum/Anzug/Hose* 'bright or light room/suit/pants', the same holds for natural phenomena and objects such as *Licht* 'light', *Flamme* 'flame', *Stern* 'star', *Holz* 'wood'. These co-occurrences are examples of free phrases, in spite of the fact that the statistical association between these nouns and the adjective is very strong.

In a large number of phrases in this list, the adjective *hell* does not retain its literal meaning. For instance, in combination with the nouns of emotion such as *Aufregung* 'excitement', *Freude* 'joy', *Empörung* 'outrage', etc., the adjective *hell* takes up the role of an intensifier. In combination with *Stimme* 'voice', *hell* describes an acoustic property of the noun. These are examples of collocations.

[7] DWDS-Wortprofil for "hell", created via the das Digitale Wörterbuch der deutschen Sprache, <https://www.dwds.de/wp/hell>, *last accessed on 23.02.2021.*

ist Adjektivattribut von	logDice	Freq.
1. Aufregung	9.9	1266
2. Licht	9.4	1807
3. Freude	9.0	858
4. Holz	8.8	570
5. Stern	8.6	709
6. Farbe	8.4	805
7. Empörung	8.4	419
8. Wahnsinn	8.0	309
9. Schar	8.0	318
10. Flamme	7.8	284
11. Mitte	7.8	357
12. Kopf	7.7	496
13. Tageslicht	7.5	203
14. Anzug	7.5	301
15. Haut	7.5	266
16. Begeisterung	7.4	228
17. Stimme	7.3	640
18. Auge	7.3	598
19. Raum	7.3	758
20. Hose	7.2	233

FIGURE 5.1: A screenshot from the Wortprofil: top-20 noun co-occurrences for the adjective *hell* 'bright, light' sorted by the logDice values.

Apart from that, there is an ambiguous phrase *heller Kopf* lit. 'bright head' which can be interpreted in two ways. Firstly, it has an idiomatic interpretation in which the noun is used figuratively as in 'a smart person'.[8] This is a clear case of metonymy because *head* stands for *person*. The second interpretation of the adjective is literal: 'a light-colored head', where *head* stands for *hair*. One of the advantages of the Wortprofil is a quick access to the examples from the corpus for each phrase. When the user clicks on a word in the list of co-occurrences, a maximum of twenty example sentences from the underlying corpus are displayed. This is a quick way to establish empirically which interpretation of a phrase is dominant. All twenty example sentences for the phrase *heller Kopf* illustrate the idiomatic interpretation of the phrase. This is an indication that the figurative use of the phrase is

[8]The English equivalent is 'bright mind'.

more dominant than the literal one. See examples (28a)[9] and (28b)[10] that illustrate that meaning.

(28) a. Beim Schach stoßen selbst die hellsten Köpfe rasch an
in chess bump self the brightest heads quickly against
ihre Grenzen.
their limits
'In chess, even the brightest minds reach their limits soon.'

b. Dabei gehe es darum, "helle und kreative Köpfe" nach
thereby go it about bright and creative heads to
Deutschland zu holen.
Germany to bring
'It was about bringing "bright and creative minds" to Germany.'

The above described functionality of the Wortprofil makes it a suitable tool for obtaining collocation candidates. Therefore, for each of the 48 adjectives described earlier in section 5.2.1, up to 100 nominal collocates have been extracted with the help of the Wortprofil. The minimum bigram frequency has been set to five occurrences. For three adjectives, less than one hundred collocates are available. Those are *steinig* 'stony' with 57 nouns, *frischgebacken* 'freshly baked' – 76 nouns, and *steil* 'steep' – 99 nouns.

The final dataset consists of 4,732 adjective-noun pairs with 48 distinct adjectives (collocates) and 3,007 distinct nouns (bases). The next step is manual annotation of these phrases in order to obtain a gold standard of free phrases and collocations.

5.3 Annotation

The obtained lists of collocation candidates (4,732 adjective-noun phrases) were given to two native speakers of German for annotation. One annotator is an experienced lexicographer and an expert on collocations (Isabel Fuhrmann), the second one is a student of Computational Linguistics with a solid background in lexical semantics and lexicography (Daniela Rossmann). The annotation task at hand is a binary classification of adjective-noun phrases into *collocations* and *non-collocations*. This section describes in detail the annotation process which is a joint work of Isabel Fuhrmann (BBAW, Berlin) and Yana Strakatova (University of Tübingen) supported by a student assistant Daniela Rossmann (University of Tübingen). The work has been briefly introduced in Strakatova et al. (2020).

[9] In *Die Zeit, 07.02.2011, Nr. 06*, from the corpus Die ZEIT of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/zeit>, *retrieved on 24.02.2021.*

[10] In *Berliner Zeitung, 18.02.2002*, from the corpus Berliner Zeitung of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/bz>, *retrieved on 24.02.2021.*

5.3.1 Annotation scheme

Figure 5.2 presents an initial decision tree that served as initial guidelines for the annotation by Isabel Fuhrmann. The annotation is binary as there are only two labels available: *collocation* and *non-collocation*. The first decision is whether a given phrase is a proper name. The decision depends on the world knowledge of the annotators and their linguistic intuition. If the phrase is not a proper name, the annotator moves on to the second step and has to decide whether the noun is used figuratively. If yes, then the expression is not a collocations, even if the adjective is used in its literal sense. The final choice depends on the adjective. If the meaning of the adjective is literal, the combination should be annotated as a non-collocation. If the meaning of the adjective is non-literal, then the phrase is labeled as collocation. It is left open to the annotator to write down any comments to explain their decision.

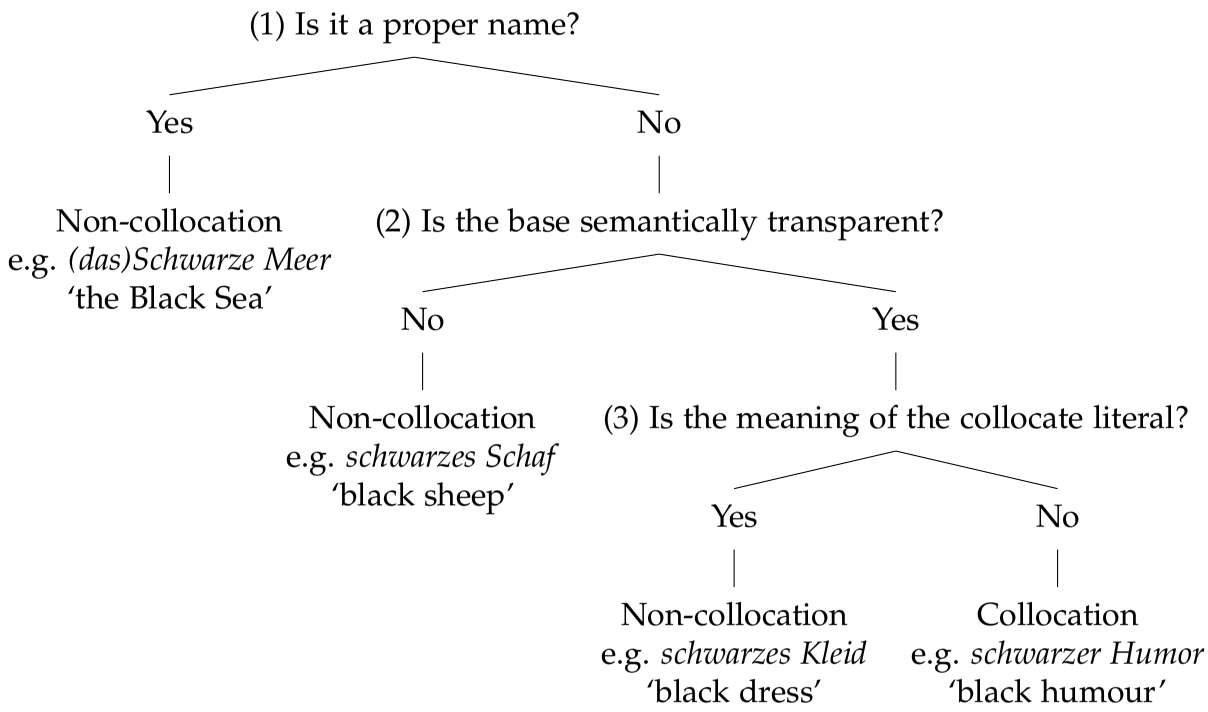


FIGURE 5.2: Annotation scheme for the binary classification of adjective-noun co-occurrences into collocations and non-collocations.

Such binary classification has one major disadvantage: phrases with very different semantic and, in part, syntactic properties are conflated into one category of non-collocations. In decision 1, a non-collocation stands for a *proper name*. In decision 2, it stands for an *idiom* because the nominal base is used in a figurative sense. Finally, decision 3 distinguishes between *collocations* and *free phrases*. Consider the definitions of these three categories that are based on the numerous definitions presented earlier in section 3.2 adapted for adjective-noun phrases:

- **Free phrase** is an adjective-noun phrase in which both constituents are lexically free and used in their literal sense.
- **Idiom** is an adjective-noun phrase in which either only the nominal base has a figurative meaning (semi-idiom) or the meaning of both constituents is figurative (full idiom).[11]
- **Proper name (named entity)** is an adjective-noun phrase used as a name or a part of the name of an organization, location, holiday or as a title of a song, a book, a movie, or similar.

Compare *free phrases* and *idioms*: free phrases are semantically fully transparent, compositional, and decomposable. Idioms, on the other hand, are non-compositional and often semantically opaque. Some idioms are syntactically (almost) fixed and non-decomposable, for instance, *mit harten Bandagen (kämpfen)* 'to play hardball' (lit. 'to fight with hard bandages'). Ideally, a distinction between free phrases and idioms should be made. Therefore, after the annotation was performed by the first annotator, it has been decided to include a more fine-grained distinction for non-collocations in the annotation scheme. The updated decision tree is depicted in Figure 7.1: instead of the general category *non-collocation*, it contains fine-grained categories *proper name*, *idiom*, and *free phrase*. This tree was used by the second annotator Daniela Rossmann who started working later than the first one.

Apart from that, one cannot rule out that any of the multi-word expression types discussed in section 3.2 may occur in the list of co-occurrences, as all MWEs are characterized by strong statistical association between their constituents. Therefore, the second annotator was encouraged to suggest other labels if none of the categories in the annotation scheme is suitable for the phrase in question.

[11] The two subtypes of idioms are borrowed from Mel'čuk's 2012a classification of phrasemes presented in section 3.2.2.

FIGURE 5.3: Updated annotation scheme that includes a more fine-grained classification of non-collocations.

5.3.2 Annotation process

The annotation was performed by the annotators independently, i.e. they did not communicate with each other until they both completed the task. The annotators received the lists of phrases sorted by adjectives. Each adjective was provided with the definition of its literal meaning (see Appendix B). The annotators were instructed to rely on this definition when making their decision about the literality of the adjective's meaning in step 3 of the decision tree. The decision about the literality of the noun's meaning was made by the annotators based on their intuition as native speakers and their expertise in lexical semantics. The phrases were given for annotation without context.

Ambiguities. As it is possible that some phrases have multiple interpretations, the annotators were encouraged to use any resources they require to disambiguate the phrases. They were free to decide what interpretation is the dominant one for an ambiguous phrase based on the corpus examples and their intuition. The annotators were also instructed to give comments and explanations about their choice in such cases.

Adjudication. During the annotation, the annotators worked independently and did not communicate with each other. After the annotation was completed, all the disagreement cases were identified, compiled with the notes of the annotators and given back to them for further adjudication. Altogether, there were 478 cases of disagreement. The annotators went through all of them in turns, changing the label if they agreed with the opinion of the other annotator. The disagreement cases were discussed taking into consideration the fine-grained categories. In this way, most of the disagreement cases were resolved with the exception of 91 cases. These remaining cases were discussed among the annotators and the author of this thesis and each phrase was assigned a label. The next subsection presents the results of the annotation in terms of the inter-annotator agreement scores and discusses some of the disagreement cases in detail.

5.3.3 Inter-annotator agreement

Based on the results of the independent annotation, the reliability of the annotators' judgements was verified by computing the inter-annotator agreement (IAA) score. Note that only one annotator performed a multiclass classification. In order to compare the annotation results and be able to calculate the agreement, the fine-grained labels *proper name, idiom, free phrase, term* had to be conflated into one category of *non-collocations*.

The observed IAA for the entire dataset is 90%. However, simple observed agreement is rarely used as the only measure of IAA because it does not take into consideration the probability of random agreement between annotators. The standard measure for evaluating the IAA in a binary classification task with two annotators is Cohen's Kappa κ (Cohen, 1960). Kappa

κ takes into account the probability of chance agreement and is computed as follows:

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \tag{5.1}$$

where p_0 is the observed agreement ratio and p_e is the expected probability of agreement if the labels are assigned randomly.

$$p_0 = \frac{A + D}{N} \tag{5.2}$$

where $A + D$ is the number of agreement cases and N is the total number of annotated instances as illustrated in Figure 5.4. The chance agreement is calculated as presented in formula 5.3.

$$p_e = \left(\frac{A_1}{N} * \frac{B_1}{N}\right) + \left(\frac{A_2}{N} * \frac{B_2}{N}\right) \tag{5.3}$$

		Rater 1	
Rater 2	Category 1	Category 2	Total
Category 1	A	B	B1(A+B)
Category 2	C	D	B2(C+D)
	A1 (A+C)	A2 (B+D)	N

FIGURE 5.4: Distribution of subjects by rater and response category from Wongpakaran et al. (2013, p.3).

If observed agreement p_0 is much larger than chance p_e, the value of Cohen's Kappa κ is high which would indicate that the annotation scheme is reliable. For the binary annotation described above the resulting κ equals **0.80**. A way to interpret κ-values is proposed in Landis and Koch (1977) according to which κ of **0.80** indicates *substantial strength of agreement*. Boleda (2006, p. 80) investigates different approaches to interpreting the agreement values and summarizes them with the following conclusion: *"there is little doubt that an agreement value exceeding 0.8 can be considered to be valid for academic purposes, but there is wide disagreement as to the meaning of values below that"*. High agreement is an indication that the annotation scheme is reliable. Now consider agreement calculation for each adjective.

IAA per adjective. The IAA-scores for each of the 48 adjectives are reported in Table 5.3. The table shows the number of collocations identified

for each adjectives, the number of synsets an adjective has in GermaNet, observed IAA given in %, and two IAA measures: Cohen's Kappa and Gwet's AC1.[12] Cohen's Kappa defined above is widely used in research, why use a second measure? The reason lies in the so-called "*Kappa paradox*" (Feinstein and Cicchetti, 1990): if the data is imbalanced, the value of Kappa may be considerably lower than high observed agreement.

Compare Kappa (κ) and observed agreement (p_0) values for the adjectives *teuer* (p_0=99% vs κ=0.0), *scharf* (p_0=92% vs κ=0.52), *barock* (p_0=85% vs κ=0.0). The difference between the two values is extreme and not intuitive. This effect can be explained by the influence of trait prevalence (Feinstein and Cicchetti, 1990). Consider the number of detected collocations for each of the three adjectives: 1 for *teuer*, 0 for *barock*, 94 for *sharf*. For the first two adjectives, non-collocations prevail. In contrast, *scharf* participates mostly in collocations. For these adjectives, the distribution of labels is highly imbalanced which results in unexpectedly low Kappa values. Investigating this issue, Gwet (2008) comes to a conclusion that Kappa works reasonably well for data with trait prevalence approaching 0.5. A skewed distribution will lead to a very low Kappa in spite of high observed agreement. Gwet (2008) proposes an alternative measure *agreement coefficient AC1* which is more stable than Kappa. The difference between the two measures lies in the way the chance agreement is computed: compare formulas 5.3 and 5.4. Kappa evaluates the probability of chance-agreement assuming that *all* observed ratings may yield an agreement by chance whereas Gwet's AC1 measure reduces "*the overall agreement by chance to the right magnitude*" (Gwet, 2008, p.47).

$$p_e = 2 * \left(\frac{A_1 + B_1}{2N}\right) * \left(1 - \frac{A_1 + B_1}{2N}\right) \tag{5.4}$$

Statistical research in various fields supports the hypothesis that AC1 is more stable than Cohen's Kappa (Wongpakaran et al., 2013; Feng, 2013; Konstantinidis et al., 2022) when the data is imbalanced. Therefore, further discussion of the IAA per adjective will be based on the AC1 scores reported in Table 5.3.

Most of the adjectives (35 out of the total of 48) yield AC1 scores over 0.8 which indicates that the distinction between the literal and non-literal senses of these adjectives is very clear. For three adjectives, the AC1 value equals 1.00: *deftig* 'savoury', *frischgebacken* 'freshly baked', and *hölzern* 'wooden'. The literal sense of *deftig* 'savoury' is easy to distinguish from the non-literal ones: it only applies when combined with food nouns as in *deftige Suppe* 'hearty/savoury soup'. The non-literal senses are found in combination with more abstract nouns. For instance, *deftige Strafe* 'severe punishment' or *deftiger Humor* 'crude humour'. Similarly to *deftig*, the adjective *frischgebacken* combines with food nouns such as *Kuchen* 'cake' or *Brot* 'bread' in its literal sense, and with nouns denoting persons in the non-literal one: *Doktor*

[12]The scores were calculated using PyCM library (Haghighi et al., 2018).

Adjective	N coll	Synsets	IAA%	Kappa	AC1
deftig 'savoury'	59	2	100%	1.00	1.00
frischgebacken 'freshly baked'	73	1	100%	1.00	1.00
hölzern 'wooden'	1	2	100%	1.00	1.00
teuer 'expensive'	1	2	99%	0.00	0.99
stürmisch 'stormy'	67	3	99%	0.98	0.98
blank 'shiny'	74	4	98%	0.95	0.97
reich 'rich'	32	3	98%	0.95	0.96
windig 'windy'	66	2	98%	0.96	0.96
steinig 'stony'	3	1	96%	0.65	0.96
rau 'rough'	79	4	97%	0.91	0.95
grob 'coarse'	74	4	97%	0.92	0.95
schlau 'smart'	16	1	96%	0.86	0.94
blind 'blind'	57	4	97%	0.94	0.94
knapp 'scarce'	79	4	96%	0.88	0.94
hell 'bright'	21	4	96%	0.88	0.94
leicht 'light'	90	7	96%	0.73	0.94
tief 'deep'	71	6	96%	0.90	0.93
spät 'late'	58	2	95%	0.83	0.93
starr 'stiff'	92	2	94%	0.67	0.93
hart 'hard'	86	10	95%	0.72	0.92
sanft 'mellow'	87	3	93%	0.63	0.91
zäh 'viscous'	90	2	93%	0.66	0.91
tot 'dead'	12	3	93%	0.72	0.91
scharf 'sharp'	94	10	92%	0.52	0.90
süß 'sweet'	54	4	96%	0.90	0.90
stark 'strong'	90	5	93%	0.77	0.90
dumm 'stupid'	22	3	93%	0.79	0.89
bitter 'bitter'	81	4	92%	0.73	0.89
offen 'open'	80	5	93%	0.75	0.88
steil 'steep'	17	2	92%	0.76	0.88
karg 'sparse'	72	2	93%	0.81	0.86
arm 'poor'	28	3	91%	0.78	0.85
barock 'baroque'	0	2	85%	0.00	0.83
mächtig 'powerful'	29	4	90%	0.75	0.80
dunkel 'dark'	29	5	89%	0.70	0.80
rund 'round'	29	4	86%	0.57	0.77
prall 'firm'	51	2	89%	0.76	0.76
schwarz 'black'	30	8	86%	0.68	0.75
zart 'soft'	79	5	85%	0.64	0.75
wild 'wild'	61	6	87%	0.74	0.74
alt 'old'	37	5	83%	0.64	0.68
frech 'bold'	29	2	79%	0.38	0.67
herrlich 'wonderful'	28	1	72%	0.00	0.63
mild 'mild'	84	3	74%	0.36	0.58
dick 'thick'	40	6	70%	0.33	0.46
frei 'free'	57	9	71%	0.42	0.40
historisch 'historical'	38	2	66%	0.25	0.38
sicher 'safe'	58	4	63%	0.28	0.26

TABLE 5.3: IAA per adjective in GerCo: Number of collocations per adjective (N coll), number of synsets listed in GermaNet (Synsets), observed agreement (IAA%), Cohen's Kappa (Kappa), Gwet's AC1 (AC1). Sorted by AC1 values. The translations are given for the *literal* sense of each adjective.

'doctor', *Eltern* 'parents'. For the adjective *hölzern* 'wooden', only one case of non-literal sense has been identified: *hölzernes Dialog* 'awkward dialogue' (lit. 'wooden dialogue').

The three adjectives with the highest agreement have one thing in common: low polysemy. One might assume that highly polysemous adjectives would be more challenging in this task. However, further inspection of the IAA scores does not support this hypothesis. Adjectives with the lowest IAA scores that are listed at the bottom of the table show various levels of polysemy: from *herrlich* 'wonderful' and *historisch* 'historical' that according to GermaNet has only one and two senses, accordingly, to *dick* 'thick' that has 6 senses and *frei* 'free' that has 9 senses in GermaNet. Which other factors can influence the decision of the annotators? As discussed in Sommerauer et al. (2020), disagreement among the annotators is an indication of ambiguities and vagueness in language, especially in semantic annotation tasks. There are several explanations for the observed disagreement cases that will be discussed further.

Ambiguity of the phrase. As no context was provided for the phrases, not all the ambiguities were resolved by the annotators in the same way. For instance, the adjective *alt* 'old' becomes ambiguous when used attributively as in *alter Trainer* 'old coach': it can either mean "of old age" or "previous". In the first case, it is a free phrase, in the second one, a collocation. Such disagreements are resolved in favor of collocations.

Interpretation of the adjective's meaning. In most cases, it is the main reason for disagreement. Consider the adjective *sicher* 'safe/secure' that has the lowest AC1 score of 0.26. Its literal meaning is "not threatened by danger". This sense was interpreted more loosely by one of the annotators: in phrases such as *sichere Rente* 'secure pension', *sicherer Job* 'secure job' and *sichere Position* 'secure position', the adjective was interpreted in its literal sense. The notion of *danger* was extended from physical danger to a more abstract interpretation such as 'danger of being fired/having no money". During the adjudication, this interpretation was adopted by both annotators and the phrases were annotated with the label "free phrase".

Another interesting example with an unexpectedly low AC1 score of 0.63 is the adjective *herrlich* 'wonderful'. According to both GermaNet and the DWDS, it has only one sense: 'magnificent, outstanding, excellent'. However, one of the annotators discovered a slight shift in meaning from purely evaluative and vague 'wonderful' to a more specific sense of "well executed/superb" often used by football commentators as in *herrlicher Kopfball* 'superb header' or *herrlicher Doppelpass* 'superb double pass'. The difference between this meaning and the one given as literal is subtle, but there is a clear pattern detected by a lexicographer and native speaker: when used with sports terms, *herrlich* implies not only evaluation as in *herrliches Wetter* 'wonderful weather', but a notion of competence of the person or their action described as 'superb'. Therefore, all such cases were annotated as collocations.

Interpretation of the noun's meaning. Not all the disagreement cases arise based on the meaning of the adjectives. Step 2 in the decision tree for the annotation requires a judgement about the meaning of the nominal base of a phrase (see figure 5.2). If the meaning is not transparent, i.e. is non-literal, the phrase should be annotated as a non-collocation, more specifically, as an idiom. For instance, the noun *Schule* 'school' is polysemous (5 senses in GermaNet). In combination with the adjective *alt* 'old', it acquires a very specific sense which in English is expressed by the adjective 'old-school' meaning 'traditional' or 'earlier'. In this case, *alte Schule* should be annotated as an idiom as in example (29a).[13]

(29) a. "Die Spezialeffekte sind total alte Schule", sagt Smith.
the special-effects are totally old school says smith
'The special effects are totally old school', - says Smith.'

b. Doch es gibt diese alte Schule nicht mehr.
yet it exists this old school not more
'But the old school is not there anymore.'

c. In meiner alten Schule kannte ich so gut wie jeden.
in my old school knew I so good as everyone
'I knew almost everyone in my old school.'

However, in some contexts, *alte Schule* can either be a free phrase or a collocation, as in examples (29b)[14] and (29c)[15] respectively. In example (29b), *old school* is a free phrase as both constituents retain their literal meanings: a building/institution that has existed for many years, whereas in example (29c), it is a collocation: *old* ≈ *former, previous*. In sum, depending on the context, there are three possible annotations for the phrase: free, collocation, or idiom. The annotators have reached an agreement by annotating this phrase as an idiom, i.e. non-collocation.

The polysemous noun *Herz* 'heart' is another example of such ambiguity: it has 6 senses in GermaNet. In combination with *wild* 'wild', it acquires a metaphorical meaning denoting a person's character or nature as in example (30).[16]

[13] In *Die Zeit, 01.12.2013, Nr. 49*, from the corpus Die Zeit of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/zeit>, retrieved on 23.03.2021.

[14] In *Die Zeit, 19.10.2017, Nr. 04*, from the corpus Die Zeit of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/zeit>, retrieved on 23.03.2021.

[15] In *Die Zeit, 23.07.2012, Nr. 30*, from the corpus Die Zeit of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/zeit>, retrieved on 23.03.2021.

[16] In *Berliner Zeitung, 15.02.1997*, from the corpus Berliner Zeitung of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/bz>, retrieved on 08.04.2023.

(30) “Das Wetter am Valentinstag war eher etwas für wilde
the weather on Valentine’s-Day was rather somewhat for wild
Herzen als für romantische Seelen.
hearts than for romantic souls
‘The weather on Valentine’s Day was more for wild hearts than romantic souls.’

In contrast to the phrase *alte Schule*, *wildes Herz* has only one reading.[17] The issue of how to annotate it arises not from the variety of interpretations that depend on the context, but rather on the decision of the annotators on whether to treat this sense of the noun as literal and label it as collocation or non-literal and label it as an idiom. The final decision was to annotate the phrase as an idiom.

Further annotation of non-collocations

A more fine-grained annotation of non-collocations depicted in Figure 7.1 was performed systematically only by one of the annotators. The other annotator made comments to justify their decision, but those comments were not consistent enough to be compared with the labels given by the second annotator. Therefore, no IAA scores could be computed for the fine-grained distinction among the non-collocations. All the labels in the final dataset were assigned based on the comments of the annotators and the discussion during the adjudication. The labels are the following: *free phrase, idiom, proper name, technical term*. Note a new label that was not present in the annotation decision tree, *technical term*. This label was suggested by the second annotator in the process of annotation. During the adjudication, it has been decided to include the new label in the annotation. All the categories will be discussed in detail in the next section (5.4).

Summary. Collocations are traditionally placed in the grey area between free phrases and idioms in terms of lexical restrictedness. The task of identifying collocations is highly challenging. However, the results of the inter-annotator study show that clear definitions and straightforward guidelines make it possible to reach substantial agreement in the task of binary classification of phrases into *collocation* and *non-collocation*. The disagreement cases are systematic and exhibit patterns that are based on ambiguities of the adjectives’ and nouns’ meanings.

[17] The phrase is also used as a proper name or as part of a proper name in the titles of movies, songs, and books. However, its meaning remains the same: "unruly character/nature (of a person)".

5.4 Analysis

The final dataset GerCo (German Collocations) comprises 4,732 adjective-noun phrases with 48 unique adjectives and 3,007 unique nouns.[18] The distribution for the binary labels is the following: 2,505 collocations and 2,227 non-collocations. Figure 5.5 shows the distribution of the binary labels in the GerCo dataset per adjective. The numbers for the more fine-grained classification are presented in table 5.4.

Label	Number	Distinct adjs	Distinct nns
collocations	2,505	47	1,837
free phrases	1,979	48	1,471
idioms	145	35	116
proper names	43	14	42
terms	18	6	17
free/idiom	37	14	35
free/proper	5	4	5
free/term	1	1	1

TABLE 5.4: GerCo dataset overview: number of instances for each label and number of distinct adjectives and distinct nouns per label.

There are four labels for non-collocations: *free phrase*, *idiom*, *proper name*, and *(technical) term*. Apart from that, there are 42 ambiguous cases in which the label varies depending on the context and which could not be resolved in the process of adjudication: *free phrase or idiom*, *free phrase or proper name*, *free phrase or term*. Each of the groups will be inspected and described separately in the following sections.

5.4.1 Collocations

Figure 5.5 illustrates the distribution of collocations in GerCo per adjective. The distribution is strikingly imbalanced for a large number of adjectives. Adjectives such as *frischgebacken* 'freshly baked', *scharf* 'sharp', or *starr* 'rigid' mostly participate in collocations, whereas *barock* 'baroque', *hölzern* 'wooden', and *teuer* 'expensive' tend to occur in free phrases. There is a weak positive correlation between the number of senses of an adjective in GermaNet and the amount of collocations the adjective is a constituent of in GerCo (Spearman r = 0.38). This indicates that the level of an adjective's polysemy plays only a secondary role in its collocability: an adjective with only two senses such as *starr* 'rigid' participates in more collocations than an adjective with 10 senses such as *hart* 'hard'.

Remember that in the data selection process, three relational adjectives were included in the dataset: *barock* 'baroque', *steinig* 'stony/rocky', and

[18] The dataset can be downloaded here: https://doi.org/10.57754/FDAT.rr563-my238.

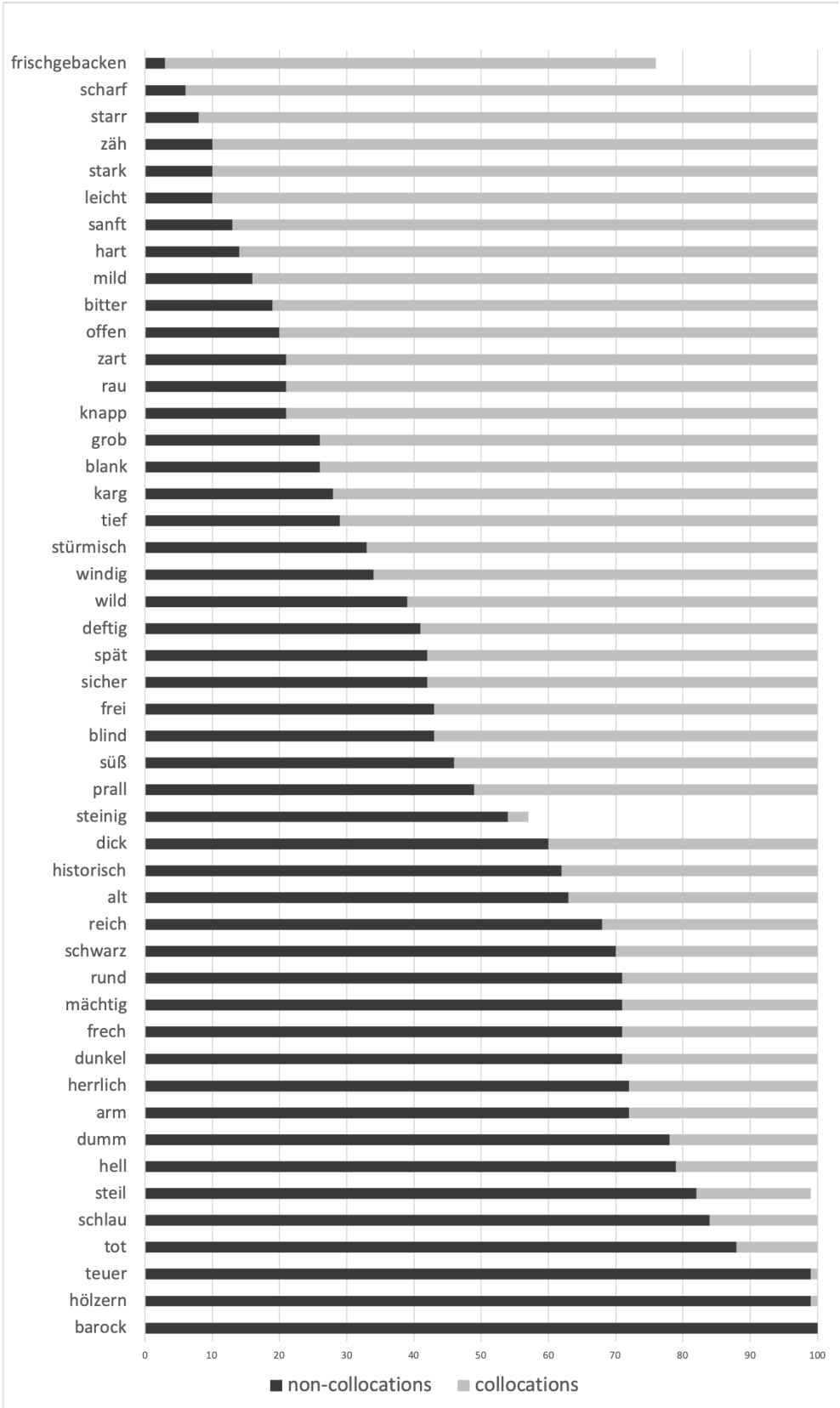


FIGURE 5.5: Distribution of collocations and non-collocations in GerCo per adjective.

historisch 'historical'. It was motivated by the fact that such adjectives can acquire non-relational readings and thus form collocations. The adjectives *barock* 'baroque' and *historisch* 'historical' have two senses both in GermaNet and the DWDS and were expected to participate in collocations. This is true for *historisch*: 38 collocations were identified. In a number of cases, it acquires the sense 'historic', as in *historische Entscheidung* 'historic decision'. Other examples illustrate the sense 'ancient, old' as in *historischer Stadtkern* 'historical city center'. For the adjective *barock*, on the other hand, no collocations were identified. Its second sense 'odd, peculiar' is not represented in the dataset. Finally, the adjective *steinig* that has only one sense both in GermaNet and the DWDS participates in two collocations: with nouns *Prozess* 'process' and *Lebensweg* 'life journey' in the sense 'difficult'.

It is important to keep in mind that the characteristics of the corpus used for the extraction of collocation candidates affect the results (see section 3.4.1). The underlying corpora for the Wortprofil are to a large extent comprised of newspaper articles. Although the news domain can be quite varied, not all word senses are equally represented there. It results in the imbalanced distribution of collocations and free phrases for some adjectives. Consider the adjective *scharf*: 94% of the phrases with this adjective are collocations. The literal meaning of the adjective – "good at cutting or piercing, pointed", is found only in four phrases out of the total of 100: *scharf* + *Gegenstand/Klinge/Messer/Schwert* 'sharp object/blade/knife/sword'. Another sense of *scharf* in German is "spicy, hot" as in *spicy chili/soup/food*. According to both annotators, it is a highly frequent sense in daily communication.[19] However, the dataset does not contain any examples of that sense. It does not imply that newspapers do not mention spicy food in their articles. Rather, such phrases are not frequent enough and thus the statistical association is not strong enough to be captured by the logDice measure. As a result, this sense of the adjective is not represented in the top-100 co-occurrences available in the Wortprofil.

The examples discussed above are cases of regular polysemy when a non-literal sense of an adjective is productive and is found in numerous phrases. The dataset also contains a few examples of very specific collocations that are so highly lexicalised that they have a dedicated Wikipedia article. One example is *wilde Ehe* (lit. 'wild marriage') which is a colloquial term for 'common-law marriage'. According to the DWDS, this expression is becoming obsolescent. Another example is *totes Gleis* lit. 'dead track' that stands for a train track that is no longer in use. They could be treated as technical terms. However, both annotators labeled them as collocations because the use of these expressions is not limited to one domain. There is a number of phrases that at the last step of the the decision tree could not be labeled neither as collocations nor as free phrases exactly for that

[19] This adjective caused a discussion among the lexicographers who are native speakers of German as to which sense is the literal one: 'sharp' or 'spicy'. This indicates that both senses are almost equally dominant in the mental lexicon of native speakers. However, the sense 'sharp' is listed first both in the DWDS and Duden (https://www.duden.de/rechtschreibung/scharf, *last accessed on 14.04.2023*) and is thus considered to be the most literal sense.

reason: their meaning is highly specific and they are used in one particular domain. This group of phrases labeled *technical terms* is described in the next subsection.

5.4.2 Technical terms

Terms were not suggested as a label option before the annotation and thus are not a part of the decision tree. The 18 expressions labeled in the final version as *terms* were identified in a data-driven manner. In the process of annotation, it became clear that some expressions do not fall into any of the categories from the annotation scheme. They were identified in the last step: the meaning of the base is literal, the meaning of the modifier is non-literal – so they might be considered collocations. However, the meaning of one of the constituents and/or the expression as a whole is very specific and its use is restricted to a particular domain. The meaning of such examples is sometimes not clear even to native speakers without additional information such as Wikipedia or dictionaries specialised in specific domains. Technical terms from the following domains were identified in the GerCo dataset:[20]

- **Vision**, e.g. *Blinder Fleck* 'blind spot' – is an obscuration of the visual field.
- **Astronomy**, e.g. *Dunkle Energie* 'dark energy' – an unknown form of energy.
- **Philosophy**, e.g. *Historischer Materialismus* 'Historical materialism' – the term used to describe Karl Marx's theory of history.
- **Economics**, e.g. *spätes Geschäft* 'late trading' – the trading that happens after the stock market closes.
- **Legal terms**, e.g. *freier Träger* 'independent sponsor' – institution that provides personnel and material resources for services and is not a public or administrative body.

5.4.3 Free phrases

In free phrases, both the modifier (adjective) and the head (noun) retain their literal meaning and the modifier is not lexically bound to the head. They form a large group in the GerCo dataset comprising 41% of all the phrases. As discussed in section 5.2.1, the distinction between literal and non-literal meaning is quite straightforward when the literal meaning is concrete (spatial and perceptional adjectives) which is reflected in the high IAA for such adjectives. The confusion is high for the adjectives that do not have a concrete sense such as *sicher* 'secure' (see section 5.3.3) or *frech* 'bold' (see section 5.2.1).

[20]The definitions stem from the dedicated Wikipedia articles.

5.4.4 Idioms

During the annotation, idioms are identified in step 2 of the decision tree (figure 5.2). If the head is has a figurative meaning, the phrase is labeled as idiom. This definition follows Mel'čuk (2012a)'s notion of an idiom. A closer inspection of the identified idioms shows that this category is not homogeneous. The following types of idioms can be identified in GerCo:

Semi-idioms

In semi-idioms, the meaning of only one constituent is figurative (Mel'čuk, 2012a). The semi-idioms in GerCo follow a similar pattern: the adjective retains its literal meaning, whereas the noun does not: (1) nouns from the semantic field ANIMAL referring to people, for example *dumm* + *Gans/Kuh/Ziege* lit. 'stupid goose/cow/goat', *frecher Hund* lit. 'cheeky/naughty dog', *schlauer Fuchs* lit. 'clever fox'; (2) metonymy where a part of the body refers to a person: *schlauer Kopf* lit. 'smart head', *freche Zunge* lit. 'cheeky/rude tongue'.

Full idioms

In full idioms, both constituents are used in a figurative sense (Mel'čuk, 2012a). There are two kinds of full idioms: *decomposable* and *non-decomposable*.

Decomposable idioms are idioms in which the meaning is composed of *"idiomatic meanings of its parts"* (Nunberg et al., 1994, p. 507). In such phrases, the meanings of both constituents are figurative, but they can be mapped onto parts of the actual meaning of an idiom. For instance, in the idiom *bitteres Erwachen* 'bitter awakening' which means 'unpleasant realisation', *bitter* refers to *unpleasant* and *awakening* refers to *realisation*. Both constituents add to the meaning of the whole phrase.[21]

Most of the phrases annotated as idioms in GerCo fall into this category. The most typical semantic fields of the noun bases in decomposable idioms in GerCo are:

- ANIMALS: *alter Hase* lit. 'old rabbit' – an experienced (*alt*) person (*Hase*); *armes Schwein* lit. 'poor pig' – someone (*Schwein*) who is pitied (*arm*); *blindes Huhn* lit. 'blind chicken' – an ignorant or stupid (*blind*) person (*Huhn*); *harter Hund* lit. 'hard dog' – a tough (*hart*) person/man (*Hund*); *schwarzes Schaf* lit. 'black sheep' – a different, odd *schwarz)* member of a group (*sheep*).

- BODY: *wildes Herz* lit. 'wild heart' – passionate (*wild*) character (*Herz*); *starke Nerven* lit. 'strong nerves' – ability (*Nerven*) to endure a lot (*stark*).

[21] Note the difference between the terms *decomposability* and *transparency* of an idiom. Decomposability refers to the possibility of segmenting the idiomatic meaning into the same (syntactic) structure the actual meaning of an idiom has. Transparency refers to a relation between the idiomatic and literal meaning of an idiom. For instance, the idiom *to saw logs* is transparent on the basis of similarity between its literal and idiomatic meaning: the sound of snoring is compared to the sound of sawing logs (Sailer, 2021).

- ARTIFACT: *bitterer Kelch* lit. 'bitter cup' – something (*Kelch*) unpleasant (*bitter*); *blindes Werkzeug* lit. 'blind instrument' – an unaware (*blind*) person used by someone (*Werkzeug*); *(mit) offenen Karten (spielen)* lit. '(to play with) open cards' – to be (*Karten spielen*) honest (*offen*).
- SUBSTANCE: *dicke Luft* lit. 'thick air' which means 'tense atmosphere'; *bittere Arznei* lit. 'bitter medicine' that stands for 'unavoidable negative measure'.

Non-decomposable idioms are holistic units the meaning of which cannot be simply decomposed into figurative meanings of its parts. Examples in the GerCo dataset are actually parts of larger idiomatic units. For instance, the phrase *harte Bandage* lit. 'hard bandage' is a part of the idiom *mit harten Bandagen kämpfen* fig. 'to play hardball'. It means 'to use every means possible to achieve a goal', so *mit harten Bandagen* refers to *every means possible*. Similarly, the adjective-noun pair *offenes Messer* lit. 'open knife' is a part of the idiom *ins offene Messer laufen (lassen)* ≈ 'to (let someone) walk straight into a trap' where *offenes Messer* stands for *trap*.

5.4.5 Proper names

Proper names were identified in most cases already in the first step of the annotation. According to Mel'čuk (2012a), complex proper names are compositional and are a subgroup of clichés. This category in GerCo contains geographical names such as *Schwarzes Meer* 'the Black Sea' and *Totes Meer* 'the Dead Sea', *Wilder Kaiser* 'Wild Kaiser Mountains'; art objects such as *Das Schwarze Quadrat* 'The Black Square' and *schwarze Madonna* 'Black Madonna'; institutions such as *Freie Universität*, *Alte Pinakothek*, *Altes Rathaus*.

One category of proper names posed a challenge to the annotators. It is comprised of movie titles that were not known to the annotators. For instance, *Bitterer Reis* 'Bitter Rice' and *Bitterer Honig* 'A Taste of Honey' (lit. 'bitter honey') are movie titles, but might as well be free phrases. The movie title *Die Süße Haut* 'The Soft Skin' (lit. 'sweet skin') might be a collocation or an idiom or a free phrase in a certain context. However, the inspection of the corpus examples containing these phrases shows that they refer only to the titles of the movies, therefore, a single label *proper name* could be assigned to them.

5.4.6 Ambiguities

During the adjudication, some of the disagreement cases could not be resolved because of their ambiguity. This concerns only the disagreement on the fine-grained labels for non-collocations. There are two groups of ambiguous phrases:

- **Free phrase/proper name**: for instance, *der Arme Poet* 'The Poor Poet' is a painting by Carl Spitzweg, and the majority of the examples in

the corpora refer to the painting. However, there are also examples that have nothing to do with the painting. Other examples of this kind of ambiguity from GerCo are *blinder Fisch* 'blind fish' (part of a book title), *wilder Schwan* 'wild swan' (book title).

- **Free phrase/idiom:** this type of ambiguity is quite common – 37 phrases in GerCo were annotated as free/idiom. In these phrases, both constituents can be used either literally or metaphorically. For instance, *runder Tisch* 'round table' can refer to a table of round shape (free phrase) or to a meeting of equal partners (idiom). More examples are *bitter Frucht* 'bitter fruit', *schwarzes Brett* 'bulletin board' (lit. 'black board'), *hartes Holz* lit. 'hard wood' can be used in the sense 'to be made of hard wood' when describing a person.

  There are examples in which an adjective in a figurative meaning combines with a number of semantically related nouns that can be either used in their literal sense or metaphorically: *steinig* + *Pfad/Strecke/Weg* 'rocky path/route/road'. The metaphoric interpretation refers to a difficult phase in life. Another productive in metaphors adjective is *steil* 'steep': *steil* + *Höhenflug/Sinkflug/Steigflug/Sturzflug* 'steep + flight/ descent/ascent/nosedive'. These nouns are metaphors for *increase/rise*. There is a large body of research on the ambiguity between literal and figurative meaning of idioms, see Sporleder et al. (2010); van Ginkel and Dijkstra (2020); Haagsma et al. (2020); Wagner (2020).

5.4.7 Concreteness

It has been observed in Strakatova and Hinrichs (2019, p.106) that abstract nouns tend to form more collocations than concrete ones. In order to investigate this observation on a larger scale, for all the nouns in the GerCo dataset, concreteness scores were obtained. The scores stem from Köper and Schulte im Walde (2016), a collection of over 350,000 German words with ratings of concreteness, arousal, imageability, and valence. Köper and Schulte im Walde (2016) calculate the ratings automatically using a supervised learning algorithm. The distribution of the concreteness scores for the nominal bases in the GerCo dataset presented in Figure 5.6 supports this observation. The lower the score, the more abstract a noun is: the most abstract noun in the dataset is *Bemühen* 'effort' with the concreteness score = 0.90 and the most concrete one is the noun *Kochlöffel* 'cooking spoon' with the concreteness score of 8.60.

There is a tendency for nominal heads of collocations to be more abstract in comparison to free phrases: the mean concreteness score for collocations is 4.2. For free phrases, the mean value is higher – 5.74. However, there is no clear-cut boundary between the two categories. There is an overlap of collocations and free phrases with medium concreteness scores. Examples of collocations and free phrases are found in both extremes. Among collocations, the most concrete noun is in the pair *süßer Vogel* 'sweet/nice bird'

(8.18); among free phrases, the most abstract one is in the phrase *freche Zumutung* 'unreasonable/cheeky demand' (1.50). Concreteness mean values of nouns for each fine-grained label are reported in Table 5.5.

Label	Concreteness mean
collocations	4.28
free phrases	5.60
idioms	5.74
proper names	5.95
terms	4.66
free/idiom	5.55
free/proper	6.56
free/term	6.75

TABLE 5.5: Mean concreteness scores of nouns in GerCo dataset for each label.

Among the unambiguous phrases, proper names and idioms yield rather higher scores as well. Figure 5.7 illustrates the distribution of concreteness scores for the idioms in the GerCo dataset. The nominal bases of most idioms belong to concrete semantic fields such as SUBSTANCE, ARTIFACTS, ANIMALS, FOOD, BODY PARTS:

- *teures Pflaster* lit. 'expensive paving/cobble', fig. 'locality or region associated with high cost of living';
- *tote Hose* lit. 'dead pants', fig. 'boredom, lack of activity, variety, or entertainment; nothing (more) going on.';
- *alter Hase* lit. 'old rabbit', fig. 'person highly experienced in a particular discipline or function';
- *hartes Brot* lit. 'hard bread', fig. 'a heavy, arduous, physically or mentally stressful job'
- *offenes Ohr (haben)* lit. '(to have) an open ear', fig. 'to listen to someone's concerns willingly';

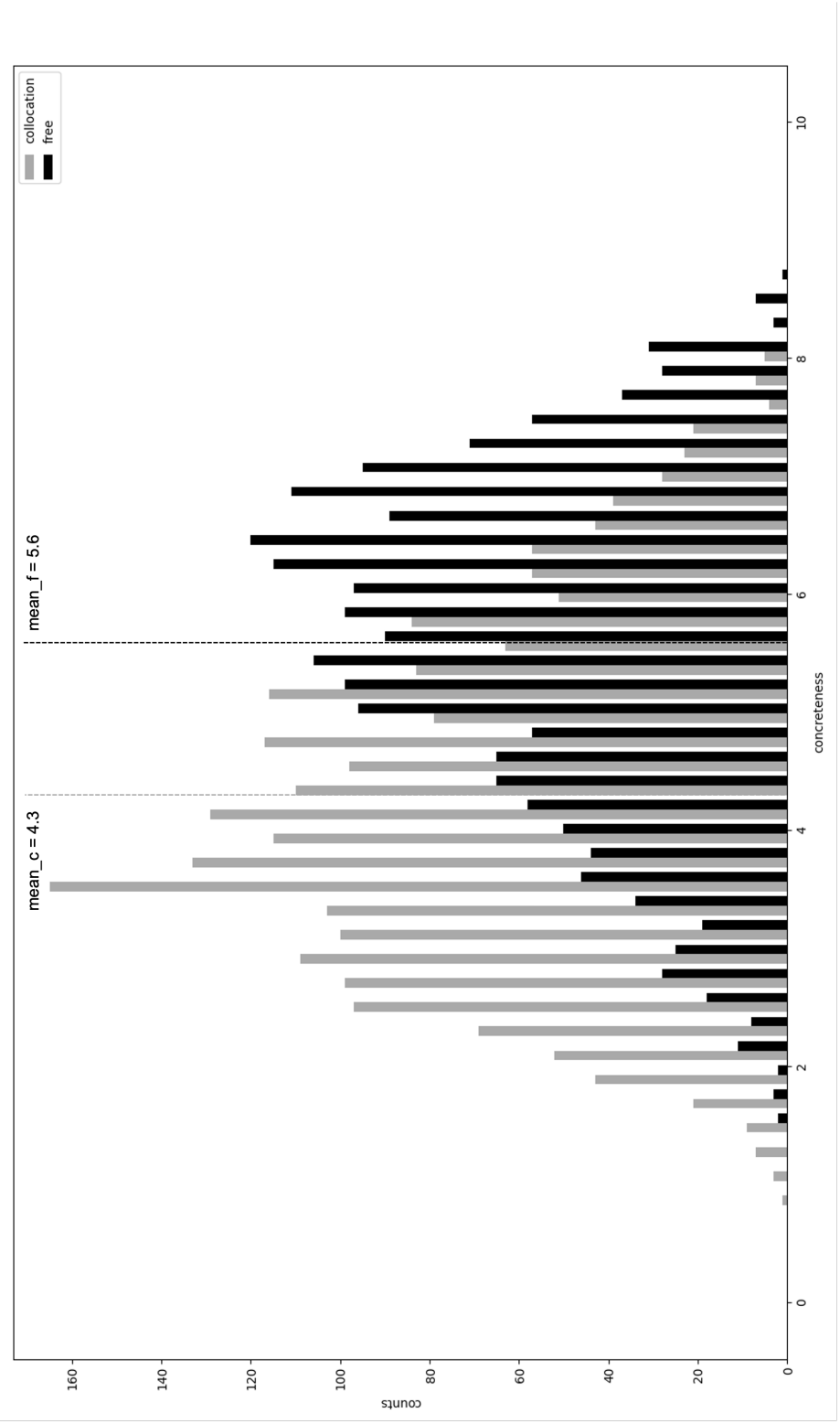


FIGURE 5.6: Concreteness scores for the nominal bases in collocations vs free phrases in GerCo: 0 score for the least concrete nouns and 10 for the most concrete ones. The dashed lines mark the mean values for collocations and free phrases.

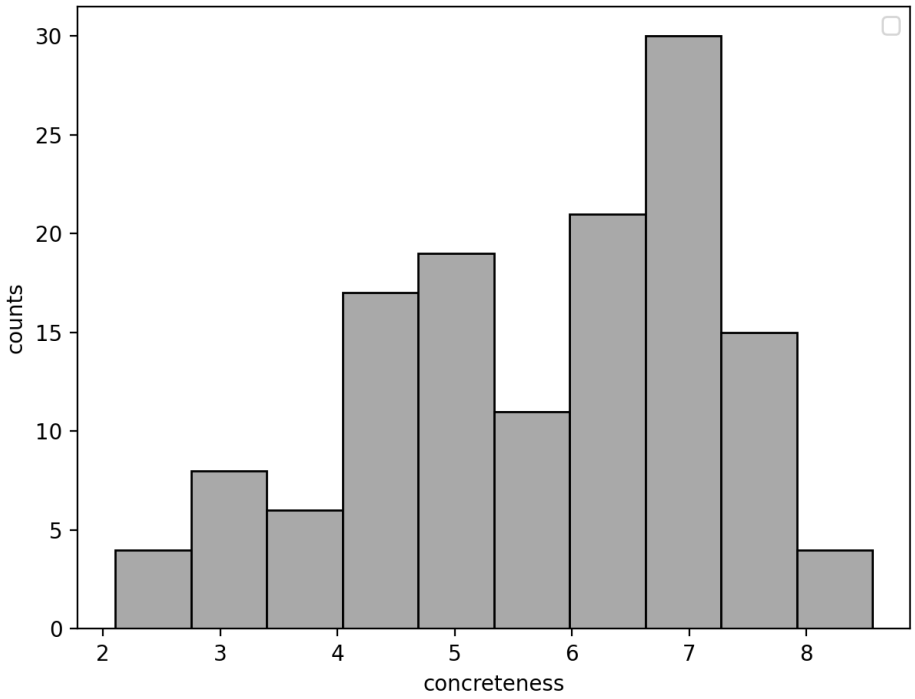


FIGURE 5.7: Distribution of concreteness scores for idioms in the GerCo dataset.

5.5 Experiments: automatic binary classification

The classification scheme for adjective-noun combinations proposed in this thesis was evaluated in terms of inter-annotator agreement and in a thorough qualitative analysis of disagreement cases and of identified categories (*collocation*, *free phrase*, *idiom*, *proper name*, *technical term*). This section will report the results of machine learning experiments conducted with the GerCo dataset. The aim of the experiment is twofold: testing the robustness of the proposed classification and evaluation of various models of automatic collocation identification. The experiments have been published in Strakatova et al. (2020), the technical implementation of the experiments has been performed by Neele Falk, one of the authors of the paper.[22]

As the dataset is imbalanced in terms of fine-grained categories (see Table 5.4), only two categories that have more than 1,500 instances will be distinguished in the experiments: *collocations* and *free phrases*. Therefore, the task at hand is a binary classification of adjective-noun phrases into free phrases and collocations. It will be investigated what feature representations provide a useful source of information to solve the task. In the first experiment, the features are represented as lexical association measures. In the second line of experiments, static and contextualized word embeddings will be used.

[22]The data can be downloaded from http://hdl.handle.net/11022/0000-0007-DABA-2.

5.5.1 Data

The first step in data preparation is enriching the dataset with context. This is motivated by use of context-aware representations in the second experiment. For each AN phrase in the dataset, a context sentence has been extracted from one of the following corpora: Wikipedia (dumps from 2017, 2018, 2019), the One Million Posts Corpus (Schabus et al., 2017; Schabus and Skowron, 2018), the German proceedings from the EuroParl corpus (Koehn, 2005; Tiedemann, 2012) and the German Political Speeches Corpus (Barbaresi, 2018). All the corpora are publicly available. In total, 3,652 sentences were extracted: one sentence per phrase. The main criterion was the length of the sentence, the limit was set to a minimum of 15 and maximum of 30 tokens.

Six different data splits were created in order to compare the performance of the models for all the 48 adjectives. In each split, the adjectives in the training and the test sets do not overlap. In this way, the models' ability to generalize can be investigated. Figure 5.8 shows the distribution of collocations and free phrases for each adjective. For most adjectives, the distribution is imbalanced.

In order to examine whether additional information about the meaning of the adjective helps to classify, the definitions of the literal sense of each adjective (listed in Appendix B) was taken into consideration in the second experimental setup. Thus, every training instance includes an AN phrase, one context sentence, definition of the adjective's literal sense, and a label: 1 for collocation, 0 for free phrase.

A random sample of 200 instances (83 free phrases, 117 collocations) was given to two native speakers and students of linguistics to measure *human performance* on this task. Based on the average number of correctly annotated instances of both annotators, the accuracy was computed. It equals 0.83.

5.5.2 Experiment 1: association measures

As presented in section 3.4, automatic collocation identification is traditionally based on a variety of lexical association measures (AMs). These measures are computed on the basis of the individual and joint frequencies of the base and its collocate. Based on the calculated AM scores, collocation candidates extracted from a corpus are ranked: the higher the score, the stronger is the lexical association between the two words. High AM scores are an indication that a phrase might be a collocation.

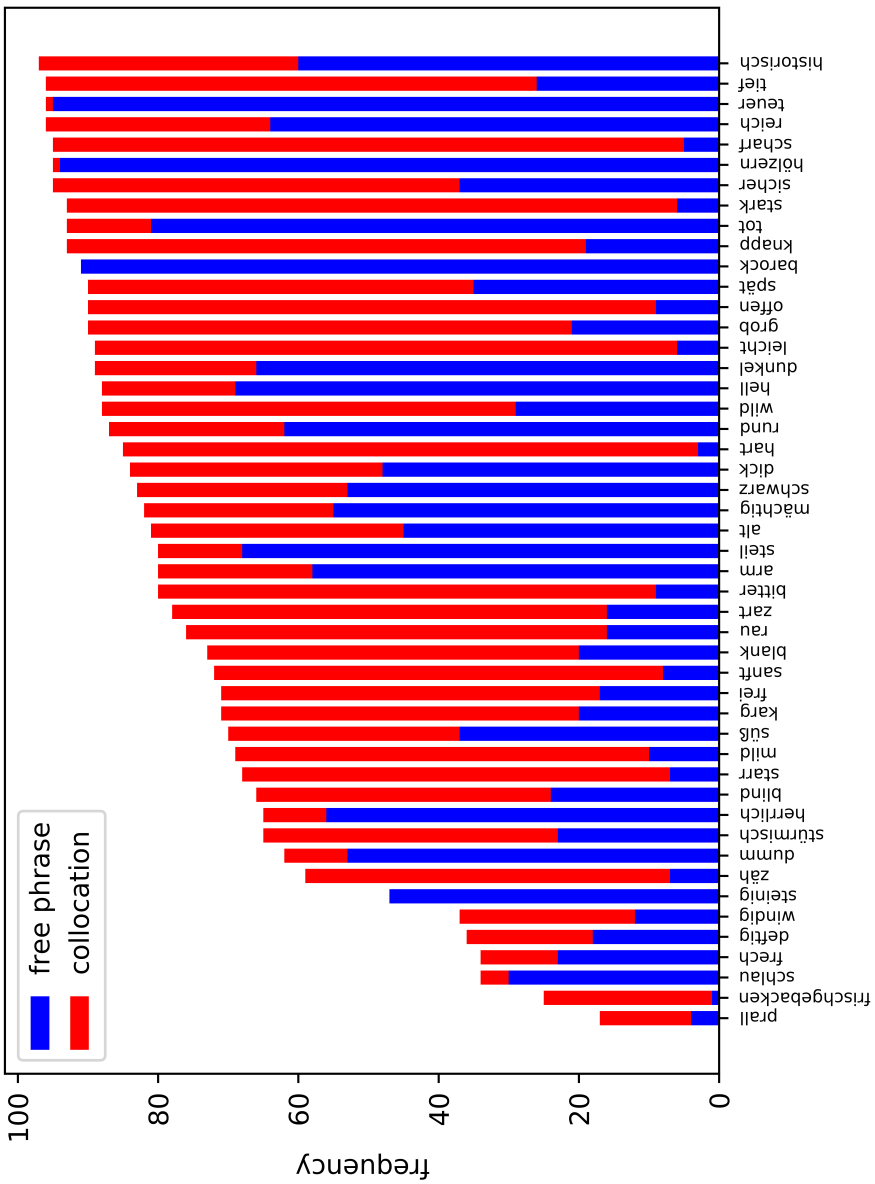


FIGURE 5.8: Class distribution for each adjective in the test set, from Strakatova et al. (2020, p. 4372).

There are two approaches to classifying phrases into collocations and free phrases using AMs. The traditional approach is to set a cut-off threshold: any phrase that has an association score above the threshold is classified as collocation, and the rest of the phrases are non-collocations. In the second approach proposed by Pecina (2008), AM scores can be used as features in a machine learning classification setting. In the experiments conducted by Pecina (2008), this method outperformed the traditional approach by a large margin. Building on this evidence, the experiment with the GerCo data set reported in Strakatova et al. (2020) relies on AM measures as features in a linear and a nonlinear classifier.

As a first step, co-occurrence frequencies of all adjective-noun pairs, 42,445,060 pairs in total, have been extracted from a number of treebanks: Wikipedia 2017 and Wikipedia 2018 (de Kok and Pütz, 2019) and de-cow16ax (Schäfer and Bildhauer, 2012; Schäfer, 2015). Association measure scores for the extracted phrases are computed using the UCS-toolkit[23] developed by Evert (2004). The following 22 AMs are available in the toolkit: log.likelihood, Dice, Jaccard, MI, MI2, MI3, MS, Poisson.Stirling, average.MI, chi.squared, chi.squared.corr, frequency, gmean, local.MI, odds.ratio, odds.ratio.disc, random, relative.risk, simple.ll, t.score, z.score. The scores are normalized and scaled between 0 and 1 in order to be used as features in the classifiers.

First, the classical approach is tested: the threshold for each AM is tuned on the training set. Each AN pair in the test set is classified depending on the best threshold. Based on a combination of the individual predictions of each AM classifier, the final prediction is made according to the majority vote (Bishop, 2006).

Secondly, following Pecina's (2006) method, Strakatova et al. (2020) use AMs as input for training both linear and non-linear classifiers which predict the class based on a combination of AMs associated with a given AN pair. The classifiers' weights are optimized to minimize cross-entropy loss on the training dataset. An adjective-noun pair is encoded as a vector of 22 dimensions, with each dimension corresponding to a type of association strength.

The linear classifier is based on a Support Vector Machine with a Radial Basis Function (RBF) kernel and l2 regularization.[24] The nonlinear classifier is a feed-forward neural network with one hidden layer of size 4 and a `ReLU` non-linearity and apply early stopping. Hidden layer size was tuned on validation (for details see Strakatova et al. (2020, p. 4376). Each classifier was trained all data splits, and overall performance was evaluated by averaging the results across these splits.

The results are displayed in Table 5.6. The majority class (collocation) baseline is not outperformed by any of the classifiers. These results indicate that association measures alone are not able to differentiate between collocations and free phrases. Combining the association measures and mapping

[23] `http://www.collocations.de/software.html`, *last accessed November 22, 2019*

[24] Using standard implementation of sklearn (Pedregosa et al., 2011).

model	validation accuracy	test accuracy
majority baseline	0.5408	**0.5344**
majority vote of threshold-classifiers	0.5350	0.5003
linear classifier (SVM)	0.5319	0.5028
combined AMs with nonlinear classifier	0.545	0.5256

TABLE 5.6: Average validation and test accuracy on all splits with association measures, from Strakatova et al. (2020, p. 4372)

them into a new feature space does not improve the performance. This result is not surprising and a plausible explanation lies in the way the GerCo dataset was created. Remember that both collocations and free phrases stem from n-best lists compiled in the Wortprofil. Thus, most of the AN pairs in the dataset, both collocations and free phrases, exhibit high statistical association between their components. The results support the hypothesis that lexical association measures alone do not suffice for differentiating between collocations and free phrases as the latter may also exhibit strong lexical association.

5.5.3 Experiment 2: word embeddings

Since the results of the first experiment indicate that additional semantic information is necessary, in the second line of experiments, Strakatova et al. (2020) use a richer source of semantic information, more specifically, word representations or *word embeddings*.

Word embeddings capture semantic information about words by encoding information about similar words and words they frequently appear with. Static word embeddings, such as GloVe (Pennington et al., 2014) or Word2Vec (Mikolov et al., 2013), derive a word's meaning from its usage across a large body of language (corpora). However, a significant drawback of static embeddings is the deficiency in handling multiple meanings, as all senses of a word are compressed into a single vector. This problem can be remedied by computing dynamic representations of words conditioned on local context (Peters et al., 2018; Devlin et al., 2019). Such contextualized representations are dynamic in the sense that they are able to capture different meanings of a word depending on the context.

The second line of experiments in Strakatova et al. (2020) tests whether semantic representations of words are capable of differentiating between free phrases and collocations. Additionally, the amount of contextual information required for accurate classification is analyzed. Strakatova et al. (2020) hypothesize that if the sense of the adjective is primarily determined by the noun it modifies, static word representations would suffice to solve the task. If the phrase is ambiguous in isolation or nearby words from the local context have a great impact on the meaning of the adjective, contextualized embeddings are likely to perform better. Also the impact of adding the definitions of the literal meaning of the adjectives is analyzed.

Strakatova et al. (2020) create four setups. In each of them, a nonlinear classifier with one hidden layer and a `ReLU` non-linearity is trained given different feature representations as input:

- adj + noun
- adj + noun + context sentence
- adj + noun + sense definition of the adjective's literal sense
- adj + noun + context sentence + sense definition

The experiments with static embeddings make use of pretrained word embeddings[25] trained with the finalfrontier utility[26] on subcorpora (Wikipedia, taz, EuroParl) of TüBa D/DP (de Kok and Pütz, 2019). For extracting contextualized embeddings, the bidirectional transformer (BERT) is used (Devlin et al., 2019). The experiment relies on the bert-base-german-cased model, trained by deepset.ai[27] on corpora of different domains. See further technical details in Strakatova et al. (2020).

An AN phrase is represented by concatinating the word embeddings of the adjective and the noun. Since BERT splits some words into smaller subword units, a single embedding for a word is obtained by averaging the corresponding subword embeddings. A context sentence and a sense definition are encoded with a bi-directional LSTM that takes the corresponding word or word piece embeddings as input. For more technical details consult Strakatova et al. (2020).

setup	embeddings	validation accuracy	test accuracy
majority baseline		0.5457	0.5087
phrase	static	0.8445	0.7138
phrase	**contextualized**	**0.8857**	**0.7415**
phrase + context	static	0.8357	0.7123
phrase + context	contextualized	0.8608	0.7337
phrase + definition	static	0.8426	0.71
phrase + definition	contextualized	0.8705	0.7390
phrase + context + definition	static	0.8414	0.7112
phrase + context + definition	contextualized	0.8645	0.7338
phrase + context + definition	*human*		*0.83*

TABLE 5.7: Averaged results on the validation and test splits for different setups with static vs. contextualized embeddings, from Strakatova et al. (2020, p. 4374).

[25] https://finalfusion.github.io
[26] https://github.com/finalfusion/finalfrontier
[27] https://deepset.ai/german-bert

The results are shown in Table 5.7. Contextualized representations slightly outperform static embeddings. Adding the context and/or the literal sense definition does not improve the general results. However, Strakatova et al. (2020) do not rule out a possibility that the additional information improves the classification for some adjectives and reduces it for others. Whether this is true or not, the results indicate that the type of information might not be robust enough to generalize well for different adjectives.

Table 5.8 presents the accuracy for selected adjectives that were either difficult to classify or exhibited significant performance variations across different embedding types or configurations. For both types of word representations similar adjectives pose a challenge: *frischgebacken* 'freshly baked', *karg* 'sparse', *mächtig* 'powerful'. The adjective *frischgebacken* stands out: it did not pose any challenges to the human annotators, but its corpus frequency is low which makes it challenging for the models. Incorporating contextual information improves classification accuracy for both embedding types. Including the sense definition further enhances performance for contextualized embeddings. However, for the adjective *karg* 'sparse', the effect is reversed: adding context and sense definition does not improve performance and even leads to a decline when using static embeddings.

	static embeddings per-class Accuracy			**contextualized embeddings per-class Accuracy**		
adjective	phrase	+context	+sensedef	phrase	+context	+sensedef
steinig 'stony'	0.08	0.12	0.15	0.85	**0.9**	0.79
karg 'sparse'	0.21	0.21	0.15	**0.25**	0.25	0.25
frischgebacken 'freshly baked'	0.27	0.31	0.31	0.12	0.27	**0.38**
mächtig 'powerful'	0.3	**0.39**	0.27	0.25	0.25	0.25
windig 'windy'	0.45	0.45	**0.47**	0.31	0.31	0.4
frech 'bold'	**0.46**	0.4	0.34	0.39	0.39	0.42
schlau 'smart'	**0.66**	0.63	0.54	0.11	0.11	0.11
dumm 'stupid'	**0.7**	0.54	0.59	0.17	0.22	0.25
mild 'mild'	0.7	0.77	0.69	**0.8**	0.74	0.69
prall 'firm'	**0.67**	0.61	0.61	0.5	0.61	0.56
deftig 'savoury'	0.86	0.84	0.84	0.81	0.86	**0.95**
herrlich 'wonderful'	0.83	0.86	**0.91**	0.5	0.52	0.59
tot 'dead'	0.52	0.49	0.6	**0.88**	0.78	0.85
teuer 'expensive'	0.65	0.64	0.74	0.89	0.8	**0.92**

TABLE 5.8: Sample of adjectives with test set accuracy, from Strakatova et al. (2020, p. 4375)

There are adjectives that are hard to model only for one embedding type. For instance, *steinig* 'stony' is a difficult adjective for static embeddings (best result 0.15) whereas contextualized embeddings perform much better (best result 0.9); the adjectives *dumm* 'stupid' and *schlau* 'smart' are difficult for contextualized embeddings, but easy for static representations. Strakatova et al. (2020) come to the conclusion that contextualized embeddings, overall, show more significant improvements when additional context and sense

representations are included, whereas models based on static embeddings often experience a decline in accuracy. These representations seem to depend more heavily on the noun. Despite the expectation that static embeddings might struggle with polysemous adjectives, they remain quite effective in classifying AN pairs into collocations and free phrases.

5.5.4 Discussion

The experiments by Strakatova et al. (2020) reveal that using word embeddings as features outperforms approaches based on association measures. The best performance (contextualized, phrase only) on the test dataset is 0.74 which is 9% less than human accuracy. However, this result is 24% higher in comparison to the baseline. The accuracy of all the tested models that rely on word embeddings is higher than 0.71. All the models are capable of detecting meaning shifts in the adjective under the influence of its head noun. Note that the models were tested on adjectives they did not see in the training sets which means that they are capable of generalizing in two thirds of the cases. This indicates that the annotation scheme for classifying AN phrases into collocations and free phrases presented in this chapter is quite robust. A closer examination of the models' performance for individual adjectives shows that the results are not homogeneous: shifts of meaning in some adjectives such as *deftig* 'savoury' and *mild* 'savoury' are detectable, whereas some adjectives such as *karg* 'sparse' pose a problem to all the models in all the setups. This maybe due to the fact that the nominal bases of *karg* in free phrases are more varied than in collocations which confuses the model. In contrast, the nominal heads of *deftig* in free phrases are homogeneous and belong to the semantic class FOOD.[28]

Additional context brings only a slight improvement which suggests that almost all the necessary information is contained in the bigram itself and the head noun suffices for distinguishing between a collocation and a free phrase in most cases. This insight goes in line with the results of the IAA study described earlier in section 5.3.3: the annotators received the AN pairs for annotation without context and the agreement was substantial.

5.6 Summary

This chapter described the construction of the gold standard dataset of German adjective-noun phrases (GerCo) that contains 2,505 collocations and 2,227 non-collocations. The inter-annotator agreement is substantial which indicates that the annotation scheme represented as a decision tree is reliable. The robustness of the annotation scheme is further backed up by machine learning experiments conducted on the GerCo dataset that yield as best result accuracy of 0.74 in the task of binary classification of AN pairs into collocations and free phrases. In both IAA study and ML experiments,

[28] See the discussion about lexical diversity in the next chapter (section 6.2.1).

a perfect agreement/accuracy could not be achieved. In most cases it can be explained by the ambiguity of either one or both elements of the phrases.

The main goal of creating a new dataset of adjective-noun phrases was to provide a semantically motivated empirical basis for studying the relations between the constituents of adjective-noun collocations and free phrases. The next chapter describes how the theoretical frameworks discussed in Chapter 4 are combined into one annotation scheme for semantic annotation of adjective-noun pairs.

Chapter 6

Attributes in adjective-noun phrases

The nature of institutionalized phrases such as collocations is different from that of free phrases. In the latter, the restrictions are only of semantic nature (Wanner, 1996) licensing phrases such as in example (31a) and preventing a language speaker from creating phrases like example (31b). The noun *rain* being a PHENOMENON and *pain* being a STATE permit the combination with the adjective *incessant*, whereas the ARTIFACT *teapot* does not allow for it.

(31) a. *incessant rain, incessant pain*
b. **incessant teapot*

Collocations, on the other hand, are restricted lexically: the nouns *rain* and *pain* semantically allow their combination with an intensifier, but the choice of the acceptable intensifier is unpredictable as illustrated in (32a) and (32b).

(32) a. *heavy rain* – **?** *exquisite rain*
b. *exquisite pain* – **?** *heavy pain*

As the main objective of this thesis is to establish an inventory with a broad semantic coverage, both collocations and free phrases should be taken into account. The GerCo dataset described in the previous chapter contains both types of phrases and is thus a suitable empirical basis for the task. A potentially suitable inventory of relations has been proposed in Chapter 4. More specifically, the set of *attributes* available in the lexical-semantic net GermaNet provides a promising basis for the sought inventory. In this chapter, this proposal will be tested.

The difference between collocations and free phrases in terms of lexical restrictedness suggests that their annotation can be approached differently. In all the free phrases in GerCo, the adjectives retain their literal meaning. For example, the adjective *schwarz* 'black' carries its color sense in all the phrases identified in the previous annotation step as 'free': *black jeans, limousine, shirt, etc*. The relations in collocations are unpredictable and have to be annotated manually as the only information available in GerCo is that the meaning of the adjectival modifier in collocations is non-literal, and most adjectives in the dataset are highly polysemous.

This chapter begins with a detailed description of semantic annotation of the GerCo dataset with attributes. The gold standard dataset was given the

acronym GerCoAt (German Collocations with Attributes) is then analysed with regard to the status of the phrases: what relations are characteristic of collocations? are they different from those typically encountered in free phrases? These questions will be first discussed in section 6.2 where a qualitative analysis of the dataset will be presented. Sections 6.4 and 6.5 will address these questions in a series of machine learning experiments with the created dataset. Apart from that, this chapter explores the possibility of creating a general annotation template for nouns that would facilitate lexicographic work on collocations that will be presented in section 6.3.

6.1 Annotation process

The annotation of the GerCo dataset has been performed separately for collocations and free phrases. In both cases, two human annotators annotated the data independently from each other so that inter-annotator-agreement (IAA) could be calculated. The two annotators are native speakers of German and advanced students of Computational Linguistics with a solid background in lexical semantics familiar with the phenomenon of collocation. They received detailed guidelines for the annotation and were encouraged to use any additional resources as necessary (such as the DWDS, GermaNet, Wikipedia, etc). The initial inventory proposed in Chapter 4 had to be modified after the first round of annotation based on the disagreement between the annotators. During the adjudication, the annotators worked together with the author of this thesis and two colleagues at the University of Tübingen, Prof. Dr. Erhard Hinrichs and Neele Falk. The labels from the initial annotation scheme were discussed and adjusted.

As discussed in the previous chapter, ambiguity is a common feature of collocations. In order to avoid differences in the interpretation of the phrases during annotation, for each adjective-noun pair, a context sentence has been extracted from publicly available corpora (Strakatova et al., 2020). The total of 3,652 pairs were provided with a context sentence with the length between 15 and 30 words. The sentences had been extracted automatically, and manually controlled by three linguists.

The semantic annotation is performed in three stages: (1) annotation of the collocations by two annotators (section 6.1.1); (2) adjudication of disagreement cases and adaptation of the annotation scheme (section 6.1.2); (3) semi-automatic annotation of the free phrases (section 6.1.3).

The initial annotation scheme is based on the *artificial concepts* from the adjectival taxonomy in GermaNet as described in sections 4.5.1 and 4.5.2. For convenience, the table containing all the artificial concepts available in GermaNet that was presented in section 4.5.1 is also presented below (table 6.1). In total, there are 80 artificial concepts.

The names of artificial concepts in GermaNet are noun-adjective compounds such as `altersspezifisch` 'age-related' and `gewissheitsspezifisch` 'certainty-related'. They were changed into names more typical for attributes by dropping the adjectival head of the compounds in the following manner:

- `altersspezifisch` 'age-related' → `Alter` 'age'
- `gewissheitsspezifisch` 'certainty-related' → `Gewissheit` 'certainty'
- `dauerspezifisch` 'duration-related' → `Dauer` 'duration'

6.1.1 Initial annotation of collocations

The empirical basis for testing this attribute inventory, the dataset GerCo, contains 48 distinct adjectives. The majority of them are polysemous and belong to at least two synsets in GermaNet.[1] Three adjectives in the dataset do not have any instances of collocations and are only represented by free phrases: *steinig* 'stone' , *barock* 'baroque', and *teuer* 'expensive'. These adjectives will be annotated in the third step of the annotation which is described in section 6.1.3.

In the annotation guidelines,[2] each attribute has been defined and provided with examples of adjective-noun pairs. In order to avoid biasing the annotators, the examples in the guidelines contain adjectives that are not present in the GerCo dataset, see examples (33a), (33b), and (33c) for illustration.[3] One attribute was added to the scheme that is not present in the set of GermaNet artificial concepts, namely, **`type`**, as in *green tea* is a `type` of tea.

(33) a. **`age`** – number of lived years (for living beings but also for inanimate objects). E.g.: *young woman*, *new car*

b. **`certainty`** – clarity, certain knowledge, without doubt. E.g.: *predictable outcome*, *clear rejection*

c. **`duration`** – temporal extension. E.g.: *longtime cooperation*

The task was explained to the annotators, and they were instructed to read the list of the available attributes with the definitions and examples and make sure that the meaning of each attribute is clear and unambiguous.

Altogether, the annotators were asked to complete three tasks:

1. **Select the suitable attribute** for each adjective-noun phrase in context. The list of phrases for annotation is given per adjective, summing up to 45 lists. In each list, all the attributes from GermaNet available for the adjective are provided. For each attribute, its definition and a couple of examples are given. If none of the suggested attributes reflects the conveyed meaning in a given phrase, it should be annotates as "unknown" and a suggestion of an attribute should be made in the comments section.

[1] Except for the adjectives *frischgebacken* 'freshly-baked', *steinig* 'stony', *schlau* 'smart', and *herrlich* 'wonderful'.

[2] The guidelines are presented in Appendix C.

[3] The attributes and definitions are given here in English. The original ones in German are presented in Appendix C.

sem class	**artificial concepts = attributes**
Perceptional	`Helligkeit` 'brightness', `Farbe` 'color', `Geräusch` 'sound', `Geschmack` 'taste', `Geruch` 'smell', `Oberfläche` 'surface', `Sicht` 'sight'
Spatial	`Dimension` 'dimension', `Richtung` 'direction', `Lokalisierung` 'localisation', `Herkunft` 'origin', `Raumverteilung` 'spatial distribution', `Form` 'shape', `Existenz` 'existence'
Temporality-related	`Zeit` 'time', `Geschwindigkeit` 'speed', `Alter` 'age', `Gewohnheit` 'habit', *Dauer* 'duration'
Motion-related	`Bewegung` 'movement'
Material-related	`Beschaffenheit` 'texture', `Zustand` 'state', `Stabilität` 'stability', `Konsistenz` 'consistency', `Reife` 'ripeness', `Feuchtigkeit` 'dampness', `Reinheit` 'purity', `Gewicht` 'weight', `Physik` 'physics', `Chemie` 'chemistry', `Temperatur` 'temperature'
Weather-related	`Klima` 'climate', `Landschaft` 'landscape'
Body-related	`Belebtheit` 'life', `Konstitution` 'constitution', `Krankheit` 'affliction', `Geschlecht` 'sex', `Körpergefühl` 'desire/feeling', `Aussehen` 'appearance', `Körperzustand` 'bodily state'
Mood-related	`Gefühl erweckend` 'emotion-stimulus', `Gefühl empfindend` 'emotion-experienced'
Spirit-related	`Intelligenz` 'intelligence/attention', `Wissen` 'knowledge/experience'
Behaviour-related	`Charakter` 'behaviour/character', `Tier` 'animal specific', `Geschicklichkeit` 'skill', `Beziehung` 'relations', `Sympathie` 'sympathy', `Neigung` 'inclination'
Social-related	`Schicht` 'stratum', `Institution` 'institution', `Religion` 'religion,' `Land` 'state,' `Region` 'region,' `Aufgabe` 'role'
Quantity-related	`Zahl` 'number,' `Menge` 'quantity,' `Kosten` 'costs,' `Ertrag` 'return'
Relational	`Geltung` 'validity,' `Gewissheit` 'certainty,' `Bedarf` 'requirements,' `Wirksamkeit` 'effectiveness,' `Aufwand` 'difficulty/energy requirement',' `Funktionalität` 'functioning,' `Sicherheit` 'security,' `Geordnetheit` 'order,' `Verknüpfung` 'linking,' `Übereinstimmung` 'correspondence,' `Genauigkeit` 'accuracy,' `Vollständigkeit` 'completeness,' `Ursache` 'cause,' `Bezug` 'reference', `Zuträglichkeit` 'tolerability'
General	`Steigerung` 'intensity', `Bewertung` 'evaluation', `Norm` 'norm'
Pertainyms	'derived from'
Privative	defined by negation

Table 6.1: Semantic classification of adjectives in GermaNet: it consists of 16 semantic classes and 80 artificial concepts that are direct hyponyms of the semantic class nodes.

2. **Disambiguate the adjective's sense** based on the senses in GermaNet. Each adjective is provided with all its sense IDs from GermaNet and the definitions of each sense.

3. **Disambiguate polysemous nouns** based on the sense IDs in GermaNet and specify the noun's semantic class. The monosemous nouns were assigned the corresponding sense IDs and semantic classes automatically.

Additionally, the annotators were instructed to mark idiomatic uses of the phrases as well as free phrases, proper names, metaphors, etc. As the original GerCo dataset had been annotated without context, it is possible that there are ambiguous phrases that are not collocations in the given context sentence.

6.1.2 Collocations: Agreement and adjudication

Altogether 1,994 adjective-noun pairs with the status "collocation" in the GerCo dataset were manually annotated. 1,259 phrases got full agreement which makes up for ≈ 63% of the data, Cohen's kappa (Cohen, 1960): κ=0.59, and Gwet's AC1 Gwet (2008) discussed in section 5.3.3 of 0.62. Table 6.2 shows raw agreement as well as Kappa and Gwets's AC1 values for each adjective and the number of attributes available for each adjective. There are always at least two attributes available because the "unknown" label was suggested as an additional option for all the adjectives.

Polysemy of the adjectives does not seem to influence the agreement much. For instance, for the adjective *stark* 'strong', six attributes are listed, and the agreement is 93%. For the adjective *frischgebacken* 'freshly baked', on the other hand, there is zero agreement, although there are only two attributes to choose from. Thus, agreement is rather a matter of interpretation of the available labels. In the case of *frischgebacken*, its collocational meaning is the same in all the phrases – "recent". One annotator labeled that meaning with the attribute **`time`**, while the other one used the label **`unknown`** in order to distinguish this collocational meaning from the literal one. However, it is entirely possible that both collocational and literal meanings of an adjective select the same attribute. In the case of *freshly baked*, it chooses the temporal dimension of, say, both *bread* and *parents*.

It has been decided to perform the adjudication in the same manner as for the binary annotation of GerCo (section 5.3.2). More specifically, instead of dismissing all the disagreement cases which is also a common practice (Hartung, 2015), to discuss them until an agreement is reached.

Adjudication and adjustment of labels

Adding new labels. During the adjudication, a few changes were made in the attributes inventory based on the disagreement cases. First of all, two new labels have been added as they had been repeatedly suggested by both

adj	attr	AC1	Kappa	IAA%
herrlich 'wonderful'	2	1.00	1.00	100
hölzern 'wooden'	3	1.00	1.00	100
schlau 'smart'	2	1.00	1.00	100
windig 'windy'	3	0.960030746	0	96
stark 'strong'	6	0.920601858	0.861573955	93
prall 'full'	3	0.908289242	0.763636364	92
grob 'coarse'	5	0.907032339	0.846491228	93
frech 'bold'	2	0.881081081	0.620689655	91
hell 'bright'	6	0.877675841	0.818181818	90
stürmisch 'stormy'	3	0.868677228	0.840593142	91
leicht 'light'	7	0.854710128	0.734863702	87
mild 'mild'	4	0.832052312	0.800897593	87
knapp 'scarce'	6	0.802956758	0.77329589	84
dunkel 'dark'	6	0.79666246	0.777607362	83
karg 'sparse'	5	0.795454545	0.682132964	82
alt 'old'	4	0.792797237	0.669291339	83
starr 'ridig'	3	0.733015495	0.41130092	78
arm 'poor'	3	0.713527851	0.693181818	85
tief 'deep'	4	0.703018626	0.483236477	75
sicher 'safe'	4	0.701765412	0.372746936	74
schwarz 'black'	10	0.697160883	0.630769231	73
historisch 'historical'	4	0.575311151	0.554371002	71
bitter 'bitter'	5	0.559863421	0.480228928	64
zäh 'viscous'	4	0.548337632	0.352517986	63
scharf 'sharp'	12	0.545926072	0.488234424	58
reich 'rich'	6	0.498039216	0.208480565	56
rau 'rough'	7	0.485311921	0.422680412	55
blind 'blind'	5	0.478814786	0.115044248	56
rund 'round'	5	0.453687609	0.484052533	56
dick 'thick'	6	0.452042861	0.475396163	54
frei 'free'	9	0.42464323	0.408798283	49
dumm 'stupid'	4	0.411764706	0.393939394	60
zart 'tender'	7	0.39827105	0.377990431	48
steil 'steep'	3	0.373913043	0	50
süß 'sweet'	6	0.349881797	0.293739968	45
tot 'dead'	4	0.33640553	0.368421053	50
spät 'late'	2	0.316573102	0	55
wild 'wild'	7	0.264671223	0.205163043	35
blank 'shiny'	5	0.239994136	0.221621622	41
offen 'open'	6	0.23965492	0.223735772	35
sanft 'soft'	5	0.165001427	0.161650903	31
hart 'hard'	10	0.124929339	0.09023882	23
deftig 'savoury'	4	-0.061662198	0	39
mächtig 'powerful'	4	-0.093978616	-0.025316456	11
frischgebacken 'freshly baked'	2	-1	0	0

TABLE 6.2: Annotation of collocations with attributes: the number of available attributes (attr), Gwet's AC1 value (AC1), Cohen's Kappa value (Kappa) and raw inter-annotator agreement for each adjective in % (IAA%).

annotators, namely the labels **manner** (Ger. *Art und Weise*) and **description** (Ger. *Beschreibung*).

> **manner** – the way in which an action is carried out/performed, e.g. *graceful movement*.
>
> **description** – the way in which an concrete or abstract entity is described, e.g. lively discussion.

The introduction of the two attributes is also prompted by the findings of the case studies in chapter 4. The labels are borrowed from the inventory of Frame Elements (section 4.4.1).

Another new label is **naturalness** (Ger. *Naturbelassenheit*). This attribute had to be introduce for examples such as *wilde Küste* 'wild coast', *wilde Natur* 'wild nature', *wilde Rose* 'wild rose'. The adjective *wild* in the sense 'not tamed, free-living, freely occurring in nature, natural, not changed by humans' is modeled in GermaNet under the artificial concept `tierspezifisch` 'animal-related' which, on the one hand, cannot really be seen as attribute. Secondly, it cannot be applied to the examples listed above that include noun from semantics fields other than ANIMAL.

The following examples caused a very long discussion as to what attribute to assign to them:

- *freie Natur* 'outdoor nature, the outdoors'. The adjective does not refer to any particular aspect of the noun *Natur*. Nor is it a **type** of nature. The native speakers rather judge the adjective here as semantically empty because it only emphasizes the meaning of the noun.
- *freie Wildbahn* is almost identical to the previous expression. It is usually used as a prepositional phrase *in freier Wildbahn* 'in the wild'. The noun is never used without the modifier, which makes this phrase highly lexicalized.
- *schwarze Nacht* 'black/dark night'. The adjective emphasizes the already existing aspect of the noun's meaning: being dark is the defining aspect of the noun *night*.

All the examples listed above have in common that the adjective's semantics is (almost) empty. This observation is reminiscent of Mel'čuk (2015)'s Lexical Function **Epit** discussed in sections 4.2.1 and 4.6. Therefore, the label **epithet** has been added to the annotation scheme to accommodate these phrases.

Finally, label **rigidity** (Ger. *Rigorosität*) was introduced for examples such as *starre Arbeitszeiten* 'fixed working hours' and *scharfe Regeln* 'strict rules'. Both for *starr* and *scharf*, the meaning 'unyielding, incapable of change, strict' is modeled in GermaNet under the artificial concept `erziehungsspezifisch` 'upbringing-related' which the annotators did not find fitting the examples in the dataset.

Splitting labels. The label suggested by GermaNet for adjectives such as *grob* 'coarse', *dick* 'thick' and *hölzern* 'wooden', is *Beschaffenheit* which is a very abstract notion that can be translated as 'quality, condition, state, nature, consistency, composition, texture'. At the same time, GermaNet also has an artificial concept *Konsistenz* 'consistency' at the same level as *Beschaffenheit*. However, it rather seems to be a hyponym of *Beschaffenheit* than its co-hyponym. This intuition is backed up by the modeling of these attribute nouns in GermaNet illustrated in Figure 6.1 where *Konsistenz* is indeed a hyponym of *Beschaffenheit*. Based on this insight, the annotators came to an agreement to define *Beschaffenheit* as "*die Struktur, Textur einer Sache*", in other words, simply **`texture`** in examples such as *grober Salz* 'coarse salt'; **`consistency`** remains a separate attribute on the same level with `texture` in examples such as *zähe Brei* 'viscous porridge'. Finally, in examples such as *hölzerne Tür* 'wooden door', the adjective expresses what the noun head is *made of*, a material, so an additional attribute is introduced, namely, **`material`**.

Merging labels. As presented in Table 6.1, there are six social-related artificial concepts in GermaNet. Two of those, *Schicht* 'stratum' and *Institution* 'instituition', were identified for the adjectives in the GerCo dataset. There was also a third option because according to the modeling in GermaNet, some adjectives belong directly to the upper node (semantic class) SOCIAL-RELATED without an artificial concept in between. An example of such a case is the adjective *schwarz* 'black' in the sense of 'having a dark skin tone'. This leads to an uneven hierarchy in the annotation inventory: a category and its subcategories are present in the same inventory which leads to the following problem. For adjectives that do have an artificial concept in their modeling, for instance, *reich* 'rich' (artificial concept `stratum`), it would be perfectly acceptable to be annotated with a more general label `social-related`, and thus, there are two equally acceptable attributes for one item. There are two strategies for solving this.[4] The first one is to only use the fine-grained artificial concepts, but then it would be necessary to introduce a label that covers those adjectives that do not have an artificial concept in their hierarchy. The second one is to merge all the fine-grained artificial concepts into one general label. The annotators agreed to the latter because most of the examples they annotated in the GerCo dataset already belonged to a more general attribute **`social`**. This solution is in no way universal, and different data might require a different approach.

Two labels available in the inventory caused an extensive discussion:

> **`emotion-experienced`** – a feeling or an emotion *experienced* by a person. E.g.: *happy child*, *sad people*;
>
> **`emotion-stimulus`** – a feeling or an emotion *caused* by external factors. E.g.: *sad movie*, *pleasant news*

[4] A third option would be to allow multiple labels for one item.

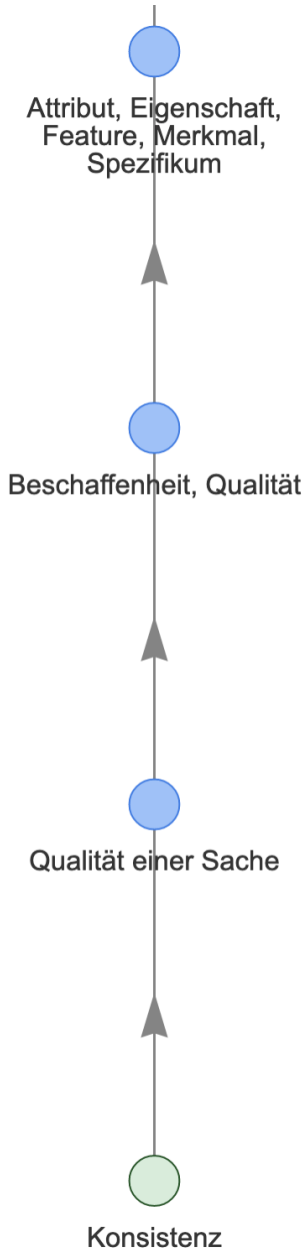


FIGURE 6.1: Screenshot from the online application GermaNet Rover (Hinrichs et al., 2020): the hierarchy for the noun *Konsistenz* which is an indirect hyponym of the noun *Beschaffenheit*.

Examples given in the guidelines are quite straightforward: *sad people* is interpreted as people who themselves experience the emotion of sadness. In such phrases, the nominal head expresses the experiencer of the emotion expressed by the modifier. In the case of *a sad movie*, the external factor movie is the stimulus for the emotion of sadness experienced by someone. The nominal head is the trigger of the emotion expressed by the adjective. This distinction suggested in GermaNet is completely justified. Warren (2003) suggests that there are implicit *links* between adjectival modifiers and their heads. For the adjective *sad*, there are three types of links: *experiencing* (34a), *causing* (34b), and *manifesting* (34c). These links correspond to the two artificial concepts in GermaNet: experiencing = `emotion`, causing = `stimulus`.

(34) a. *sad girl*: **experiencing** sadness
b. *sad song*: **causing** sadness
c. *sad eyes*: **manifesting** sadness

However, not all phrases in the dataset are as straighforward as the ones listed above. The collocations that caused confusion are with the adjective *bitter* 'bitter', for example *bitterer Fazit* 'bitter conclusion' and *bitterer Abend* 'bitter evening' as shown in table 6.3.

	bitter conclusion	**bitter evening**
Annotator 1	`emotion-stimulus`	`emotion-experienced`
Annotator 2	`emotion-experienced`	`emotion-stimulus`

Table 6.3: Disagreement example: annotation of two collocations with the adjective *bitter*.

It is unclear which of the two labels is suitable for these phrases. In *bitter conclusion*, conclusion does not cause bitterness, but rather the person who makes a conclusion experiences bitterness. In the second example, *bitter evening* is a evening during which someone experiences bitterness. It has been decided to merge the two labels into a more general label **`emotion`** and not to make a finer distinction between the underlying links.

Heuristic. The approach taken in this thesis aims at maintaining an even level of hierarchy of attributes. Thus, some sort of heuristic is needed for making a decision about the granularity of the labels. A concept frame structure in Barsalou (1992)'s sense can serve as such and has proven to be useful during the adjudication process. Consider figure 6.2 that presents a sketch for a concept frame for the noun *evening*. Basically, the annotator's task is to represent the meaning of the head noun in terms of attributes and check the levels of the nodes in the hierarchy. For instance, figure 6.2 depicts a possible frame structure for the noun *evening* where the attribute **`feeling`** is at the same level as the three other ones.
Similarly, for the noun *conclusion*, a frame structure presented in figure 6.3 is suggested. These frames are in no way final and can be extended depending on the collocates of the nouns. After the annotation of the dataset is finished, it will be possible to create more reliable templates for such frames

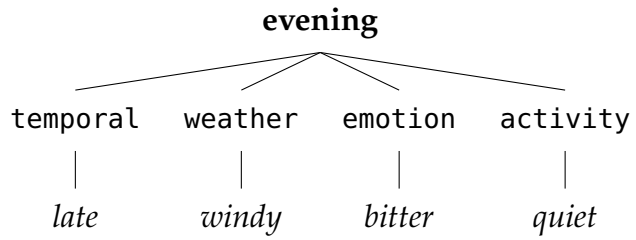


FIGURE 6.2: A sketch of the frame structure for the noun **evening** with attributes and their possible values as adjectives.

of nouns. As each noun is annotated with its sense ID and with its semantic class, generalized templates can be created for the semantic classes present in the dataset. For instance, the semantic class of the noun *Abend* 'evening' in GermaNet is *Zeit* 'time'. There are also other nouns from the same semantic class in the dataset: e.g. *Jahr* 'year', *Nacht* 'night', *Winter* 'winter', *Wochenende* 'weekend'. Based on the annotation of these nouns and their collocates, a template for the nouns from the semantic class *Zeit* 'time' will be created. The templates will be discussed in section 6.3.

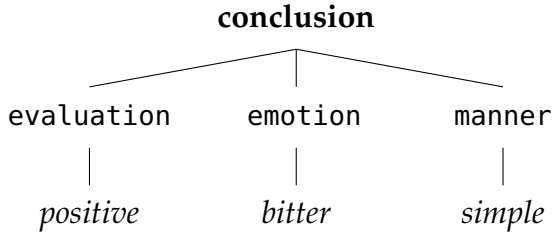


FIGURE 6.3: A sketch of the frame structure for the noun **conclusion** with attributes and their possible values as adjectives.

"Unknown" cases. The label **unknown** was introduced for the cases when GermaNet does not have a suitable label in the modeling of a given adjective. For example, the adjective *deftig* lit. 'hearty, substantial' combines in its literal meaning with nouns from the semantic class FOOD. When used collocationally, as in *deftige Strafe* 'stiff punishment', the adjective functions as an intensifier. However, the intensity meaning is not (yet) modeled for the adjective *deftig* in GermaNet. GermaNet includes only two senses for the adjective *deftig* as presented in example (35).

(35) a. Synset1: [*deftig*, *derb*, *rustikal*] in the sense 'nutritious, substantial (food)'
b. Synset2: [*deftig*, *herzhaft*] in the sense 'rough and direct; rustic'

The annotators were asked to mark all such cases as unknown.

Another example is the adjective *hell* lit. 'bright'. GermaNet models 4 senses: 'bright/luminous', 'bright (describing sound)', 'lucid (sane)', 'smart'. There are, however, two more senses: (1) 'intense, of high degree' as in *helle Empörung* 'sheer outrage' and *helle Freude* 'sheer joy'; (2) positive evaluation as in *helle Zukunft* 'bright future'. Thus, for *helle Freude* and *helle Empörung*, the label `unknown` was changed to **`intensity`**; for *helle Zukunft*, it was changed to **`evaluation`**.

All the phrases annotated as **`unknown`** by at least one annotator, 676 in total, were then given to the annotators again. In most cases, the labels already available in the inventory of attributes were used.

The annotation described in this section has been performed for collocations. The second half of the GerCo dataset comprised of free phrases has undergone a different annotation procedure that was partially automated. It will be described in the next subsection.

6.1.3 Annotation of free phrases

The constituents of free phrases, in contrast to collocations, retain their literal meaning (5.4.3). For each adjective in the GerCo dataset, their literal meaning had been defined before the annotation (see the discussion in section 5.2.1 and the definitions in Appendix B). For example, the adjective *rund* 'round' is defined as "spherical, circular, ring-shaped, without corners and edges".[5] It retains this meaning in all the phrases marked as 'free' in the GerCo dataset: e.g. *runder Fleck* 'round spot', *runder Ball* 'round ball', *runder Knopf* 'round button', etc. The attribute corresponding to this sense of the adjective is **`shape`**. Consider a different adjective: *blind* 'blind' – "without sight, vision", as in *blindes Kind* 'blind child', *blinder Hund* 'blind dog'. The corresponding attribute is **`health`**, and thus all the free phrases with the adjective *blind* can be annotated with this attribute. In this manner, each adjective that has instances of free phrases in the GerCo dataset has been assigned a label from the annotation scheme. All the free phrases were then labeled automatically.

Two relational adjectives, *barock* 'baroque' and *steinig* 'stone', have been completely excluded from the dataset. The decision has been motivated by the fact that the relations between such modifiers and their head nouns are different from those of descriptive adjectives (see section 2.1) and should rather be accommodated by an inventory of relations in noun-noun compounds because relational adjectives denote objects and not properties. Both adjectives are modeled in GermaNet as belonging to the semantic class of PERTAINYMS. The artificial concepts in their modeling are pointers to the semantic classes of the nouns the adjectives are derived from: EVENT for *barock* and NATURAL OBJECT for *steinig*.

The third relational adjective in the dataset, the adjective *hölzern* 'wooden', is different. Although it is denominal as well, it belongs to the adjectival semantic class SUBSTANCE and the artificial concept in its

[5] Translated from German.

modeling is `texture`. The difference is most likely motivated by the fact that the adjective *hölzern* denotes the same relation in all free phrases independently of the noun it combines with, namely, the relation *made-of* as in *hölzerne Decke* 'wooden ceiling', *hölzener Kasten* 'wooden box', *hölzerne Puppe* 'wooden doll'. In contrast, the adjective *barock* can be in various relations with its head nouns, e.g. denoting `style` in *barocke Fassade* 'baroque facade' = facade in baroque *style*, `typical for` in *barocke Lebensfreude* 'baroque zest for life' = zest for life *typical for* baroque epoch.

In order to safeguard against errors in ambiguous cases that are context dependent, all the free phrases were checked manually. Only a few cases of collocations, proper names, and idioms were identified. Manual verification shows that the automated approach to free phrases is highly reliable. The total of 1,425 free phrases were annotated by the same two annotators who worked with the collocations. Additionally to attributes, each head noun was annotated with its semantic class, as it had been done for the nouns in collocations.

6.1.4 Semantic classes of nouns

The annotators' task included assigning a semantic class to each head noun based on the semantic classes available in GermaNet. The agreement on this task was 92% and all the disagreement cases were successfully resolved. The tendency for disagreement cases is similar to the one for the adjectives: apart from the level of polysemy, a major source of confusion is the abstractness of the nouns' meanings. Semantic classes that were most frequently confused are EVENT, COMMUNICATION, and COGNITION. Apart from that, the semantic class ARTIFACT is often confused with various classes when annotating ambiguos nouns such as *Kunst* 'art', *Roman* 'novel', *Satire* 'satire'.[6] Consider the noun *Kunst* 'art': it may either denote an ability or a skill in which case its semantic class is COGNITION, or product of an activity - ARTIFACT. The semantic classes will be later used for creating general templates for grouping adjectival collocates in section 6.3.

6.1.5 Summary

The initial annotation inventory was solely based on the semantic classification of adjectives in GermaNet. In the process of adjudication, the disagreement cases indicated some disadvantages of the initial scheme and, as a result, it was slightly changed: new labels were added, too fine-grained labels were merged into more general categories. The overview of the resulting inventory consisting of 49 labels is presented in table 6.4.

[6]One example stands out, namely *hölzernes Pferd* 'wooden horse', where the meaning of the nominal head is shifted from *horse as an animal* to *an object that looks like a horse* – see discussion of such phrases in Kamp and Partee (1995).

attribute	example free	example collocation
age	*alte Dame* 'old lady'	-
appearance	*sanftes Gesicht* 'gentle face'	*blanker Oberkörper* 'bare torso'
brightness	*dunkles Zimmer* 'dark room'	-
certainty	-	*windiges Versprechen* 'windy promise'
character	*frecher Bursche* 'cheeky fellow'	*scharfer Hund* 'fierce dog'
climate	textitstürmischer Herbst 'stormy autumn'	*raue Region* 'rough region'
color	*schwarzer Rock* 'black skirt'	-
completeness	-	*offene Rechnung* 'open account'
consistency	*zähe Masse* 'thick mass'	*dicker Nebel* 'thick fog'
constitution	*starker Mann* 'strong man'	*zäher Bursche* 'tough guy'
cost	*teueres Auto* 'expensive car'	*freier Eintritt* 'free entrance'
description	*freche Antwort* 'impertinent answer'	*raue Poesie* 'rough poetry'
dimension	*tiefes Loch* 'deep hole'	*knappes T-Shirt* 'tight t-shirt'
duration	-	*alter Freund* 'old friend'
effort	-	*leichter Sieg* 'easy victory'
emotion	-	*bitteres Lachen* 'bitter smile'
epithet	-	*freie Wildbahn* '(in) the wild'
evaluation	*herrlicher Tag* 'wonderful day'	*süßes Kind* 'sweet child'
function	-	*scharfe Munition* 'live ammunition'
health	*blinder Pianist* 'blind piano player'	-
importance	-	*historische Entscheidung* 'historic decision'
intelligence	*schlauer Junge* 'smart boy'	-
intensity	*mildes Sonne* 'mild sunlight'	*grober Fehler* 'grave mistake'
manner	-	*wilder Tanz* 'wild dance'
material	*hölzerne Decke* 'wooden ceiling'	-
movement	*starrer Körper* 'rigid body'	*wildes Wasser* 'wild water'
naturalness	*wildes Tier* 'wild animal'	*wilde Küste* 'wild coast'

Continued on the next page

Table 6.4: Continued from previous page

attribute	example free	example collocation
power	*mächtige Familie* 'powerful family'	*starke Gruppe* 'powerful group'
precision	-	*grobe Abschätzung* 'rough estimation'
quantity	*karge Rente* 'meager pension'	*reiche Ernte* 'rich harvest'
rigidity	-	*starre Vorgaben* 'rigid guidelines'
security	*sichere Distanz* 'safe distance'	-
shape	*runder Mond* 'round moon'	-
smell	-	*süßer Geruch* 'sweet smell'
social	*reiche Verwandten* 'rich relatives'	*offener Brief* 'public letter'
sound	-	*tiefe Stimme* 'deep voice'
space	-	*offenes Feld* 'open field'
speed	-	*stürmische Industrialisierung* 'rapid industrialisation'
stability	-	*starke Währung* 'strong currency'
state	*totes Tier* 'dead animal'	-
surface	*blankes Aluminium* 'shiny aluminium'	*blinder Spiegel* 'cloudy mirror'
taste	*süßes Brötchen* 'sweet bun'	*milder Käse* 'mild cheese'
temperature	-	*milder Abend* 'mild evening'
texture	*grober Sand* 'coarse sand'	*dicker Pullover* 'thick pullover'
time	*später Abend* 'late evening'	*alter Name* 'old name'
tolerability	-	*milde Seife* 'mild soap'
type		*schwarzer Tee* 'black tea'
validity	-	*blinder Passagier* 'stowaway'
weight	*leichtes Gepäck* 'lightweight luggage'	-

TABLE 6.4: Inventory of attributes after adjudication with examples from the dataset.

6.2 Gold standard dataset GerCoAt

6.2.1 Overview

The resulting gold standard dataset referred to by the acronym GerCoAt (German Collocations with Attributes)[7] contains 3,305 adjective-noun phrases[8] that are either collocations or free phrases. The dataset provides two kinds of annotations: (1) relations between the modifier and the head – 49 relations in total, and (2) semantic classes of the nouns – 23 classes in total. Table 6.6 presents an overview of the gold standard GerCoAt. More than half of the dataset is represented by collocations. The number of semantic classes of head nouns is equal for both collocations and free phrases. All the semantic classes listed in GermaNet are represented in the dataset which indicates that both collocations and free phrases are semantically highly diverse. However, the number of attributes is much higher for collocations (40 attributes) than for the free phrases (29 attributes). The fact that almost all the attributes from the inventory are present in collocations (41 out of the total of 49) shows that the relations between the constituents of collocations are very diverse.

Each adjective in the dataset combines with nouns from various semantic classes. Table 6.5 illustrates the variety of nominal semantic classes for each attribute identified for the adjective *bitter* 'bitter'. In order to find this information for other adjectives, consult the full dataset available at https://doi.org/10.57754/FDAT.76krc-egt63.

	collocations	free phrases
phrases	1,870	1,435
unique nouns	1,358	1,073
unique adjectives	45	45
attributes	41	29
sem classes nouns	23	23

TABLE 6.6: Overview of the GerCoAt dataset.

Concreteness. Mean concreteness scores of the nouns for each attribute are presented in figure 6.4. The scores are taken, as in the previous chapter (see section 5.4.7), from the collection of affective norms by Köper and Schulte im Walde (2016). The attributes with the most concrete nouns are **surface** (6.84), **color** (6.8), **material** (6.75), and **appearance** (6.73). These attributes are typical for nouns from the semantic classes ANIMAL, ARTIFACT, NATOBJECT, SUBSTANCE, etc. The most abstract attributes are **speed** (3.31), **epithet** (3.33), **rigidity** (3.5), and **intensity** (3.58) characterizing nouns of COGNITION, COMMUNICATION, EVENTS, FEELING, etc.

[7] The data can be downloaded here: https://doi.org/10.57754/FDAT.76krc-egt63

[8] All the idioms, proper names, named entities, and technical terms have been excluded from the final dataset.

adj	noun	attribute	NN-semClass	status
bitter	*Erinnerung* 'memory'	emotion	cognition	collocation
bitter	*Kommentar* 'comment'	emotion	communication	collocation
bitter	*Erfahrung* 'experience'	emotion	event	collocation
bitter	*Gefühl* 'feeling'	emotion	feeling	collocation
bitter	*Realität* 'reality'	emotion	location	collocation
bitter	*Abend* 'evening'	emotion	time	collocation
bitter	*Notwendigkeit* 'necessity'	intensity	attribute	collocation
bitter	*Träne* 'tear'	intensity	body	collocation
bitter	*Kritik* 'criticism'	intensity	communication	collocation
bitter	*Niederlage* 'defeat'	intensity	event	collocation
bitter	*Enttäuschung* 'disappointment'	intensity	feeling	collocation
bitter	*Kälte* 'cold'	intensity	natPhenomenon	collocation
bitter	*Feind* 'enemy'	intensity	person	collocation
bitter	*Armut* 'poverty'	intensity	possession	collocation
bitter	*Feindschaft* 'hostility'	intensity	relation	collocation
bitter	*Ende* 'end'	intensity	time	collocation
bitter	*Geschmack* 'taste'	taste	attribute	free
bitter	*Mandel* 'almond'	taste	food	free
bitter	*Schokolade* 'chocolate'	type	food	collocation

TABLE 6.5: Exmples of nouns from each semantic class per attribute encountered with the adjective *bitter* 'bitter' in the GeCoAt dataset.

Polysemy. As discussed in Chapter 5, one of the main characteristics of the GerCo dataset is the high level of polysemy of the adjectives. This feature is reflected in the variety of attributes that the adjectives in the dataset can select. Figure 6.5 gives an overview of the number of attributes for each adjective in the GerCoAt dataset.
There are four adjectives in the dataset that select only one attribute: *frischgebacken* 'freshly baked' and *spät* 'late' – the attribute **`time`**; *teuer* 'expensive' – **`cost`**; *herrlich* 'wonderful' – **`evaluation`**. The rest of the adjectives have at least two options when combined with nouns with the adjective *hart* 'hard' being the most polysemous as it has 12 attributes to choose from. Appendix I provides a detailed overview of adjectival senses found in the dataset with examples, corresponding artificial concepts in GermaNet, final attributes after assigned after the adjudication, and collocational status of each sense.

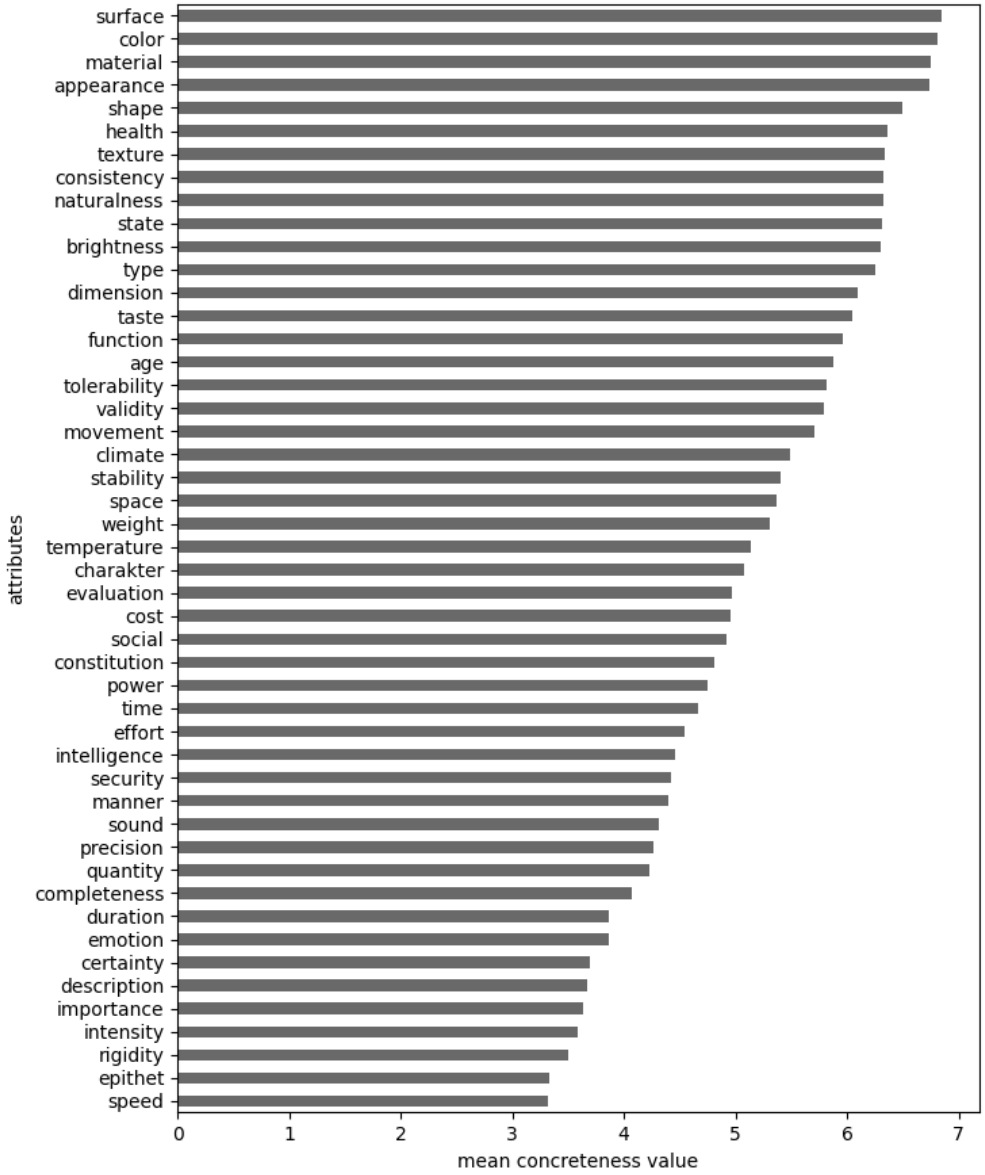


FIGURE 6.4: Mean concreteness scores for the nominal bases for each attribute in GerCoAt.

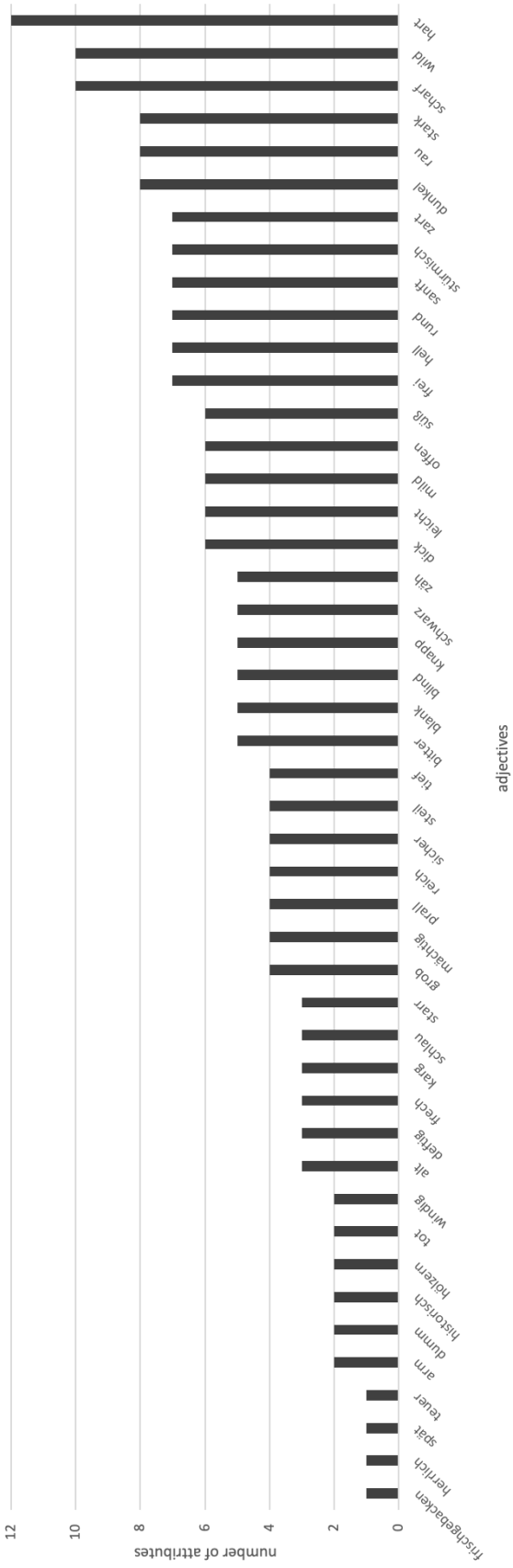


FIGURE 6.5: The number of attributes per adjective in the GerCoAt dataset.

Lexical diversity of attributes. In the Meaning-Text Theory (see section 4.2.1), Lexical Functions are divided into the two groups **standard** and **non-standard** based on two criteria: broadness of the domain and broadness of range of a function. The two criteria will be employed for characterizing the attributes as well. No strict grouping into standard and non-standard attributes is intended here because the gold standard is not large enough to make such assumptions. It is, however, large enough to outline the tendencies in the lexical diversity for the attributes, both in terms of adjectives that select the attributes (=range) and the semantic classes of nominal bases that are characterized by the attributes (=domain).

The domain and the range are measured here in terms of entropy instead of raw counts. The reason for that is the imbalance of both the number of adjectives for each attribute and the number of semantic classes. For instance, when relying on raw counts, the attributes **`evaluation`** and **`sound`** are equally varied in terms of range (table 6.7): both attributes are represented by six different adjectives in the dataset. However, the inspection of the data shows that 62 out of the total of 96 phrases with **`evaluation`** are represented by the adjective *herrlich* 'wonderful', and 19 by the adjective *süß* 'sweet', the rest of the adjectives only have a few instances. In contrast, the distribution of adjectives for the attribute **`sound`** is much more balanced. Similar examples can be found in the distribution of semantic classes. For example, **`taste`** and **`security`** both have 10 semantic classes, but the former is highly imbalanced with 50 out of the total of 68 instances belonging to the class FOOD.

In order to take into account such imbalances, the measure of entropy (6.1) was adopted here from Kilgarriff (2006) where it was proposed to measure *collocationality* of words.

$$H(x) = -\sum p(x) \times log_2(p(x))) \tag{6.1}$$

Applied to range, entropy *H(x)* indicates how much information is contained in the distribution of adjectives for a given attribute. For the domain, the distribution of semantic classes of the nominal bases is taken into consideration. The lower the entropy, the less information there is in the distribution. For the above discussed examples: the entropy of the range for **`evaluation`** (1,5) is less than for **`sound`** (2,4), which means that the latter is semantically more varied in terms of its range. Similarly, the domain for **`security`** (H=2,6) is more diverse than for **`taste`** (H=1,6).

Table 6.7 displays the range of the attributes in the GerCoAt dataset: entropy and the number of adjectives for each attribute. The most lexically diverse attributes are **`description`** and **`intensity`**. In contrast, the attributes **`age`**, **`color`** and a few others are represented only by one adjective in the GerCoAt: *alt* 'old' and *schwarz* 'black' respectively. Bear in mind that during the selection of adjectives, semantic classes of adjectives were balanced: per each semantic class, three adjectives were selected to make sure that the dataset is heterogeneous. All adjectives are polysemous, and the fact that the final dataset is unbalanced indicates that the meanings tend to shift

attribute	entropy	n adjectives
age	0	1
color	0	1
duration	0	1
health	0	1
importance	0	1
material	0	1
naturalness	0	1
security	0	1
temperature	0	1
weight	0	1
cost	0.082836553	2
speed	0.543564443	2
shape	0.577653903	2
state	0.636211337	3
movement	0.764114055	3
power	0.80609147	2
intelligence	0.922259647	2
consistency	0.954434003	2
validity	0.985228136	2
space	0.996791632	2
brightness	0.99867467	2
smell	1	2
stability	1	2
precision	1.26273173	3
surface	1.268265045	3
completeness	1.295737801	3
rigidity	1.324531805	3
texture	1.493162865	4
tolerability	1.5	3
evaluation	1.527874642	6
taste	1.56906161	4
constitution	1.632105718	4
quantity	1.789236952	5
epithet	1.871602261	6
effort	1.886308282	4
function	1.891061112	4
climate	1.910036815	4
time	1.93657298	7
emotion	2.031186251	7
certainty	2.126271749	6
social	2.132684057	5
appearance	2.271233797	7
dimension	2.328519303	7
manner	2.363752358	8
sound	2.371895669	6
type	2.59297866	8
character	3.122459275	11
intensity	3.994002812	24
description	4.163897057	30

TABLE 6.7: The range of attributes in the GerCoAt dataset: entropy and the number of unique adjectives per attribute.

attribute	entropy	n semclasses
smell	0	1
manner	0.205592508	2
material	0.336073374	5
health	0.453716339	2
appearance	0.824980257	3
space	0.905587262	3
speed	1.198192411	3
stability	1.251629167	3
temperature	1.325011211	3
weight	1.459147917	3
character	1.460527218	4
tolerability	1.5	3
taste	1.590367989	10
age	1.694051166	6
climate	1.750413285	8
function	1.879964949	5
effort	1.969811065	5
color	1.980089233	8
surface	2.081556038	7
duration	2.084962501	5
power	2.106597425	11
intelligence	2.106734298	7
sound	2.155221529	5
naturalness	2.176698209	7
constitution	2.28166966	7
importance	2.292071752	6
shape	2.32260112	12
consistency	2.352217001	6
type	2.449311833	8
state	2.491960614	14
validity	2.521640636	6
security	2.607925867	10
dimension	2.657350164	14
completeness	2.661226256	7
epithet	2.682588731	8
movement	2.693020192	9
social	2.711882222	14
rigidity	2.813076647	11
cost	2.847622736	15
brightness	2.987553173	13
emotion	3.003692615	13
time	3.021390994	16
precision	3.037230147	11
texture	3.127152387	11
description	3.189365884	18
quantity	3.221219268	15
intensity	3.335940103	18
certainty	3.34078638	16
evaluation	3.8455914	20

TABLE 6.8: The domain of attributes in the GerCoAt dataset: entropy and the number of semantic classes per attribute.

towards a small set of attributes: **description**, **intensity**, **character**, **type**, **sound**, and **manner**.

Table 6.8 shows the broadness of the domain for each attribute in terms of entropy and the number of semantic classes of the nominal bases. The most diverse attributes are **evaluation**, **certainty**, **intensity**, **quantity**, **description**, whereas **smell**, **manner**, and **material** are more homogeneous. This might correlate with the status of phrases if for a given adjective, the attribute present in free phrases is restricted to fewer semantic classes than the attribute present in collocations. According to table 6.6, there is an equal number of semantic classes in collocations and free phrases: 23 classes. The general entropy of semantic classes both for collocations and free phrases is also almost equal: 3.86 and 3.71, respectively. However, zooming into the semantic class entropy for each adjective shows that such balance does not hold for all the adjectives. Table 6.9 compares the semantic class entropy for each adjective in the dataset between collocations and free phrases. The greatest difference is found in adjectives *stark* 'strong' and *spät* 'late': the entropy is higher in collocations. For much fewer adjectives, the situation is reversed; for instance, the entropy for *schlau* 'smart' and and *herrlich* 'wonderful' is higher in free phrases.

Collocations versus free phrases. Figure 6.6 depicts the distribution of collocations and free phrases across all the attributes in the GerCoAt dataset. The distribution is highly unbalanced. First of all, it is unbalanced in terms of the number of instances per attribute: **intensity** has over 450 instances, **description** and **social** – over 200 phrases, but more than half of the attributes have less then 100 instances. Secondly, there are attributes that are found overwhelmingly either in collocations or in free phrases. The number of attributes equally present in both types of phrases is very small.

The following discussion will start with attributes that are typical for collocations and that were challenging for the annotators, followed by the attributes found mostly in free phrases. The balanced attributes will be discussed last.

6.2.2 Attributes in collocations

Collocations are highly diverse in terms of the relations that hold between their constituents. This section will discuss the most challenging attributes.

Intensity denotes intensification or diminution of the most salient property of the modified noun. In the GerCoAt dataset, this attribute is one of the most diverse attributes in terms of its lexical realization: as shown in table 6.7, 24 adjectives in the dataset can select it. Bear in mind that this attribute can only describe nouns that have a feature that can be intensified. As table 6.8 shows, the domain of this attribute is very broad covering 18 semantic classes out of the total of 23 available in GermaNet. The most frequent semantic class for this attribute is EVENT as in *dickes Lob* 'high praise', *deftige Strafe* 'severe punishment', *grober Unsinn* 'utter nonsense', *starker Einfluss*

	collocations		free phrases		
adjective	entropy	n semclass	entropy	n semclass	difference
stark	3.454242311	15	0	1	3.454242311
spät	3.108028908	13	0.183122068	2	2.92490684
hart	2.676730902	14	0	1	2.676730902
blind	2.76774325	8	0.453716339	2	2.314026911
blank	3.188539673	13	1.105013997	3	2.083525676
bitter	3.041123937	14	1	2	2.041123937
starr	3.21828966	14	1.251629167	3	1.966660492
offen	3.075252072	12	1.148834854	3	1.926417218
scharf	3.145435202	14	1.378783493	3	1.766651709
süß	3.037156818	11	1.347792586	8	1.689364232
leicht	3.141656531	17	1.459147917	3	1.682508614
deftig	1.936278124	5	0.297472249	2	1.638805876
mild	3.288566682	15	1.657742727	4	1.630823956
reich	3.074198347	10	1.46731624	4	1.606882107
alt	3.207142961	12	1.694051166	6	1.513091795
zart	3.670276897	15	2.173557262	7	1.496719635
frei	3.447392523	14	2.022579762	5	1.424812761
dick	3.107257305	11	1.806156463	7	1.301100841
wild	3.091012236	14	1.913890987	6	1.17712125
mächtig	2.772017969	10	1.720190233	8	1.051827736
tot	2.725480557	7	1.707547375	8	1.017933182
sanft	3.248359626	13	2.251629167	5	0.996730459
rund	2.984697003	10	2.133046511	11	0.851650492
zäh	2.755734257	12	1.950212065	4	0.805522193
stürmisch	2.512896066	9	1.792487689	4	0.720408377
schwarz	2.614786979	8	1.980089233	8	0.634697747
sicher	3.182576621	14	2.573568764	10	0.609007857
dunkel	3.28418372	11	2.742674139	11	0.541509581
frischgebacken	0.529360865	2	0	1	0.529360865
tief	3.086474159	14	2.636463246	10	0.450010913
prall	1.78067213	5	1.370950594	3	0.409721535
windig	1.902174814	7	1.625814584	4	0.276360231
rau	2.93375442	13	2.686278124	7	0.247476295
knapp	2.78396923	11	2.68418372	9	0.09978551
dumm	2	5	1.971912754	7	0.028087246
grob	2.381598663	10	2.36711147	6	0.014487194
frech	2.16197818	5	2.229871195	7	-0.067893015
hell	2.855953643	10	3.009336101	12	-0.153382458
karg	1.659458321	7	1.932156076	6	-0.272697755
hölzern	0	1	0.336073374	5	-0.336073374
steil	1.505876256	5	1.866139923	6	-0.360263668
historisch	2.764201431	8	3.381190239	15	-0.616988808
arm	1.162533673	5	1.840540331	5	-0.678006658
herrlich	2.05881389	5	3.157376474	13	-1.098562584
schlau	0.811278124	2	2.195375354	6	-1.384097229
barock	-	-	2.466246459	10	-
teuer	-	-	2.843945069	15	-
steinig	-	-	1.118078209	3	-

TABLE 6.9: Entropy of nominal semantic classes for each adjective in the GerCoAt dataset: for collocations and free phrases sorted by the difference in entropy values (column "difference").

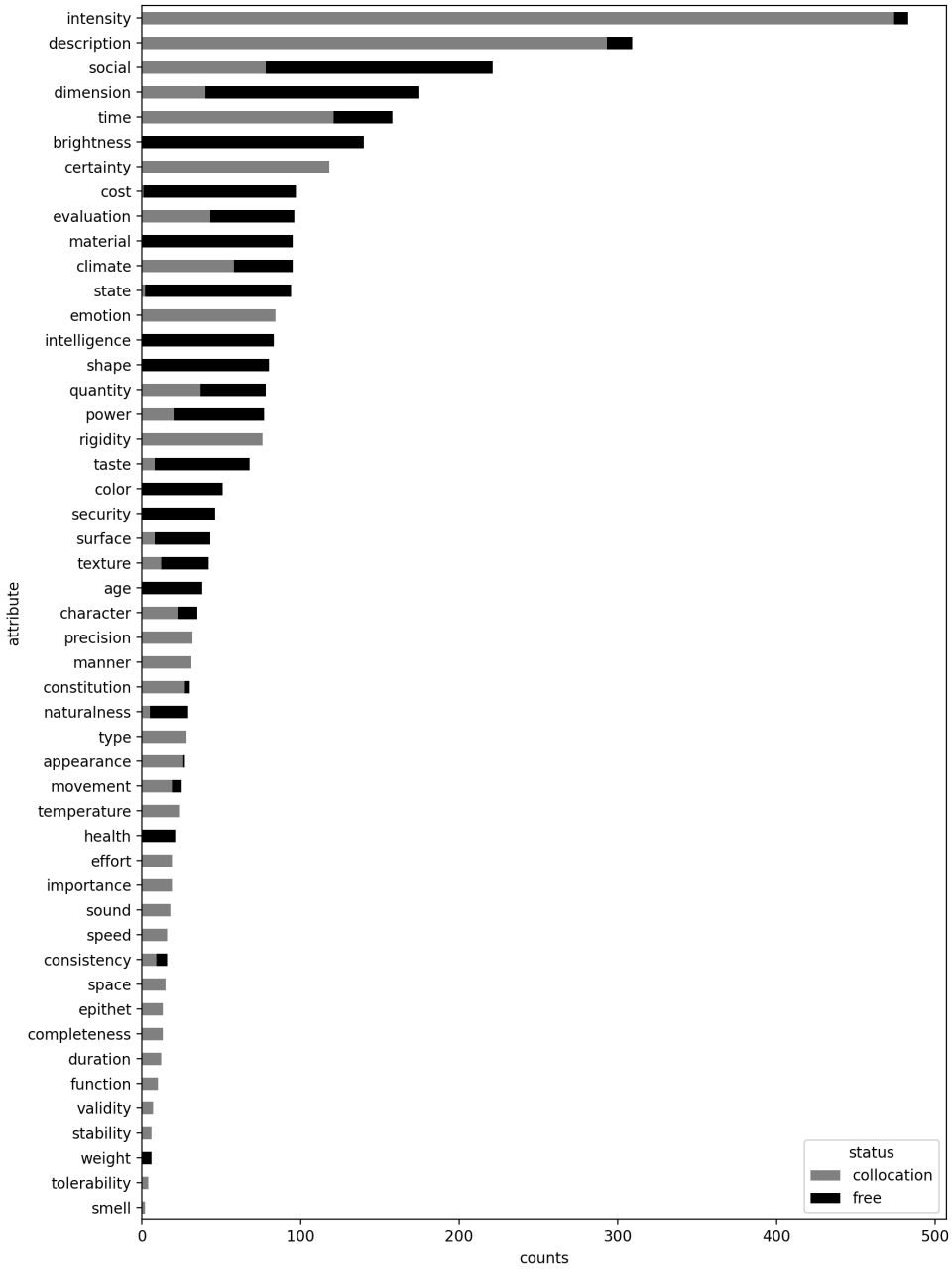


FIGURE 6.6: Distribution of collocations and free phrases in the GerCoAt dataset for each attribute. Sorted by the total number of instances an attribute has in the dataset.

'strong influence', etc. `Intensity` is also a frequent attribute of the nouns of FEELING such as *Angst* 'fear' or *Freude* 'joy'. These semantic classes are obvious candidates for having the attribute `intensity`. There are, however, some examples that are less straightforward, such as PERSON.

One such example has been already discussed in chapter 4: *dicke Freunde* 'close friends'. The reference is made to the relationship – *friendship*, not the friends as human beings. Similar examples are found in GerCoAt: *bitterer Feind* 'bitter enemy', *dicke Kumpels* 'best buddies'. Another kind of reference is to the occupation or rather typical activity of a person: *scharfer Kritiker* 'harsh critic' is someone who criticises harshly. There are also a few TIME nouns that can be characterized in terms of `intensity`: e.g. *tiefer Winter* 'deep winter' and *tiefe Nacht* 'deep night', where the modifier intensifies the temporal aspect of the nominal base.

The meaning of intensification has been discussed in literature on collocations before: Mel'čuk (1996) includes the meaning of intensification in the inventory of Standard Lexical Functions as **`Magn`**. The same idea is advocated in van der Wouden (1997) who notes that most intensifiers in languages are of collocational nature. He illustrates this observation with a list of adverbs that can serve as intensifiers, as well as a few adjectives and idioms. van der Wouden (1997) emphasizes that the reason behind selecting a certain intensifier for particular bases can be explained as there is a pattern based on a common metaphor. For instance, the choice of the adjective in *deep sleep* is motivated by the metaphor ´´sleep is an abyss": to *fall* asleep. It is outside of the scope of this thesis to look for such patterns and explanations of the choice of collocates, but the GerCoAt dataset does provide rich data for investigating this issue.

`Description/ manner`. The two attributes borrowed from FrameNet have been shortly discussed in section 4.4.2: both of them have basically the same meaning. They differ only in terms of the nouns they characterize: **`description`** is an attribute of abstract entities such as *Poesie* 'poetry' and *Schönheit* 'beauty' and **`manner`** describes events such as *Tanz* 'dance' and *Kuss* 'kiss'. The distinction is very subtle and, therefore, the two categories can be merged into one. In terms of domain broadness, the attribute **`description`** is highly varied: it is found in 20 different semantic classes. Its range is also very broad with 30 various adjectives as its values.

Both attributes refer to a quality or characteristic that cannot be reduced to a more specific attribute defined in the annotation guidelines. They may be compared, in a way, to the Lexical Function **`Bon`** from the Meaning-Text Theory (see section 4.2.1) which is defined as subjective evaluation. However, they are semantically more rich and cannot be reduced to a positive or negative subjective evaluation. Consider the following examples: *zarte Poesie* 'soft/tender poetry', *raue Schönheit* 'rugged beauty', *wilder Tanz* 'wild dance', *stürmischer Kuss* 'passionate kiss' (lit. 'stormy kiss'). A *wild dance* does not contain a positive or negative connotation, or at least not without a context. It simply conveys the manner of the dance. Similarly, *soft poetry* is a description of the manner in which a piece of poetry is subjectively

perceived. The same holds for *passionate kiss* and *rugged beauty*.

The two attributes seem very generic. However, an attempt to create more fine-grained attributes for such examples failed: only deadjectival attributes specific to each phrase such as `wildness, softness, ruggedness,` etc. were suggested by the annotators. Thus, no intermediate level of granularity could be found. As stated in (Ruimy et al., 2001) and discussed in section 4.3.1 of this thesis, the semantics of abstract entities is very complex. The `manner/description` dimension of their meaning thus incorporates a large number of very specific characteristics denoted by specific adjectives.

`Certainty` is the dimension of a noun's meaning that expresses how confident the speaker is about the existence of the noun or of some salient aspect of the noun's meaning. First of all, it is represented by the adjective *sicher* 'certain' that basically expresses **epistemic modality**, i.e. "the degree of confidence that the speaker has in the state of affairs expressed by a particular utterance" (Pustejovsky and Batiukova, 2019, p.273) Consider the example *sicherer Tod* 'certain death':

(36) "Eine Lungenentzündung bedeutet den **sicheren Tod** für einen
"a pneumonia means the certain death for
Python", erklärte die Biologin.
a python", explained the biologist.
'Pneumonia means **certain death** for a python," the biologist explained.'

In example (36),[9] the speaker expresses her certainty about the event of death taking place in the future. Consider a different example:[10]

(37) Manche Anleger lassen sich von **windigen Versprechen**
some investors let themselves of windy promises
locken.
lure.
'Some investors are lured by **windy promises**'

The sentence in (37) illustrates the opposite polarity of the attribute: a *windy promise* is unlikely to be fulfilled, thus extending the category **`certainty`** to adjectives that are not considered epistemic modal adjectives, but convey a similar meaning.

`Effort` refers to the energy or work needed to carry out an action, to achieve a goal. Consider the collocation *leichte Aufgabe* 'easy task' (lit. 'light task'): the adjectival modifier selects the attribute directly from the semantic structure of the noun: the **`effort`** to fulfill a task.

[9] In *Die Zeit, 18.03.2016 (online)*, from the corpus Die Zeit of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/zeit>, *retrieved on 15.11.2021.*

[10] In *Die Zeit, 17.04.2017, Nr. 14*, from the corpus Die Zeit of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/zeit>, *retrieved on 15.11.2021.*

`Function` refers to the use or purpose of an entity.[11] Most of the nominal bases in this category belong to the semantic class of ARTIFACTS and all the collocations are highly lexicalized: e.g. *totes Gleis* 'dead track', *scharfe Munition* 'live munition'.

(38) a. *totes Gleis* 'dead track' – TELIC role: attribute `function`
b. *totes Tier* 'dead animal' – FORMAL role: attribute `state`

Artifacts are functional words (Pustejovsky and Batiukova, 2019, p.159), the availability of the TELIC dimension in their semantic structure makes the collocation in (38a) interpretable. Compare it to (38b): the qualia template for the semantic type ANIMAL does not have a telic dimension, thus the modifier selects its FORMAL dimension.

`Validity` refers to the property of being legal. What makes it remarkable is that all the collocations labeled with that attribute are highly lexicalized. Consider the example *blinder Passagier* 'stowaway' (lit. 'blind passenger'): the modifier *blind* has a very specific meaning in this collocation not found in any other combinations. Another adjective that selects this attribute, *wild* 'wild', is quite productive in that sense and combines with a few nouns to express `validity`: *wild* 'wild' + *Siedlung* 'settlement'/*Parken* 'parking'/*Müllkippe* 'waste dump'/*Ehe* 'marriage'. As it has only 5 instances in the GerCoAt dataset, it could be considered a specification of the **`social`** aspect.

`Type`. This category was borrowed from FrameNet and already shortly discussed in chapter 4. In the GerCoAt dataset, it is mostly represented by nouns from the semantic classes ARTIFACT and FOOD, similarly to the examples from the case study in chapter 4. Unlike other labels in the annotation scheme, **`type`** cannot be referred to as *attribute*. Consider the example *schwarzer Tee* 'black tea': the adjective does not select one particular aspect of tea, but targets a number of them: mainly its `processing`, which, however, also influences its `color, flavour, effect`. Thus, **`type`** should be seen as a relation rather than an attribute – the hyponymic relation between the adjective-noun phrase as a whole and the noun head:

(39) a. black tea *is_a_type_of* tea
b. light beer *is_a_type_of* beer
c. black comedy *is_a_type_of* comedy

Indeed, all the three examples are listed in the Princeton WordNet[12] as direct hyponyms of their head nouns. In German, such phrases sometimes become closed nominal compounds with an adjectival modifier: *schwarzer Tee* → *Schwarztee* 'black tea', *bittere Schokolade* → *Bitterschokolade* 'bitter chocolate'.

[11] Similarly to **`effort`**, it is a specification of the TELIC dimension of a noun's qualia structure.

[12] `http://wordnetweb.princeton.edu/perl/webwn`, *last accessed on 05.08.2023*

6.2.3 Attributes in free phrases

The annotation of free phrases did not pose any problems to the annotators (see section 6.1.3). The attributes in free phrases refer mostly to physical properties of entities. van der Wouden (1997) points out that certain meanings expressed by modifiers are rarely collocational: for instance, color and material names. The GerCoAt dataset provides empirical support for this observation and extends it to other meanings: **`shape`**, **`texture`**, **`dimension`**, **`surface`**, **`brightness`**, **`taste`**. Such meanings are referred to by Mel'čuk (1999) as **LF-constants**, i.e. meanings that are expressed in non-constrained phrases. LF-constants have the same expression for any base: for instance, the adjective *rund* 'round' can be used to describe **`shape`** of any noun that has this dimension and a corresponding shape (resembling a circle). Thus, the adjectives selecting these attributes are lexically not highly varied in the GerCoAt dataset although the number of instances is rather high (see table 6.7): **`color`** and **`material`** have entropy equal to 0, **`brightness`** and **`shape`** less than one.

One attribute deserves a more detailed discussion, namely, **`material`**. In the GerCoAt dataset, the attribute **`material`** is represented by the denominal adjective *hölzern* 'wooden'. In the German wordnet's taxonomy of adjectives, *hölzern* 'wooden' is a hyponym of the node `beschaffenheitsspezifisch` 'texture-related'. The adjective is connected to the synset of the noun it is derived from, *Holz* 'wood', via the lexical relation *pertains to*. Among other adjectives listed under `beschaffenheitsspezifisch` 'texture-related', there are more similar examples: *steinern* 'stone', *kupfern* 'bronze', *silbern* 'silver', etc. All those adjectives are relational, and the common assumption about relational adjectives is that they are not property-denoting. Therefore, strictly speaking, the attribute **`material`** is not an attribute, but a relation. As it is straightforward that the relation is *made_of* for all the **`material`**-related adjectives listed above, it was decided to keep it in the inventory.

6.2.4 Mixed attributes

The attributes discussed further are found both in free phrases and collocations.

`Evaluation` refers to a subjective positive or negative assessment of an entity (concrete or abstract), process, or action. The phrases in (40) illustrate this attribute with free phrases and, in example (41) – with collocations.

(40) a. *herrlicher Wein* 'wonderful wine'
b. *herrlicher Duft* 'wonderful scent'
c. *herrliches Wetter* 'wonderful weather'

(41) a. *dummes Gefühl* 'dumb feeling'
b. *helle Zukunft* 'bright future'
c. *süße Katze* 'sweet cat'

For some phrases, it is possible to give a more precise interpretation of which aspect is being evaluated: for *wine*, it is likely its taste, however, it could also be its quality. The positive evaluation of *weather* can refer to the temperature, the precipitation, the wind, or the sun, etc. The collocations in (41) pose an even greater challenge due to the level of abstractness of the head nouns *feeling* and *future*. The nominal base in the collocation *sweet cat*, on the other hand, is not abstract. However, the modifier *sweet* does not provide a lot of additional information about the noun. Its only role is to indicate a positive evaluation by the speaker.

`Social`. The label `social` stands for "social aspect". It is a generalization over a few more fine-grained attributes such as `nationality`, `status`, `occupation`, `assets`, `religion`, etc. The more general category **`social`** has been included in the inventory in order to maintain an even level of granularity. For instance, some of the phrases cannot be reduced to the more specific categories listed above: *freies Land* 'free country' and *freie Bürger* 'free citizens'.

Other attributes that are almost equally frequent both in collocations and free phrases are **`quantity`**, **`climate`**, **`dimension`**. The latter is rather unexpected to be found in collocations as it refers to physical properties of concrete objects. In fact, it is a rare case where an adjective has a more abstract meaning as its literal meaning and acquires a concrete meaning when combined with concrete objects: the adjective *mächtig* 'massive'. Its literal meaning is 'powerful' as it is derived from the noun *Macht* 'power'. When combined with nouns such as *building, wall, tree*, it selects the attribute **`dimension`** from their semantics.

6.3 Nominal bases in collocations and free phrases

Abstract nouns are more prone to taking collocates than concrete ones. The analysis of the gold standard dataset GerCo in section 5.4.7 provided empirical support for this observation. A more detailed account of nominal bases in collocations and in free phrases is presented in this section. More specifically, the discussion will focus on the semantic classes of nouns and the typical attributes for the nouns of different classes. The main motivation for this analysis is to identify trends in the collocational behavior of nouns. The trends will be presented in the form of rough templates for each semantic class based on the frequencies in the gold standard dataset.[13]

Figure 6.7 depicts the distribution of collocations and free phrases across the semantic classes in the gold standard dataset GerCoAt. Similarly to the distribution across attributes (figure 6.6), three groups of semantic classes

[13] A possible application of such templates is in lexicography. Using ready-made templates with potential attributes for each semantic class of noun would facilitate the process of grouping the numerous collocates extracted automatically.

can be identified: (1) present to a large extent in collocations; (2) present to a large extent in free phrases; (3) mixed. The three groups will be discussed further and, for each of them, one example of an annotation template will be presented. The templates for all the classes are presented in Appendix D.

Collocational semantic classes

The semantic classes of nouns that exhibit a tendency to form collocations are EVENT, COMMUNICATION, COGNITION, FEELING, QUANTITY, ATTRIBUTE, RELATION. In this section, one of them will be discussed, namely, the template for the semantic class EVENT displayed in figure 6.8. The length of the dashed bars in the figure corresponds to the proportion R of each attribute for the given semantic class based on the GerCoAt dataset (formula 6.2):

$$R = \frac{n_{attribute}}{n_{instances}} \tag{6.2}$$

where the number of instances of the attribute **intensity** for the semantic class EVENT ($n_{attribute}$) is divided by the total number of instances of the semantic class EVENT ($n_{instances}$). Only the attributes with R higher than 1% have been considered in the templates. All the attributes with scores lower than that have been merged into the category **other**.

As shown in figure 6.8, the most frequent attribute for EVENT nouns is **intensity**, followed by **description**, **manner**, and **certainty**. Apart from that, events can have emotional, social, temporal, financial aspects, etc. It does not imply, that every event noun has all of these attributes. Consider the examples from the dataset: the noun *Lächeln* 'smile' has the attributes **emotion** (*bitter* 'bitter'), **evaluation** (*süß* 'sweet'), **intensity** (*leicht* 'light'), **intelligence** (*schlau* 'smart'), and **manner** (*sanft* 'soft'). The noun *Strafe* 'punishment', on the other hand, only has one attribute **intensity** for all its collocates: *deftig* 'severe', *hart* 'hard', *scharf* 'harsh', *mild* 'mild'.

Semantic classes in free phrases

The most non-collocational semantic classes are FOOD, ANIMAL, SUBSTANCE, and NAT OBJECT, i.e. mainly concrete entities. Such nouns are expected to be mainly characterized in terms of the four qualia roles (see section 4.3.1). Consider the template for the semantic class FOOD depicted in figure 6.9. The most typical attribute is **taste**, followed by **type** which is, strictly speaking, not an attribute. However, as it is a recurrent pattern for adjectives to modify food nouns in that way, it is worth including in the template. Apart from that, FOOD is characterized by **quantity, texture** and **consistency**, it may have a **cost**, and a **temporal** aspect. For instance, *Brei* 'porridge' has the attributes **taste** (*süß* 'sweet') and **consistency** (*zäh* 'thick') in the GerCoAt dataset.

Mixed semantic classes

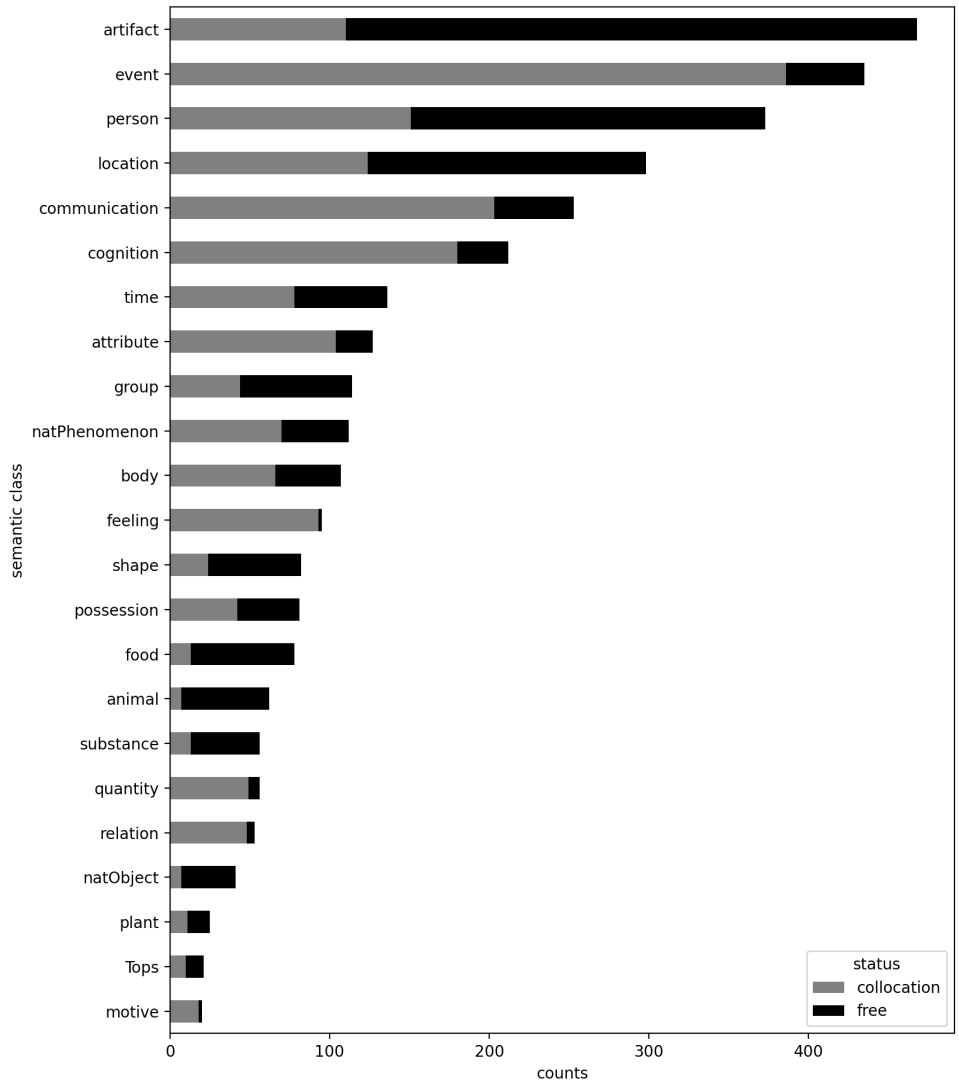


FIGURE 6.7: Distribution of collocations and free phrases in the GerCoAt dataset for each nominal semantic class. Sorted by the total number of instances a semantic class has in the dataset.

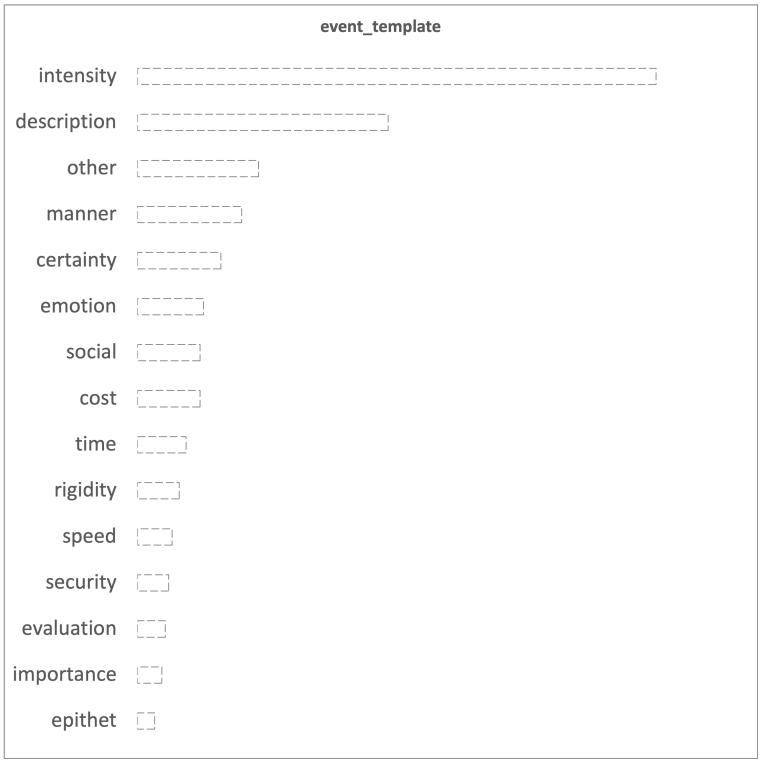


FIGURE 6.8: Attribute annotation template for nouns from the semantic class EVENT based on the frequencies in the gold standard dataset GerCoAt.

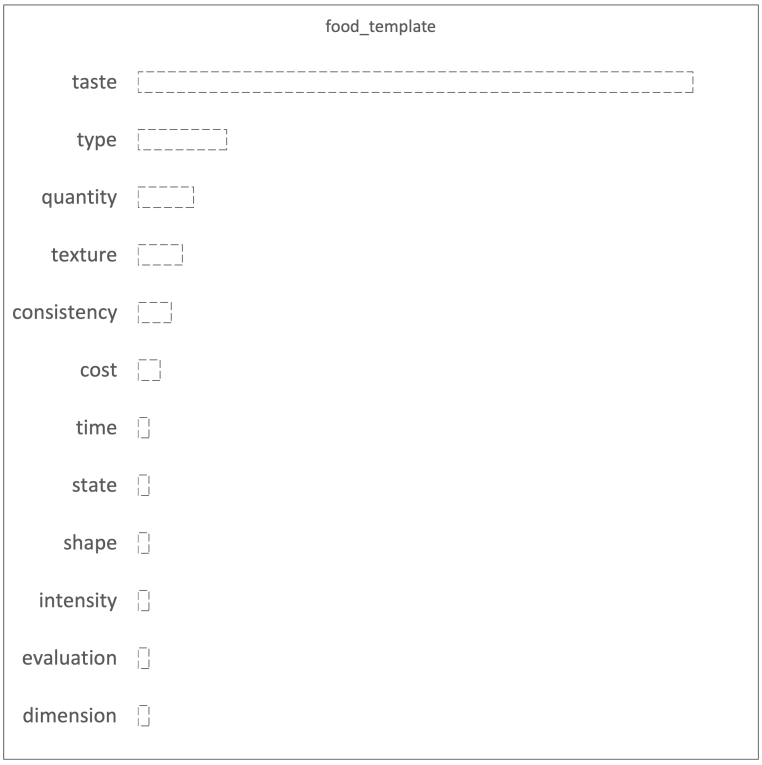


FIGURE 6.9: Attribute annotation template for nouns from the semantic class FOOD based on the frequencies in the gold standard dataset GerCoAt.

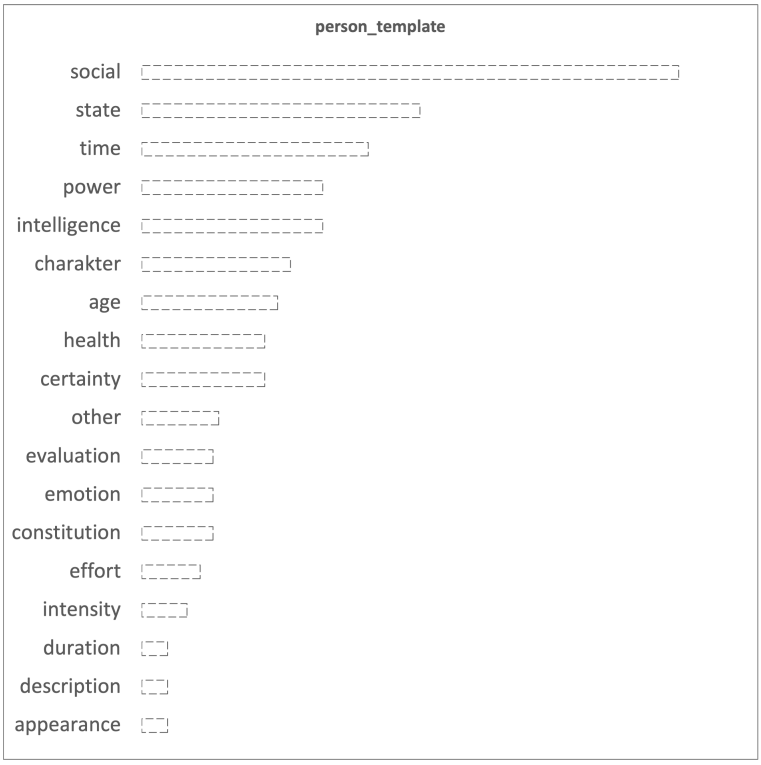


FIGURE 6.10: Attribute annotation template for nouns from the semantic class PERSON based on the frequencies in the gold standard dataset GerCoAt.

Semantic classes such as LOCATION, PERSON, ARTIFACT, TIME, GROUP, BODY, and NAT PHENOMENON are almost equally frequent both in collocations and in free phrases. The reason for that is that these classes are indeed of dual nature and comprise both concrete entities or abstract concepts. Figure 6.10 presents the annotation template for the nouns from the semantic class PERSON. As expected, the prevailing attribute is **`social`**, followed by **`state`**, reflecting the dual nature of the nouns in this semantic class: social and biological. Moreover, some nouns are ambiguous and can either refer to a human being or their social role depending on the modifier: the example mentioned a few times in this thesis is the noun *Freund* 'friend'. In the GerCoAt dataset, the reference is made to a friend's social role (=friendship) by the collocates *alt* 'old' and *dick* 'close', as well as to a friend as a human being by the adjective *tot* 'dead'.

Other semantic classes in this group are also represented by both concrete and abstract entities. ARTIFACTS such as *Komödie* 'comedy' and *Brief* 'letter' are less concrete than *Sonnenbrille* 'sun glasses' and *Teppich* 'carpet'. Similarly, the semantic class BODY includes concrete nouns such as *Mund* 'mouth' and *Haut* 'skin' as well as body-related states and processes such as *Schlaf* 'sleep' and *Fieber* 'fever'.

General remarks

The templates have been created for each of the 23 nominal semantic classes.

They reflect the major tendencies in the combinatorial behaviour of nouns modified by adjectives. These templates can be seen as rough concept frames that were discussed in Chapter 4 in section 4.5. As the selection of the data for the dataset was systematic and balanced (three adjective per semantic class, for details see section 5.2), the resulting dataset is unbiased and is expected to manifest the general tendencies in language. Therefore, the templates can assist lexicographers in semantic grouping of adjectival collocates of nouns.

6.4 Machine learning experiments

Similarly to the previous chapter (see section 5.5), the robustness of the annotation scheme will be tested in a number of machine learning experiments. In contrast to the binary classification task with categories *collocations* and *free phrases*, the new task is way more challenging as the classification scheme contains 49 labels. A conventional multi-class classification is hardly feasible with such a small dataset. The experiments in this section are inspired by those reported in Hartung et al. (2017) and rely on composition functions.

The main assumption in Hartung et al. (2017, p. 56) is that "*adjectives, nouns and attributes (via attribute nouns) can be embedded in the same semantic space*". In other words, the models learn a composed representation based on the adjective and noun vector, which is intended to reveal the implicit attribute in an AN phrase. This representation should be as close as possible to the vector for the corresponding, gold-standard attribute. For example, applying the function to the input words *grob* 'coarse' and *Sand* 'sand' should result in a vector that is close to `texture`.

In this approach, the vector space in which adjectives, nouns and attribute vectors are located plays an important role. It is expected that models will perform differently for collocations versus free phrases, since the function has to induce a semantic shift for adjectives in a collocation if the original vector of the adjective does not represent the shifted meaning. In contrast, the function does not need to perform a large meaning shift for adjectives in free phrases because the individual vectors already show a high degree of similarity in the underlying vector space.

Two experiments are conducted. First, the overall success of two composition models in the attribute-selection task is explored with different amounts of training data. Based on the results of the first experiment, the best model and the best training setting are selected for the second experiment. More specifically, the performance of the model is tested separately for collocations and free phrases.

The work reported in this section has been carried out by the author of the thesis in collaboration with Eva Huber and Neele Falk who were responsible for the technical implementation of the experiments, and Prof. Erhard Hinrichs who gave theoretical and conceptual advice at each step of the experiments.

6.4.1 Data

The gold standard dataset GerCoAt described above is the empirical bases for the experiments. However, the size of the dataset is quite small considering the number of labels (see Table 6.6). A higher number of training samples could significantly improve the performance of the models as neural networks generally benefit from large training data.

Recall that during the annotation, the annotators were instructed to annotate lexical unit IDs in GermaNet for all the adjectives and nouns in the dataset. Thus it is possible to automatically extract semantically related lemmas from GermaNet.[14] The attributes annotated in the original dataset are expected to hold between semantically related adjectives and nouns. This logic applies both to free phrases and collocations. For instance, the original dataset contains the collocation *tiefer Ton* 'low-pitched sound' whose attribute is `sound`. This particular sense of the adjective *tief* has a dedicated entry in GermaNet. The antonym of this sense of *tief* is the adjective *hoch* 'high-pitched'. In this way, the phrase *hoher Ton* 'high-pitch sound' is automatically constructed annotated with the attribute `sound`. Additionally, it is also possible to substitute the noun with a semantically related LU, for instance, a co-hyponym of the given sense of the noun *Ton* in GermaNet is *Pfeifen* 'whistle'. Based on that, two more phrases are constructed: *hohes Pfeifen* 'high-pitched whistle' and *tiefes Pfeifen* 'low-pitched whistle', both annotated with the attribute `sound`.

Consider another example: the free phrase *schwarze Hose* 'black pants'. The adjective can be substituted by its co-hyponyms such as *blau* 'blue', *gelb* 'yellow', *rot* 'red, etc. The noun is sibstituted by its hyponyms such as *Schlaghose* 'flared trousers' and *Stoffhose* 'cloth pants', and its co-hyponyms such as *Bluse* 'blouse' and *Pullover* 'sweater'. All the new phrases are annotated with the same attribute as the original phrase, `color`.

In this manner, for all the LUs in the original dataset, semantically related adjectives and nouns are extracted from GermaNet using conceptual (*hypernymy* and *hyponymy*) and semantic (*synonymy* and *antonymy*) relations. The resulting new phrases are annotated automatically with the attributes from the original phrases.

In order to eliminate nonsensical phrases that inevitably get constructed in the automatic augmentation,[15] the adjective-noun phrases are filtered by their bigram frequencies in a large corpus consisting of several German treebanks (Wikipedia 2017 and Wikipedia 2018 (de Kok and Pütz, 2019), and the corpus DECOW16AX (Schäfer, 2015; Schäfer and Bildhauer, 2012)). The frequency threshold has been set to five.

[14] `https://github.com/Germanet-sfs/germanetpy`, *last accessed on 28.08.2023.*

[15] Co-hyponyms introduce an especially large amount of noise into the data. For instance, a co-hyponym of the adjective *alt* 'old' in the sense 'existing for a long time' is the adjective *isochron* 'isochronous'. Adding this adjective to the sample resulted in the nonsensical phrase *isochrone Freunde* 'isochronous friends'.

Two augmented datasets have been created which, from the smallest to the biggest, include incrementally more types of semantically related LUs:[16]

- **augmented small**: synonyms, antonyms and hyponyms
- **augmented large**: synonyms, antonyms, hyponyms, co-hyponyms and hypernyms

The exact numbers for each dataset are presented in Table 6.10. Five different train/dev/test splits are created in order to make sure that different samples of the datasets are tested. For each split, the training data is filtered such that no adjectives or nouns overlap between training and dev/test sets.

data	phrases	adj	nn
original	3,094	46	2,030
augmented small	18,783	655	7,264
augmented large	84,305	2,152	14,669

TABLE 6.10: Data overview: augmented and filtered by frequency.

6.4.2 Models

Both composition functions used in the experiments take the input vector for the adjective **a** and the input vector for the noun **n** and compose them to retrieve a representation for the hidden attribute **p**. The vector of the hidden attribute has the same embedding dimension, d, as **a** and **n**. The composition models are trained to minimize the cosine distance between the composed representation of the hidden attribute **p** and the vector of the correct attribute $\bar{\mathbf{p}}$:

$$argmin(1 - \cos(\mathbf{p}, \bar{\mathbf{p}})) \tag{6.3}$$

Transformation Weighting Model (TWM) This model was proposed in Dima (2019), who compared it to various other composition functions on the task of reconstructing the corpus-induced vectors of short phrases on which it has produced the best results. In this model, a number t of non-linear transformations $\mathbf{T} \in \mathbb{R}^{t \times 2d \times d}$ are applied to the concatenated input vectors **a** and **n** resulting in a matrix **H** of t transformed composed representations:

$$\mathbf{H} = g(\mathbf{T}[\mathbf{a}; \mathbf{n}] + \mathbf{B}) \tag{6.4}$$

These transformed representations are weighted into a final representation **p** with a double-tensor contraction and a third order tensor $\mathbf{C} \in \mathbb{R}^{t \times d \times d}$:

[16] Augmented datasets can be downloaded here: https://owncloud.gwdg.de/index.php/s/PXKlLRG7UJ0BOB2.

$$\mathbf{p} = \mathbf{C} : \mathbf{H} + \mathbf{b} \tag{6.5}$$

Dropout is applied to the transformed representations in **H** and a `ReLU` non-linearity is used. The different transformations allow capturing different interactions between the input vectors. The number of non-linear transformations is smaller than the size of the input vocabulary, thus the model is forced to reuse similar transformation-weighting combinations for similar inputs.

Full Weighted Additive Model (FWAM) This model, developed by Guevara (2010), was the best-performing model on the task of attribute selection in Hartung et al. (2017). The core of this model lies in a position-specific transformation of the input vectors **a** and **n**: two linear maps, one for the modifier (**A**) and one for the head (**N**), are trained and applied to the corresponding input vectors. The transformed vectors are combined via addition:

$$\mathbf{p} = \mathbf{A}\mathbf{a} + \mathbf{N}\mathbf{n} \tag{6.6}$$

The linear maps are initialized with identity matrices so that the model starts off with simple addition.

6.4.3 Training and evaluation

As described in the previous section, the models are trained to create composed representations that are close to the underlying hidden attribute of the AN phrases. The final composed representations are normalized to unit length. Note that the vectors of the adjectives, nouns and attributes originate from the same vector space. Following the evaluation by Baroni and Zamparelli (2010) for a related task, a rank-based evaluation is conducted. More specifically, a sorted list of all possible attributes is computed based on the cosine similarity between the composed representation and the attribute vectors. The rank of a given test instance equals the position of the correct attribute in that sorted list.

For each attribute, the F1-score is computed considering the most similar (F1@1) and the five most similar (F1@5) vectors. The F1-scores are averaged and the F1-macro for each rank is reported. The models are trained with early stopping.

In the experiments, pretrained word embeddings with subwords are used. They are provided by finalfusion[17] and were trained with the skip-gram algorithm Mikolov et al. (2013). The embeddings were trained on the the corpus TüBa-D/DP de Kok and Pütz (2019).[18]

The embeddings for attributes are precomputed as a centroid of a paraphrase for each attribute. Thus the vector for `character`, for example, was

[17] https://finalfusion.github.io/pretrained

[18] The embeddings have a dimension of 300 and were trained with negative sampling and a context size of 10.

computed as the average of the vectors for *character, trait, personality, temperament, behaviour, reaction*. The paraphrases are stem from the annotation guidelines (see Appendix C). The dropout rate and the number of transformations in the transformation weighting model were tuned on the development sets.[19]

6.4.4 Experiment 1

Table 6.11 reports the results averaged over all splits. As baselines, a nearest neighbour search was performed based on a) the vector of the adjective, b) the vector of the noun and c) the centroid of the adjective- and noun vector.

	F1@1	**F1@5**
baselines		
adjective	0.092	0.278
noun	0.141	0.357
avg. phrase	0.157	0.438
original data		
FWAM	0.219	0.552
TWM	0.230	0.495
augmented small		
FWAM	0.227∗	0.505
TWM	0.350∗	0.578∗
augmented large		
FWAM	0.238	0.557∗
TWM	**0.427**∗†	**0.623**

TABLE 6.11: Average results on the random test set splits: macro F1-score at rank 1 and at rank 5. The TWM with the largest training dataset significantly outperforms all other experimental settings.

Statistical significance ($p < 0.05$) was computed using the paired-bootstrap test (Efron and Tibshirani, 1994), comparing (1) whether the difference for the same model trained on the different training data is significant (significance is marked with a $*$ in the table) and (2) whether the best model is significantly better than all the other models (significance is marked with a † in the table).

The results show that both composition functions outperform all baselines. With respect to data augmentation, the results show that automatic enrichment is helpful for several models. The largest improvement can be seen for the TWM: on the original data set, the model can only achieve a macro-F1 score of 0.23 despite a lexical overlap; with the large synthetic training data, there is an improvement of about 20 percent.

[19] 50 transformations, 0.4 dropout rate. Dropout rates were tuned between 0.0 and 0.8 in increments of 0.2. The number of transformations was tuned given the following options: 10, 30, 50, 100, 200

Comparing the composition functions reveals that the TWM outperformes FWAM for all datasets. The gap between the performance of the two composition functions increases with more training data. This suggests that the TWM has a higher generalization power when more training data is available, as it has the flexibility to use different combinations of its parameters for different AN combinations. The positive influence of more training data can be seen more clearly when looking at the results for rank 1.

Since the TWM trained on the largest training set achieves the highest performance, that model will be used in the next experiment. For each experiment, we create three random train/dev/test splits with no lexical overlap.

6.4.5 Experiment 2

In this experiment, the attributes are investigated with regard to the phrases' collocational status. The performance of the model is tested on the attributes that are encountered both with free phrases and collocations (*mixed attributes* discussed in section 6.2.4). The number of instances of collocations and free phrases per attribute is balanced.

The overall results are presented in Table 6.12. The difference in performance between collocations and free phrases supports our hypothesis that literality of meaning is an important factor in modelling attributes in AN phrases. When the meaning of the phrase is literal, it is easier for the model to correctly predict the attribute: e.g. *freundlicher Hund* 'friendly dog' is predicted as `character`, or *herrliche Kombination* 'great combination' as `evaluation`. However, the model faces a challenge when tasked to predict attributes in collocations: e.g. in *süßer Hund* 'sweet dog', `taste` is predicted instead of `evaluation`, or in *raue Witterung* 'rough weather', `surface` is predicted instead of `climate`. The difficulty for collocations is increased by the absence of lexical overlap between the training and test data: the model does not see similar examples in the training set that it can learn from.

	F1@1	**F1@5**
free phrase	0.74	0.91
collocation	0.51	0.73

TABLE 6.12: Overall results for free phrases vs. collocations: macro F1-score at rank 1 and at rank 5.

The detailed results for each attribute are presented in Table 6.13. Although the overall tendency is clearly toward better performance for free phrases, there are some exceptions. Several examples indicate that there is a phenomenon that can be referred to as *the dominance of the non-literal sense* reflected in the embeddings. For instance, the attribute `climate` is represented in collocations by the adjective *karg*. Its literal sense is 'sparse' expressing `quantity`, but the collocational sense of 'barren' when describing natural objects like mountains, hills, islands, etc. is well-established in the German

language and does not pose any problems for the model to predict the correct attribute `climate`. Another example is the adjective *reich* 'rich': both of its attributes, `quantity` as in 'rich income/harvest' and `social` as in 'rich family/banker', are correctly predicted which indicates that both senses are equally present in the embedding.

	free phrases		**collocations**	
attribute	**F1@1**	**F1@5**	**F1@1**	**F1@5**
`character`	**0.86**	**0.93**	0.36	0.71
`climate`	0.26	0.55	**0.94**	**0.97**
`dimension`	**0.80**	**0.98**	0.43	0.82
`evaluation`	**0.89**	**0.97**	0.33	0.41
`power`	**0.73**	**0.96**	0.00	0.28
`quantity`	0.50	0.90	**0.77**	**0.97**
`social`	**0.88**	**1.00**	0.50	0.79
`time`	**0.96**	**0.98**	0.72	0.90

TABLE 6.13: Results for free phrases vs. collocations for each attribute: macro F1-score at rank 1 (@1) and at rank 5 (@5).

6.4.6 Summary

This study investigated which factors have an effect on the success of attribute selection in adjective-noun phrases in distributional semantics. A composition function which is flexible enough to use different parameter combinations for different types of phrases can benefit from an automatic enrichment of training data (Experiment 1). The composition model can accurately predict attributes that occur in free phrases, if enough training data is available. However, collocations pose a problem for the distributional models. As it is less likely for similar adjective-noun combinations to be encountered in training, the models cannot easily generalize for unseen data (Experiment 2). Another insight gained from Experiment 2 is that there are some attributes that are predictable in free phrases as well as in collocations. These attributes occur frequently in both the original and the augmented datasets and can be selected by many different adjectives.

6.5 Modifying the level of granularity

The annotation scheme discussed until this point can be considered rather fine-grained with almost 50 labels available. Modifying the level of granularity for this inventory is feasible in the direction of higher generalization, i.e. merging the labels into more coarse-grained categories. This step may resolve the issue of high imbalance of the categories that is characteristic of the GerCoAt dataset. The additional level of granularity introduced for the GerCoAt dataset is based on the taxonomy of adjectives in GermaNet: using the top-nodes as coarse-grained attributes – 16 categories in total.

This layer of annotation has been used in machine learning experiments on attribute selection in adjective-noun phrases presented in Falk et al. (2021). The inventory of labels is based on the top-nodes in the adjectival taxonomy in GermaNet as illustrated in figure 6.11.

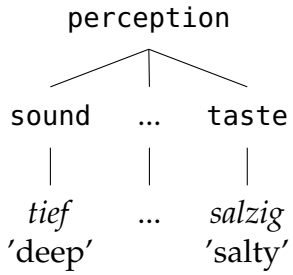


FIGURE 6.11: A part of the taxonomy of adjectives in GermaNet for *tief* 'deep' and *salzig* 'salty'. The top node **perception** is used as an attribute label (Falk et al., 2021, p.241).

In this manner, an inventory comprising 16 labels has been extracted from GermaNet. The label set is presented in table 6.14 with the corresponding fine-grained attributes from the original annotation scheme. The coarse-grained label that comprises the highest number of fine-grained attributes is **relation**: it includes 11 specific attributes. In contrast,the attributes **emotion**, **climate**, **society**, etc. are mapped to only one or two fine-grained categories. Three of the coarse-grained categories stand out as they cannot be aligned with the top-nodes in GermaNet: **intensity**, **evaluation**, and **manner**. The first two labels are subcategories of the top-node **klassenübergreifend** 'general' which is way too general and is practically undefinable. The third attribute, **manner** is not a part of GermaNet taxonomy of adjectives at all.
In terms of balance, the dataset with the coarse-grained granularity is slightly more balanced than the original one. Figure 6.12 depicts the distribution of phrases across the coarse-grained labels. There are six categories that have less than 100 instances. **Intensity**, however, remains the most frequent category. The most collocational categories are **intensity**, **manner**, **relation**, and **emotion**.

The experiments on automatic prediction of attributes in Falk et al. (2021) further support the hypothesis that due to the idiosyncratic lexical behavior, collocations are more challenging for neural models than free phrases. In order to test this, Falk et al. (2021) use a subset of the dataset balanced in terms of the amount of collocations and free phrases for each attribute. The subset contains the following attributes: `climate`, `quantity`, `time`, `society`, `location`, `behaviour`, `evaluation`. The models are trained on two automatically augmented datasets: small (21,498 AN pairs) and large (232,398 AN pairs). In order to explore the impact of lexical memorization on the attribute selection task, the experiment is conducted in three lexical settings: (1) **No overlap** The validation/test and training have distinct vocabulary.

coarse-grained	**fine-grained**
behaviour	character, naturalness, power
body	appearance, health, constitution
climate	climate
evaluation	evaluation
feeling	emotion
intensity	intensity
location	space, dimension, shape
manner	manner, description
motion	motion
mind	intelligence
perception	color, smell, taste, brightness, sound, surface
quantity	quantity, costs
relation	accuracy, certainty, stability, difficulty, safety, rigorousness, completeness, functionality, importance, validity, compatibility
society	social
substance	material, texture, consistency, temperature, state, weight
time	duration, time, age, speed

TABLE 6.14: Second layer of annotations in the GerCoAt dataset, based on GermaNet.

(2) **Modifier overlap** The validation/test and training share modifiers (adjectives). (3) **Head overlap** The validation/test and training share heads (nouns) (Falk et al., 2021, pp. 243-244). Table 6.15 shows the results of the experiments. In all settings, a model's performance is better for free phrases than collocations. However, Falk et al. (2021) report that despite the commonly perceived idiosyncrasy of collocations, there are regular patterns present. For instance, the attribute **`climate`** consistently achieves a high F1 score for collocations across all experimental settings, ranging from 0.67 to 0.87. This suggests that the shifts in meaning for adjectives associated with this attribute follow regular patterns. Another illustration of such consistent meaning shifts is observed with the polysemous adjective *süß* 'sweet'. In its literal sense, it selects the attribute **`perception`** as in *süße Torte/Tee* 'sweet cake/tea'. However, *süß* can also select the attribute **`evaluation`** when combined, for example, with nouns from the semantic field PERSON, as in *süßes Kind* 'sweet child'.

training data	**fastText**		**BERT**	
	free phrase	collocation	free phrase	collocation
		small		
no overlap	0.66	0.53	0.59	0.44
modifier overlap	0.74	0.57	0.67	0.59
head overlap	0.80	0.73	0.73	0.67
		large		
no overlap	0.73	0.61	0.62	0.58
modifier overlap	0.84	0.73	0.87	0.72
head overlap	0.75	0.61	0.67	0.63

TABLE 6.15: Average Macro F1 score for the balanced set in terms of collocations and free phrases for each training set (Falk et al., 2021, p.247).

Systematic attributes in collocations such as **`intensity`** or **`evaluation`** are predicted with scores higher than 50% on unseen data. Figures 6.13 and 6.14 show the results of the experiments for each attribute in different settings on the full unbalanced dataset.

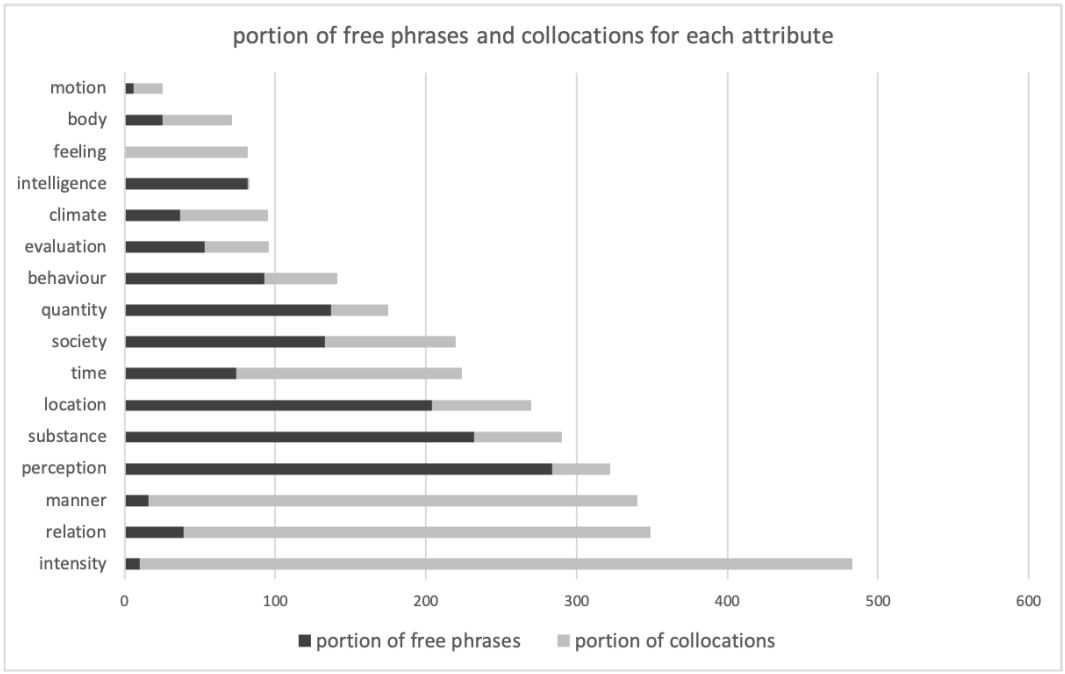


FIGURE 6.12: Distribution of free phrases and collocations in the GerCoAt dataset across the second-layer categories. The figure is from Falk et al. (2021, p.243).

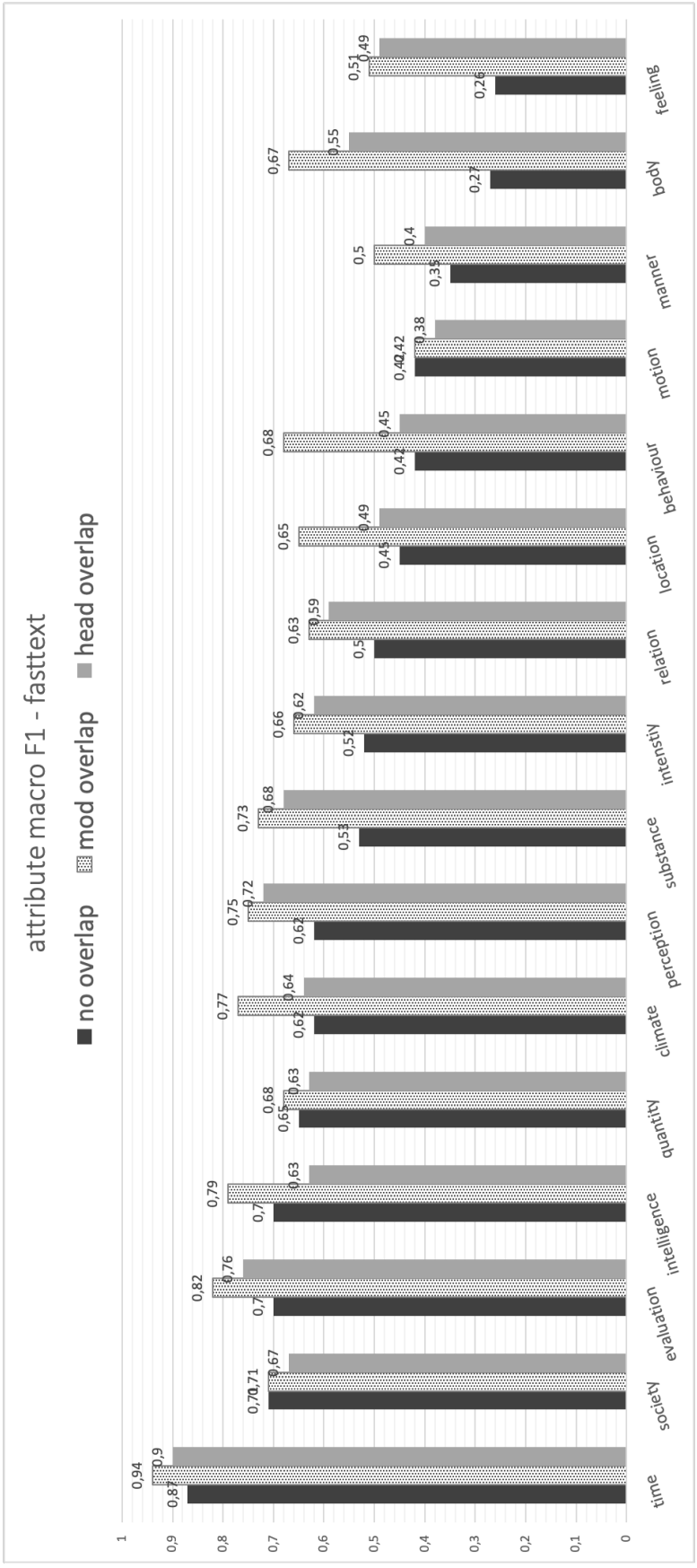


FIGURE 6.13: The results of the experiments with static embeddings in Falk et al. (2021, p.244).

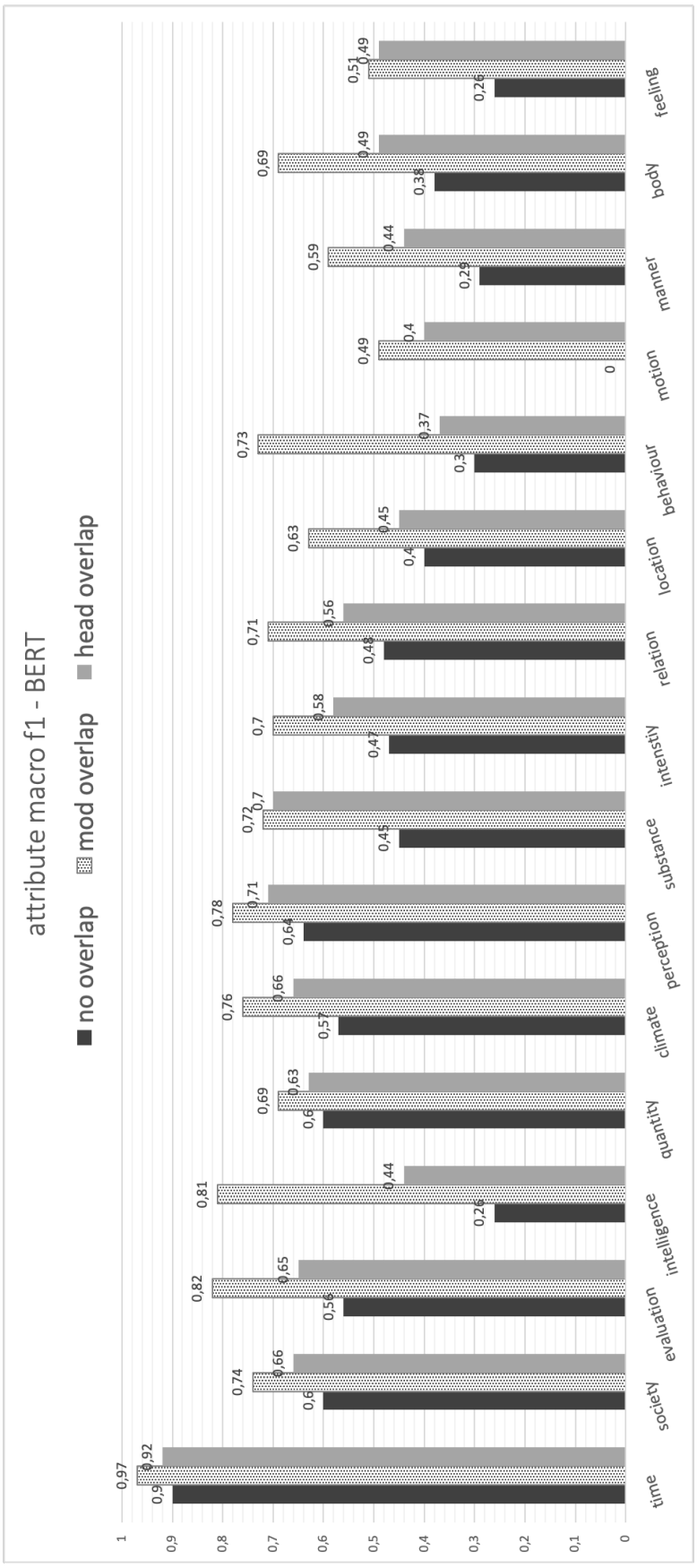


FIGURE 6.14: The results of the experiments with contextualized embeddings in Falk et al. (2021, p.244).

6.6 Summary

The gold standard dataset GerCoAt described in this chapter contains lexical and semantic information about more than three thousand German adjective-noun collocations and free phrases in context. There are several layers of granularity of the manual annotation:

1. **Sense IDs**: the most fine-grained annotation. Each adjective and noun has been provided with its sense ID in the German wordnet GermaNet;[20]
2. **Fine-grained attributes**: the annotation scheme contains 49 labels;
3. **Coarse-grained attributes**: this level of annotation is represented by 16 labels;
4. **Literality**: binary annotation into two groups: "collocation" and "free phrase" (described in chapter 5).

The analysis of the data has been performed for levels 4 and 2 of granularity: the literality and the fine-grained attributes.

Concreteness: the data provides empirical support for the earlier observations made in chapter 4 of this thesis that the nominal bases in collocations are on average more abstract than those in free phrases. There are also noticeable tendencies among the fine-grained attributes: the average concreteness scores for the nouns that have the attributes **`surface`**, **`color`**, and **`material`** are almost twice as high as for the ones characterized by **`speed`**, **`epithet`**, **`rigidity`**, **`intensity`**.

Semantic classes: in the annotation process, a necessity arose to use some sort of heuristic to make sure that the proposed attributes are of same level of granularity. Based on the annotated dataset, annotation templates have been created for each semantic class of nouns. The inspection of the proposed templates goes in line with the intuition and world knowledge of the annotators. For instance, the nouns of the semantic class FOOD have the following aspects according to the template: `taste`, `type`, `quantity`, `texture`, `consistency`, `cost`, `time`, `state`, `shape`, `intensity`, `evaluation`, and `dimension`.

Attributes: none of the attributes can be considered purely collocational or not. For that reason, it was important to compile an annotation scheme accommodating both collocational and non-collocational meanings. Rather, certain meanings have a tendency to be expressed in collocations, for instance, the meaning of intensification, whereas meanings such as color or dimension are overwhelmingly expressed in free phrases.

[20] This layer of annotation could be used as gold standard in tasks on Word Sense Disambiguation.

Various machine learning experiments (reported in section 6.4 and inFalk et al. (2021)) show that the difference between free phrases and collocation is also picked up by the tested models. Attributes in collocations are more challenging for the models to learn than those in free phrases. However, if a meaning is expressed systematically, the prediction scores are also high for collocations. Strongly lexicalized collocations such as *blinder Passagier* 'stowaway' (lit. 'blind passenger') and *helle Zukunft* 'bright future' still pose a challenge to the models (Falk et al., 2021).

Chapter 7

Validating the inventory with different data

The annotation scheme for adjective-noun phrases described in the previous chapter was developed based on a very specific dataset (GerCo), where the adjectives have been selected in a very systematic manner (see section 5.2.1). Notwithstanding, the proposed inventory of attributes is expected to be general enough to accommodate lexically diverse data in any language.

In order to investigate the generalization power of the label set, an already existing dataset of adjective-noun phrases have been chosen – the English dataset HeiPLAS (Hartung, 2015) that has been introduced in section 4.5. The dataset is lexically rich: it contains a large number of unique adjectives and nouns. This chapter investigates to what extent the inventory proposed in the previous chapter of this thesis can accommodate all the meanings available in the HeiPLAS dataset.

7.1 Overview of the HeiPLAS dataset

The construction of the HeiPLAS dataset (Hartung, 2015) is described in section 4.5 of this thesis. The AN pairs in HeiPLAS are annotated with attributes that stem from WordNet. Table 7.1 gives an overview of the dataset: it contains a large number of diverse adjectives and nouns, and, in this respect, differs from the GerCoAt dataset.

phrases	adjectives	nouns	attributes
1,598	849	923	254

Table 7.1: Overview of the HeiPLAS dataset.

A much larger number of attributes (254) in comparison to the GerCoAt dataset can most likely be explained by the finer level of granularity of the labels. An alternative, yet less plausible explanation is that there are simply much more various adjectival meanings in the HeiPLAS dataset than in the GerCoAt dataset. Bear in mind that GerCoAt has only 46 unique adjectives in 3,307 adjective-noun pairs (46/3,307: the ratio is about 0.01) whereas HeiPLAS contains considerably more diverse adjectives (849/1,598: the ratio=0.53).

7.2 Annotation

In order to investigate whether the GerCoAt annotation scheme is applicable to the new data, the HeiPLAS dataset has been annotated manually on two levels relying on the guidelines developed in chapters 5 and 6.

1. **Level 1**: GerCoAt attributes;
2. **Level 2**: binary annotation into "collocations" and "free phrases".

Due to time constraints, the annotation has been performed by two annotators for the two levels in parallel: one annotator for each level. Thus, the IAA could not be calculated and only the qualitative feedback of the annotators who are both experienced linguists will be used for analysis.

7.2.1 Annotation of attributes

This annotation was performed by the author of this thesis. The big advantage of using the HeiPLAS dataset for validation is the fact that it already has annotations of attributes. All the phrases were grouped according to the original labels and thus could be annotated in batches of phrases with similar meaning. There are three cases of mapping of the HeiPLAS attributes to the ones from the GerCoAt dataset: (1) **splitting** an attribute with a general into more specific attributes; (2) **merging** specific attributes; (3) introducing **new** attributes.

Splitting. There are several examples where the interpretation of an attribute in HeiPLAS is very general and the attribute is mapped to several more specific ones from GerCoAt.[1] Such attributes are AGE and SIZE. The attribute AGE in HeiPLAS has 11 instances in total. As table 7.2 shows, the single attribute AGE in HeiPLAS is mapped to three different attributes in GerCoAt. It indicates that AGE is interpreted much more broadly in HeiPLAS: AGE for the phrase *new friend* refers metaphorically to the age of friendship which corresponds to the attribute `duration` in GerCoAt.

The attribute SIZE in HeiPLAS comprises three different readings: (1) literal: the physical size of a concrete object as in *gigantic redwood* that maps to the attribute `dimension` in GerCoAt; (2) metaphorical: degree of an abstract concept as in *gigantic disappointment* that maps to `intensity`; (3) metaphorical: amount that maps to `quantity` as in *slender wage*.

Merging. Examples of splitting HeiPLAS attributes are few in comparison to the amount of highly specific attributes that had to be merged when

[1] Note that HeiPLAS does have the attributes corresponding to the ones borrowed from GerCoAt as well: DEGREE, DURATION, QUANTITY, TIMING.

attr HeiPLAS	adj	nn	attr GerCoAt
AGE	brand-new	shoe	`age`
AGE	little	boy	`age`
AGE	mature	cell	`age`
AGE	new	car	`age`
AGE	new	comet	`age`
AGE	young	people	`age`
AGE	new	law	`time`
AGE	new	year	`time`
AGE	newborn	fear	`time`
AGE	recent	addition	`time`
AGE	new	friend	`duration`
SIZE	big	city	`dimension`
SIZE	deep	cut	`dimension`
SIZE	gigantic	redwood	`dimension`
SIZE	small	car	`dimension`
SIZE	gigantic	disappointment	`intensity`
SIZE	stupendous	demand	`intensity`
SIZE	large	family	`quantity`
SIZE	slender	wage	`quantity`

TABLE 7.2: Splitting: Mapping of the attributes AGE and SIZE from HeiPLAS to the GerCoAt attributes.

aligned with the GerCoAt attributes. For example, the attribute `manner` comprises the total of 77 very specific attributes from HeiPLAS. Several examples are presented in table 7.3: some of them are indeed simply very fine-grained, such as CONSPICUOUSNESS, DULLNESS, FELICITY, FRIENDLINESS, and LIVELINESS, whereas some, such as ANIMATION and TASTE, are both fine-grained and used in a non-literal sense.

Another example is the GerCoAt category `dimension` that comprises more specific attributes such as HEIGHT, LENGTH, WIDTH, SIZE and STATURE. Similarly, the GerCoAt label `evaluation` comprises NICENESS, BEAUTY, PLEASANTNESS, QUALITY, WORTHINESS, and a dozen of other attributes from HeiPLAS.

New. Finally, as expected, the inventory from GerCoAt does not reflect all the aspects of nominal meanings expressed by descriptive adjectives and had to be extended. Seven additional attributes were introduced. Table 7.4 provides two examples for each of the new attributes.

- **`cause`** – similar to the AGENTIVE dimension in the Qualia Structure (Pustejovsky, 1991); refers to the origin of an entity, the factors responsible for bringing it about;

- **composition** – similar to the CONSTITUTIVE dimension in the Qualia Structure (Pustejovsky, 1991); refers to the chemical composition of an entity;
- **continuity** – steadiness and/or regularity of a process or event;
- **location** – refers to spacial characteristics of an entity having to do with its location such as distance, direction, position;

attr HeiPLAS	adj	nn	attr GerCoAt
ANIMATION	sprightly	dance	manner
CONSPICUOUSNESS	conspicuous	necktie	manner
DULLNESS	dull	mood	manner
FELICITY	clumsy	apology	manner
FRIENDLINESS	cozy	room	manner
LIVELINESS	lively	discussion	manner
TASTE	flashy	ring	manner
HEIGHT	high	building	dimension
LENGTH	short	hair	dimension
SIZE	immense	snake	dimension
STATURE	tall	tree	dimension
WIDTH	narrow	bridge	dimension
NICENESS	nice	day	evaluation
NICENESS	nice	dress	evaluation
BEAUTY	beautiful	country	evaluation
PLEASANTNESS	pleasant	evening	evaluation
QUALITY	good	secretary	evaluation
WORTHINESS	worthwhile	book	evaluation

TABLE 7.3: Merging: Mapping of various attributes from HeiPLAS to the GerCoAt attributes description, dimension and evaluation.

- **quality** – an objective or subjective estimation of whether something is correct. In a way, it is inspired by the Lexical Function Ver (Mel'čuk, 2015) that is an objective qualifier (see section 4.2.1). Objective qualifiers are adjectives such as *wrong*, *false*, *correct*, *right*, i.e. assessments that can be verified. Subjective qualifiers are those that rather refer to some standards or expectations of how something should be, such as *perfect*, *proper*, *inappropriate*.[2]

[2]In a way, it can be interpreted as deontic modality expressed by an adjective. As defined in Pustejovsky and Batiukova (2019), it "expresses an attitude towards how a state of affairs should be, with respect to social and ethical norms". See Linden (2012) for a discussion of deontic modality in adjectives.

attr HeiPLAS	adj	nn	new attr
NATURALNESS	unnatural	death	`cause`
INTENTIONALITY	intentional	damage	`cause`
PURITY	contaminated	water	`composition`
COMPLEXITY	simple	substance	`composition`
CONTINUITY	constant	pain	`continuity`
CONTINUITY	continuous	fear	`continuity`
POSITION	exterior	paint	`location`
DISTANCE	near	neighbor	`location`
TRUTH	false	testimony	`quality`
PERFECTION	perfect	specimen	`quality`
NORMALITY	abnormal	circumstance	`relation`
DIFFERENCE	different	meeting	`relation`
SEX	female	heir	`sex`
SEX	male	infant	`sex`

TABLE 7.4: New coarse-grained attributes for specific attributes in HeiPLAS.

- **`relation`** – puts the modified noun in context of other entities, a general comparison in terms of norm. Typical adjectives that express this meaning are *different, similar, foreign, ordinary, unusual*.
- **`sex`** – property that distinguishes organisms on the basis of their reproductive roles.

Not all the attributes from the GerCoAt inventory were used for the annotation of the English data. For instance, the attributes `rigidity`, `tolerability`, and `security` are probably too rare or too specific. The attribute `material`, on the other hand, is a very frequent property of concrete objects: all artifacts are characterized by the material they are made of. However, it is absent from the HeiPLAS dataset.

The resulting dataset, codenamed HeiPLAS+, comprises 1,447 adjective-noun pairs. The attribute inventory contains 49 labels. The frequency distribution across the attributes is depicted in figure 7.2. Similarly to the German dataset GerCoAt, HeiPLAS+ is imbalanced in terms of attributes: whereas the attributes `manner`, `relation`, `character`, `intensity` have about 100 and more instances, such attributes as `movement`, `sex`, and `weight` are found only in a few phrases. The English dataset is lexically more varied than the German one. The fact that it is imbalanced in a similar way, indicates that the imbalance is not an artifact of a particular dataset. First of all, it is a reflection of the adjectives' combinatorial behaviour in a language. Secondly, the skewedness of the data also suggests the attribute `manner` is indeed very general. However, as discussed in section 6.2.2, `manner` can only be split into numerous highly fine-grained attributes.

7.2.2 Annotation of collocations

The classification of the adjective-noun pairs into collocations and free phrases was performed based on the guidelines from section 5.3.1 by a native speaker of English and a linguist Richard Lawrence, PhD. According to the definition of collocation adopted in this thesis, there are two types of collocations: (1) the collocate's sense is not its literal sense; (2) the collocate is fully dependent on the base. Until this point, only collocations of the first type have been investigated. There are more adjectives in the English dataset than in the German one, which means that there is possibility to encounter the second type of collocations as well.

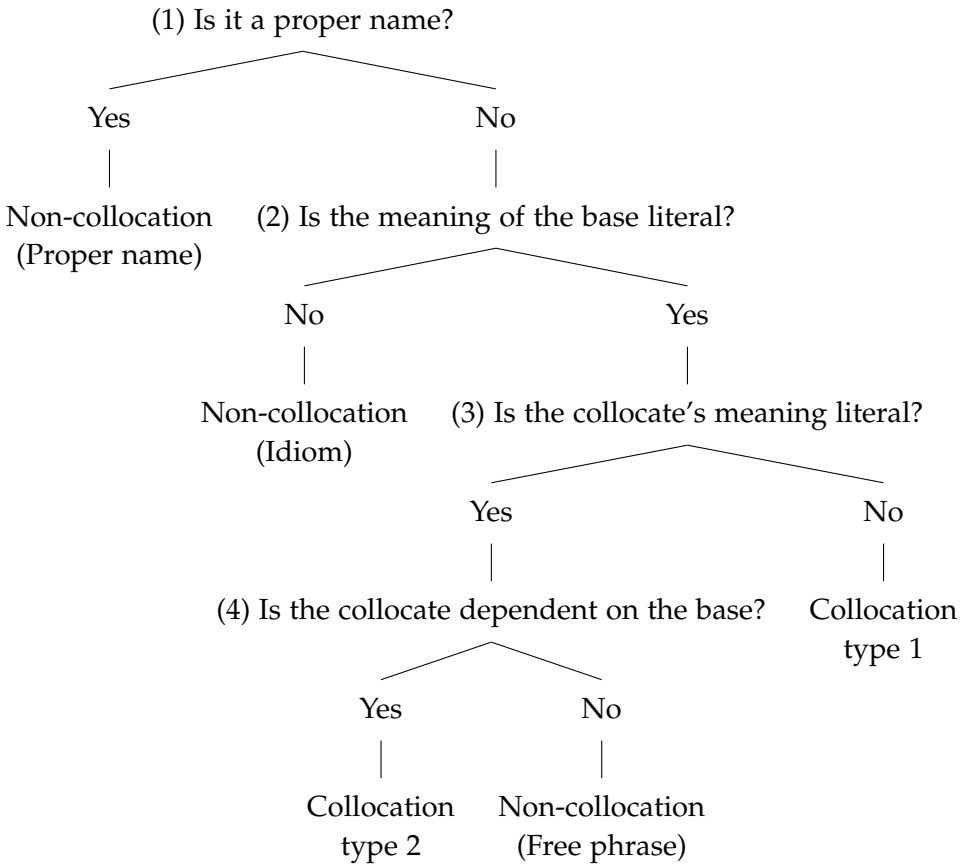


FIGURE 7.1: Extended annotation scheme for the classification of the collocation candidates.

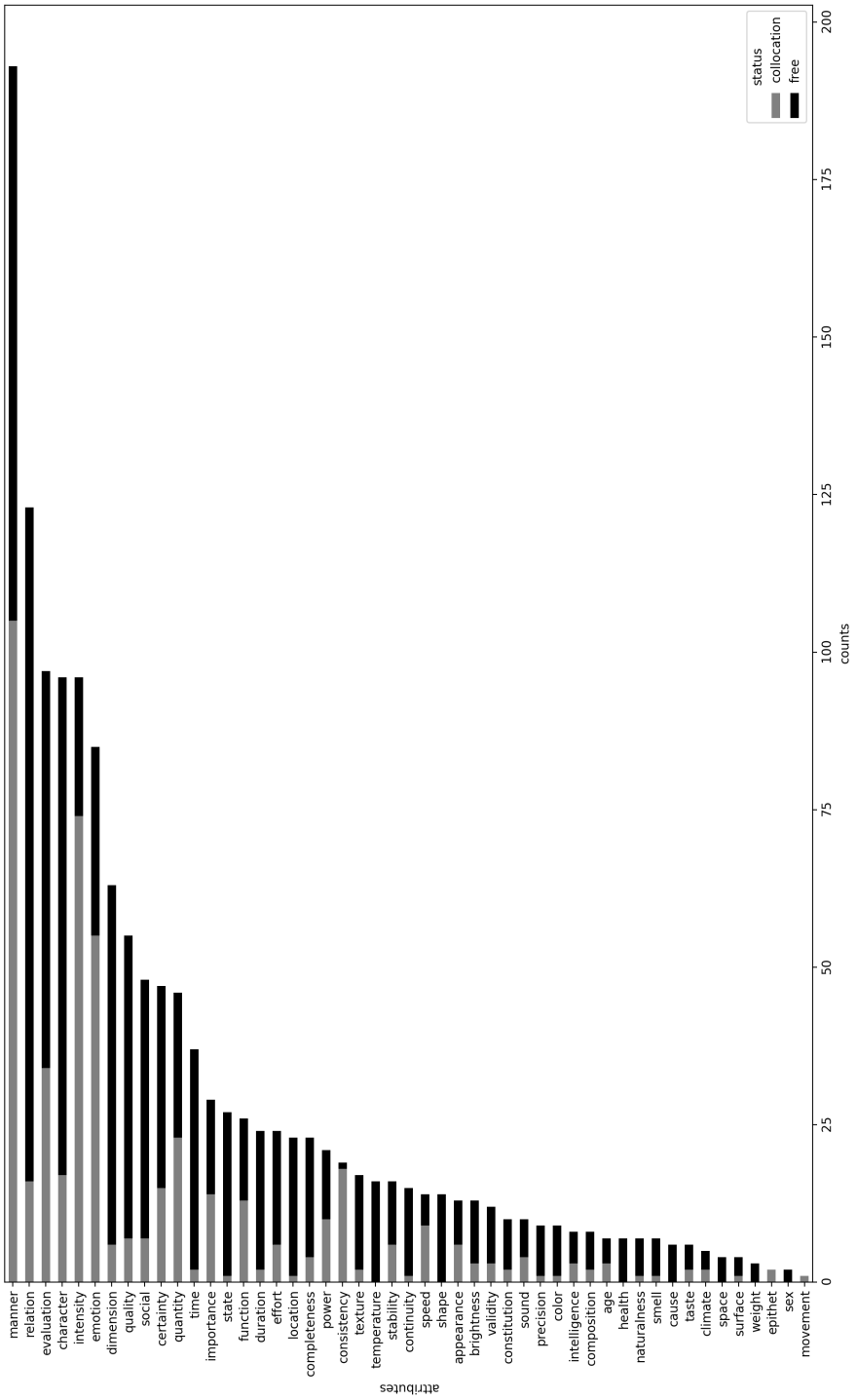


FIGURE 7.2: Distribution of collocations and free phrases for each attribute in HeiPLAS+.

The decision tree that served as the guideline for annotating collocations in section 5.3.1 has thus been extended to accommodate collocations of type 2. Figure 7.1 depicts the new decision tree with the additional step 4: "Is the collocate dependent on the base?".

> An adjectival collocate is **dependent on the base** if it can be combined only with very few nominal bases, for instance, *artesian well*. (from definition 6)

The total of 1,447 adjective-noun pairs were classified into 485 collocations and 961 free phrases. Only a couple of collocations were marked as "type 2" by the annotator, two of them are presented in example (42). The annotator's intuition that the given adjective-noun pair is a collocation of type 2 is backed up by two factors: usage of the phrase as a typical and often the only example in the adjectives' entries in different monolingual dictionaries and high association score in the Sketch engine in various corpora (Kilgarriff et al., 2014).

(42) a. *ill-gotten gains*: the adjective can also be combined with semantically-related nouns, such as *wealth* and *goods*. However, the phrase is so strongly lexicalized that in Macmillan online dictionary,[3] there is no separate entry for the adjective, but only an entry for the collocation as a whole;

b. *cushy job*: sometimes the expression *cushy number* is used instead, where *number=job*;

The distribution of collocations and free phrases across attributes is depicted in figure 7.2. As the dataset is skewed and contains twice as much free phrases as collocations, there are no purely collocational attributes. However, the trends observed in the English dataset are similar to the ones in GerCoAt:

- typical attributes in collocations are **`intensity`** as in *exquisite pain, bad headache*; **`emotion`** as in *acid comment, blue day*; **`speed`** as in *smart pace, lazy river*, etc.
- typical attributes in free phrases are **`dimension`** as in *large city*; **`quality`** as in *incorrect calculation*; **`time`** as in *recent addition*; **`texture`** as in *coarse sand*; **`shape`** as in *straight road*; **`color`** as in *white dress*; **`relation`** as in *similar price*; etc.
- attributes found (almost) equally in collocations and free phrases: **`manner`** as in *bouncy tune, soapy compliment, lively discussion*; **`quantity`** as in *short money, slender wage, ample supply*; **`function`** as in *potent liquor, powerful drug, ineffectual therapy*; **`evaluation`** as in *rosy prediction, bright future, nice day*.

[3] `https://www.macmillandictionary.com`, *last accessed on 18.12.2021*

7.3 Conclusion

The GerCoAt inventory of attributes has been successfully applied to a different dataset in a different language – English dataset of adjective-noun pairs and attributes HeiPLAS.[4] The mapping of the fine-grained attributes in HeiPLAS to more coarse-grained ones from GerCoAt was not homogeneous. Not all the attributes in HeiPLAS can be considered specifications of more coarse-grained attributes in GerCoAt. A few attributes in HeiPLAS were split into several different attributes from GerCoAt. In those cases, the discrepancy is due to the inclusion of a metaphorical meaning into the meaning labeled by an attribute in HeiPLAS. For instance, **`taste`** as physical sensation and as aesthetic quality.

Furthermore, as expected, the inventory had to be extended: seven new attributes have been added. Taking into consideration the high lexical variety of the English dataset, seven additional attributes are not many.

The extended inventory contains 54 labels that are suitable for classifying adjectival collocates of concrete and abstract nouns. Note that the scheme can be applied only to descriptive adjectives. The annotation of two dataset has shown that practically any attribute can be found both in collocations and in free phrases. The final extended inventory with definitions, examples, and additional comments is presented in Appendix E.

It is likely that the inventory is not completely exhaustive and the need for additional attributes might arise when dealing with unseen data. The intuition is, however, that there are not many meanings not covered by the inventory because the HeiPLAS dataset exhibits a very high level of lexical diversity and makes use of all the attributes available in WordNet covering all the semantic classes.

The annotation of HeiPLAS AN pairs into collocations and free phrases has shown that it is challenging to determine whether some pairs should be annotated as collocations of type 2 or free phrases. Is there a way to operationalize the restrictedness of use? In Chapter 3, examples similar to *cushy job* were discussed in terms of statistical asymmetry (section 3.4.2). It was suggested that an asymmetric statistical association measure such as ΔP_2 or D_{KL_2} can reflect restricted combinability in binary word combinations. The next chapter will investigate this suggestion in a corpus study.

[4]The resulting dataset HeiPLAS+ can be downloaded here: https://github.com/itsobvious/HeiPLASplus/

Chapter 8

Asymmetric AN combinations: a corpus study

As discussed in section 3.4.2, several statistical association measures have been proposed that take into consideration the directionality in co-occurrences. The most recent studies on asymmetric AMs introduced into collocation research the measure ΔP (formula 8.1) based on conditional probabilities (Ellis, 2006; Gries, 2013).

$$\Delta P = P(outcome|cue) - P(outcome|\neg cue) \tag{8.1}$$

The unidirectional measure ΔP has been compared to bidirectional ones in the task of automatic collocation extraction (Garcia et al., 2019). The results indicate that neither of the AMs can be considered the "best" one in this task. Moreover, the performance of ΔP is even lower than that of the bidirectional ones. However, poor performance in one particular task is hardly a valid reason for discarding the measure completely. One cannot rule out that a systematic study of this relatively new measure can provide insights into the nature of words' combinability that are valuable from lexical-semantic, lexicographic, and language learning and teaching perspectives.

Until now, this thesis only considered adjective-noun pairs that exhibit high mutual attraction which is based on the most popular and well studied statistical association measure logDice. The logDice proved to be effective and is now integrated into co-occurrence extraction tools such as Sketch Engine and the Wortprofil. This section will step away from this well-trodden path and turn to a less studied unidirectional association measure ΔP. The main interest of this study does not lie in examination of the measure's ability to detect or fail to detect a larger number of "true collocations" in comparison to other AMs. Rather, the investigation aims at identifying the strengths of ΔP. In other words: what types of co-occurrences make their way into the n-best lists based on ΔP? This section will provide a qualitative analysis of German adjective-noun combinations with high ΔP values.

8.1 Data

In order to calculate ΔP scores, co-occurrence frequencies have been computed for all the combinations of adjectives and nouns that are related to each other via the *attributive modifier* relation in a number of German tree banks: Wikipedia 2017 and Wikipedia 2018 (de Kok and Pütz, 2019) and decow16ax (Schäfer and Bildhauer, 2012; Schäfer, 2015).

	word2	**¬word2**	**total**
word1	a	b	a+b
¬word1	c	d	c+d
total	a+c	b+d	a+b+c+d=n

TABLE 8.1: Contingency table of observed frequencies. *a* is the frequency of a bigram, *a+c* is the frequency of the noun and *a+b* is the frequency of the adjective

In total, there are 42,445,060 pair types. The created data sample will be analysed manually which is highly time consuming. Out of this purely practical consideration, a minimal frequency cut-off is set: only the pairs with a frequency > 400 will be investigated. Apart from that, all the pairs containing digits and special characters have been removed. These steps reduced the size of the sample to 36,945 AN combinations. For these pairs, three scores are computed based on the notation in contingency table 8.1:

- unidirectional ΔP_1 (8.2): indication for right-predictive co-occurrences of the type $\overrightarrow{collocate\ base}$, e.g. $\overrightarrow{artisian\ well}$

$$\Delta P_1 = \Delta P(w_1|w_2) = \frac{a}{a+c} - \frac{b}{b+d} \tag{8.2}$$

- unidirectional ΔP_2 (8.3): indication for left-predictive co-occurrences of the type $\overleftarrow{collocate\ base}$, e.g. $\overleftarrow{good\ omen}$

$$\Delta P_2 = \Delta P(w_2|w_1) = \frac{a}{a+b} - \frac{c}{c+d} \tag{8.3}$$

- bidirectional logDice (8.4): does not distinguish between left-predictive and right-predictive co-occurrences, e.g. *strong feeling*

$$logDice = 14 + log_2 D = 14 + log_2 \frac{2a}{(a+c)+(a+b)} \tag{8.4}$$

The next section will explore two subsets of the data sample and present a comprehensive analysis of left- and right-predictive adjective-noun combinations. The results of the study will be shortly summarized in section 8.2.3.[1]

8.2 Asymmetric combinations: n-best lists

The traditional method in research on statistical association measures implies creating n-best lists of candidates by sorting them based on the calculated scores: the higher the score, the stronger the association between the two elements is. The goal of this exploratory study is not limited to determining which n-best list contains a greater number of true collocations. Rather, it aims at examining what types of multi-word expressions are found based on these measures and what challenges arise depending on the measure. Each list will be manually annotated using the categories described in Chapter 5: *proper name*, *free phrase*, *idiom*, *technical term*, and *collocation*.

The sample should be large enough to be able to make generalizations about asymmetric AN combinations. The frequency cut-off threshold of 400 is quite arbitrary. A similar approach has to be taken for setting the number of examples to investigate, n will be set to 600. Therefore, for each type of asymmetry, 600 top combinations based on the calculated scores will be considered.

8.2.1 Left-predictive co-occurrences

The first 600-best list is based on the scores of ΔP_1. The values of this asymmetric measure can vary between -1 and 1. The closer the value is to one, the stronger is the predictive power of the base word, in this case, a noun. Within the top-600 scores, the values of ΔP_1 lie between 0.5579 and 0.99999. Twelve phrases are excluded due to lemmatizer errors leaving 588 phrases for manual annotation.

The very first aspect that stands out is that over 80% (475 out of 589) of the base nouns in the list are compounds. This is not surprising: the more specific the compound is, the higher is the probability that it combines with a small number of adjectives.

The distribution of labels in the list is presented in table 8.2. The majority of the left-predictive combinations are proper names such as *Deutsche Umwelthilfe* 'Environmental Action Germany', *Olympische Sommerspiele* 'Summer Olympic Games', and *Europäisch Südsternwarte* 'European Southern Observatory'. Another well represented category comprises technical terms such as *frei Arztwahl* 'right to choose a physician' and *vorläufige Vollstreckbarkeit* 'preliminary enforceability'. The distinction between terms and free phrases in this list is quite vague due to the fact that the nouns are very specific compounds. For instance, in the phrase *lokales Kleingewerbe*

[1] Described data samples can be downloaded here: https://github.com/itsobvious/asymmetric-ADJ_NN.

Label	Number	Examples
proper name	240	*Olympische Sommerspiele* 'Summer Olympic Games', *Schwäbisch Gmünd*, *(der) Letzte Mohikaner* 'The Last of the Mohicans', *Hohe Tatra* 'High Tatras', *Vereinigtes Königreich* 'United Kingdom'
term	208	*vorläufige Vollstreckbarkeit* 'preliminary enforceability', *chronische Polyarthritis* 'chronical polyarthritis', *absoluter Nullpunkt* 'absolute zero', *Neue Sachlichkeit* 'new objectivity'
free phrase	65	*passender Staubsaugerbeutel* 'suitable vacuum cleaner bag', *neuer Teilnehmerrekord* 'new attendance record', *verschiedene Veranstaltungsformen* 'various event formats', *langer Blütenstiel* 'long flower stem', *früher Sonntagmorgen* 'early Sunday morning'
collocation	40	*reine Formsache* 'mere formality', *gesunder Menschenverstand* 'common sense', *stehende Ovation* 'standing ovation', *schwere Kopfverletzung* 'serious head injury'
idiom	28	*innerer Schweinehund* 'one's weaker self', *eierlegende Wollmilchsau* 'something that meets all requirements', *ewige Jagdgründe* 'happy hunting ground', *fliegende Untertasse* 'flying saucer'
cliché	6	*zu guter Letzt* 'last but not least', *gutes Gelingen* 'good luck', *herzlichen Glückwunsch* 'congratulations', *gute Besserung* 'get well soon'
compound	1	*rote Beete* 'beetroot'

Table 8.2: Left-predictive AN combinations: distribution of labels in the sample.

'local small business', the base itself might be considered a term. An indication for that is the existence of a Wikipedia article explaining what *Kleingewerbe* means in legal terms. Examples of free phrases in the list are *durchschnittliche Haushaltsgröße* 'average household size', *passender Staubsaugerbeutel* 'suitable vacuum cleaner bag', *angegebener Qualifikationspunkt* 'specified qualification point'. These phrases are not established technical terms, their base nouns are simply very specific compounds.

Over forty phrases were annotated as collocations. In some of them, the meaning of the base is literal whereas the meaning of the collocate is not literal, as in *schwere Kopfverletzung* 'serious head injury' (lit. 'heavy') and *fester Ensemblemitglied* 'permanent ensemble member' (lit. 'solid'). Apart from such typical examples of collocations, there are AN pairs that are highly lexicalized as illustrated in (43).

(43) a. The first example is *freie Wildbahn* which stands as a whole for 'wild' as in 'in the wild'. This is a part of the fixed phrase *in der freien Wildbahn*.
b. Similarly, *üble Nachrede* is a lexicalized phrase which can be translated as 'defamation, slander'.
c. The collocation *reine Formsache* 'mere formality' is also highly lexicalized.
d. The pair *letzte Ruhestätte* 'last resting place' is slightly more flexible than the two examples above. The noun allows modification by a few other adjectives, e.g. *würdig* 'worthy/dignified', *ewig* 'eternal', and *endgültig* 'final/ultimate'. However, the phrase is still highly asymmetric with a ΔP_1 value of 0.92.
e. Similarly, in the phrase *frühe Morgenstunde* 'early morning hour', a number of adjectival modifiers is acceptable: *gestrig* 'yesterday's', *heutig* 'today's', *kühl* 'cool'. However, the morpho-syntactic form is fixed: in most examples, the noun and, therefore, its adjectival modifier are used in plural form as in *frühe Morgenstunden* 'early morning hours', and alternative translations are 'small hours' and 'wee hours'. ΔP_1 value for this pair is 0.92.

Among highly asymmetric left-predictive phrases there are also idioms. Typical idioms are those in which the base noun has a figurative meaning, for instance *hässliches Entlein* 'ugly duckling' where *Entlein* refers to a person and not an actual duck. Two examples are combinations in which the base noun is a non-compositional compound such as *innere Schweinehund* 'weakness of will, laziness' which literally means 'inner dog-pig' and can be used as an insult. Another example is *eierlegende Wollmilchsau* 'something that meets all requirements' which literally means 'egg-laying pig that gives wool and milk'.

The remaining six AN pairs were first marked as unknown:

(44) a. the pair *gut+Letzt* is only used in the phrase *zu guter Letzt* which stands for 'last but not least'.
b. The pair *eigen+Bekunden* is only used in the phrase *nach eigenem Bekunden* which can be translated as 'by someone's own admission'.
c. The pair *eigen+Gutdünken* is only used in the phrase *nach eigenem Gutdünken* which can be translated as 'as someone sees fit'.
d. *gutes Gelingen* 'good luck', *herzlichen Glückwunsch* 'congratulations', *gute Besserung* 'get well soon' are phrases used in particular social situations.

Examples in (44) can be considered *clichés* in the sense of Mel'čuk (2012a): both lexically and semantically restricted phrases. Clichés are compositional. According to Mel'čuk, proverbs and complex proper names also belong to the category of clichés.

Summary. The exploratory study of the asymmetric statistic association measure ΔP_1 for AN pairs in German shows that this measure is capable of detecting various types of multi-word expressions. It is especially suitable for detecting proper names and technical terms. Free phrases can also be highly asymmetric when the base is a very specific nominal compound such as *Staubsaugerbeutel* 'vacuum cleaner bag'. The measure also detected a number instances of collocations, idioms, and clichés. Their number is quite low compared to terms and proper names, therefore, the measure is not very successful at extracting them. However, it would be very useful for language learners and lexicographers to know that a given AN pair is highly asymmetric and the use of the nominal head is extremely restricted.

8.2.2 Right-predictive co-occurrences

The second 600-best list to explore contains right-predictive AN pairs, i.e. those in which the use of the collocate is highly restricted, based on the measure ΔP_2. A high ΔP_2 value is an indication of asymmetry and, thus, of restrictedness of the modifier. The values in the list lie between 0.4131 and 0.9999. Note that the lower score is lower than that in the left-predictive list (0.5579) which means that left-predictive asymmetry is more common in German than the right-predictive one.[2] Apart from that, in contrast to left-predictive pairs, this list contains fewer compound nouns (65 in total). This is a reflection of the asymmetry: in right-predictive combinations, it is the adjective that is very specific.

During the annotation, 59 errors were identified including lemmatizer errors and typos, they will be excluded from further analysis. The manual annotation has been performed by the author of this thesis. The distribution of labels in the data sample is presented in table 8.3.
Roughly a half of the combinations were annotated as technical terms. There are 87 proper names and a small number of idioms in the data sample. The second large group of right-predictive AN combinations is represented by collocations.

Collocations.

There are various subgroups of collocations identified in the data sample and the lower the value of ΔP_2, the more vague is the distinction between a collocation and a free phrase because the amount of acceptable head nouns rises.

First of all, there is a number of asymmetric right-predictive collocations in which the adjective's use is highly restricted and the adjective is used in a non-literal sense. In such cases, the difference to a free phrase is clear and follows the logic underlying the decision tree in chapter 5 (Figure 7.1). Consider a few examples:

(45) a. *brenzlige Situation* 'tricky situation' (lit.'burnt'), $\Delta P_2 = 0.83$

[2] There are 1,057 AN phrases with ΔP_1 value > 0.413.

Label	Number	Examples
term	256	*molare Masse* 'molar mass', *rheumatoide Arthritis* 'rheumatoid arthritis', *festverzinsliche Wertpapiere* 'fixed-income securities', *zinsloses Darlehen* 'interest-free loan', *glykämischer Index* 'glycemic index'
collocation	182	*brenzlige Situation* 'tricky situation', *himmelweit Unterschied* 'stark difference', *klirrende Kälte* 'biting cold', *mulmiges Gefühl* 'queasy feeling', *faustdicke Überraschung* 'huge surprise'
proper name	87	*Timmendorfer Strand*, *(der) Gestiefelte Kater* 'Puss in Boots', *Brühlsche Terrasse* 'Brühl's Terrace', *Sixtinische Kapelle* 'Sistine Chapel', *Fruchtbringende Gesellschaft* 'Fruitbearing Society', *Das Wohltemperierte Klavier* 'The Well-Tempered Clavier'
idiom	15	*gordischer Knoten* 'Gordian knot', *zweischneidiges Schwert* 'double-edged sword', *trojanisches Pferd* 'Trojan horse', *(am) seidenen Faden (hängen)* '(to hang by a) thread'

TABLE 8.3: Right-predictive AN combinations: distribution of labels in the sample.

b. *faustdicke Überraschung* 'huge surprise' (lit.'thick as a fist'), ΔP_2 = 0.69
c. *klirrende Kälte* 'biting cold' (lit.'clanking'), ΔP_2 = 0.57
d. *klangvoller Name* 'famous name' (lit.'sonorous'), ΔP_2 = 0.54

In these collocations, the non-literal meaning of the modifier has become dominant. Moreover, the adjectives in these particular meanings have a strong preference for a very limited number of nouns.

Secondly, there is a large number of combinations in which the adjective retains its literal meaning, which is its only meaning according to the DWDS and GermaNet illustrated by example (46) and (48).

(46) a. *geraume Zeit* 'fairly long time', ΔP_2 = 0.98
b. *unverrichteter Dinge* 'empty-handed', ΔP_2 = 0.96
c. *unerfindliche Gründe* 'unknown reasons', ΔP_2 = 0.95
d. *helllichter Tag* 'broad daylight', ΔP_2 = 0.95
e. *himmelweiter Unterschied* 'stark difference', ΔP_2 = 0.94
f. *sehnlichster Wunsch* 'fondest wish', ΔP_2 = 0.89

In extremely asymmetric combinations (ΔP_2>0.8), the restrictedness is obvious: in each pair in example (46), only very few paraphrases are possible and none of them can compete with the given noun. In example (46e), the adjectival modifier *himmelweit* expresses intensification and can be translated into English as 'enormous'. Its combinability is extremely restricted: the adjective is mostly used with the noun *Unterschied*. According to the Wortprofil,[3] it is the only possible head noun. In the corpus underlying this study, the family size of *himmelweit* in attributive position is 21 and the distribution of frequencies is zipfian as there are only five examples with frequency higher than 1: *Ferne* 'distance' (2), *Niveauunterschied* 'difference of level' (2), *Lücke* 'gap' (2), *Abstand* 'distance' (9), *Unterschied* 'difference' (620). Mel'čuk (2023, p. 46) terms such LUs ***unilexemes***, i.e. LUs that appear "*only in one particular collocation (or in a handful of collocations)*".

Some of such highly asymmetric right-predictive AN pairs are extremely restricted and have a fixed morpho-syntactic form. Consider example (46b): the phrase *unverrichteter Dinge* is only used adverbially in genitive case of its plural form as illustrated in example (47).[4]

(47) "Die Fahrer mussten
the-NOM.PL driver-NOM.PL must-PST-PL
unverrichteter Dinge abreisen.
without-achieving-GEN.PL thing-GEN.PL leave-INF
'The drivers had to leave empty-handed'

[3] <https://www.dwds.de/wp/himmelweit>, *last accessed on 14.06.2023*

[4] In *Berliner Zeitung, 10.05.2003*, from the corpus Berliner Zeitung of the Digitales Wörterbuch der deutschen Sprache, <https://www.dwds.de/d/korpora/bz>, retrieved on 14.06.2023.

Another type of highly restricted AN combinations is shown in (46f). The adjectival modifier is used in superlative form of the adjective *sehnlich* 'fervent'. The preference for the superlative form is also reflected in raw corpus frequencies: positive form *sehnlicher Wunsch* – 125; superlative form *sehnlichster Wunsch* – 575. The DWDS entry for *sehnlich*[5] also gives as example the superlative form of the adjective in combination with the noun *Wunsch*.

Now consider a few collocations with less restricted adjectives.

(48) a. *missliche Lage* 'difficult situation', $\Delta P_2 = 0.68$
b. *ungebetener Gast* 'uninvited guest', $\Delta P_2 = 0.54$
c. *landläufige Meinung* 'general opinion', $\Delta P_2 = 0.45$
d. *erschwinglicher Preis* 'reasonable price', $\Delta P_2 = 0.44$

The lower the ΔP_2 score, the easier it is to substitute the base by a different, but semantically related noun. If for the adjective *misslich*, the Wortprofil provides the total of 28 possible nominal bases, for *erschwinglich*, 50 options are listed (see Appendix G). The more options there are, the closer is a combination to a free phrase rather than a collocation. The boundary between a collocation and a free phrase for AN pairs in which the adjective retains its literal meaning can be drawn on a purely quantitative basis by defining *restrictedness* in terms of ΔP_2 value. In this study, the cut-off value has been set to 0.4, therefore, all AN combinations that make it to the list and that are not technical terms, proper names, or idioms should be considered collocations.

However, there is still one large group of phrases that could not be assigned to any of these categories. During annotation they were labeled *unknown*. Following the logic described in the previous paragraph, these phrases should be considered collocations. They can be roughly divided into three groups:

1. **Participles**
   - **Participle I**: *schallendes Gelächter* 'roaring laughter', *aufgehende Sonne* 'rising sun', *vertrauensbildend Maßnahme* 'trust-building activity', *kochendes Wasser* 'boiling water'.
   - **Participle II**: *rausgeschmissenes Geld* 'wasted money', *vorgefasste Meinung* 'preconception', *durchzechte Nacht* 'drunken night', *angeleint Hund* 'dog on a leash', *gezapftes Bier* 'drought beer'.
2. **Pertainyms**, e.g. *hellseherische Fähigkeit* 'clairvoyant ability', *alkoholischer Getränk* 'alcoholic drink', *tatkräftige Unterstützung* 'active support', *leibliches Wohl* 'physical well-being'.
3. **Compounds**, e.g. *pflegebedürftiger Mensch* 'person in need of care', *krebskrankes Kind* 'child suffering from cancer'.

[5] https://www.dwds.de/wb/sehnlich, *last accessed on 17.06.2023.*

The first and the largest group consists of pairs with a present (I) or past (II) participle as modifier. As discussed in section 2.2, both types of participles are considered adjectives in German when used attributively. Consider a few examples. Present participle *aufgehend* 'rising' is derived from the verb *aufgehen* which is highly polysemous. The verb has 9 senses in GermaNet and 14 senses in the DWDS; just to name a few: 'to tally', 'to rise', 'to sprout', 'to open'. Its combinability with nouns as subjects is thus highly varied: *Rechnung* 'invoice', *Kalkül* 'calculation', *Sonne* 'sun', *Vorhang* 'curtain', *Saat* 'seed', *Tür* 'door'. However, the derived present participle *aufgehend* is much more selective than the verb itself and has a strong preference[6] for the noun *Sonne* 'sun'. A similar observation holds for the participle *kochend* 'boiling'. The verb it is derived from, *kochen*, can be used transitively by taking a human being as subject (*Koch* 'cook', *Oma* 'grandma') in the meaning 'to cook' or intransitively, and take food nouns such as *Suppe* 'soup' as subject in the meaning 'to boil'. The derived present participle *kochend* exhibits a preference[7] for the noun *Wasser* 'water' and, thus, intransitive use.

Past participles exhibit similar behavior. For instance, the past participle *gezapft* 'tapped': compared to the participle, the verb it is derived from (*zapfen*) combines with a much larger variety of direct objects: *Bier* 'beer', *Sprit* 'fuel', *Benzin* 'petrol' *Strom* 'electricity', *Wasser* 'water'.[8] A shift in meaning is also found in participles, for instance, in the collocation *rausgeschmissenes Geld* 'wasted money' where the participle is derived from the verb *rausschmeißen* whose literal meaning is 'to throw away'. When this verb takes *Geld* 'money' as its direct object, the meaning shifts to a less literal 'to waste'. The derived participle is used only in the non-literal meaning 'wasted' with *Geld* 'money' as its head.[9]

The second group is much smaller, ten combinations in total. The adjectival modifiers are pertainyms, i.e. they are derived from nouns and retain the meaning of the noun. For instance, *alkoholischer Getränk* 'alcoholic drink' is a *drink* that contains *alcohol*. Pertainyms were so far excluded from study of AN combinations in this thesis because the main focus lies on descriptive adjectives. Relational adjectives such as *barock* 'baroque' that were discussed in previous chapters (section 5.2 and 6.1.3) are different from the examples given here in terms of their combinability. There is a large variety of possible heads for *barock*,[10] but the pertainyms such as *alkoholisch* 'alcoholic' and *hellseherisch* 'clairvoyant' are extremely selective and their combinability is restricted. This selectiveness should be recognized.

Finally, there is a number of compound adjectives that made it to the

[6] https://www.dwds.de/wp/aufgehend, *last accessed on 26.06.2023.*
[7] https://www.dwds.de/wp/kochend, *last accessed on 26.06.2023.*
[8] https://www.dwds.de/wp/gezapft, *last accessed on 26.06.2023.*
[9] https://www.dwds.de/wp/rausgeschmissen, *last accessed on 26.07.2023.*
[10] https://www.dwds.de/wp/barock, *last accessed on 26.06.2023.*

n-best list that belong to the semantic field KÖRPER 'body': *schwerbehinderter Mensch* 'severely handicapped person', *krebskrankes Kind* 'child suffering from cancer', *pflegebedürftiger Mensch* 'person in need of care', *demenzkranker Mensch* 'person with dementia'. They have a very limited combinability: only combination with nouns denoting living beings is possible for such adjectives. However, it is unclear whether this kind of restrictedness can be compared to any discussed before. The constraints are predictable: any living being can suffer from cancer or be in need of care. The intuition here is that such phrases are not collocations, but rather free phrases although the ΔP_2 value is quite high. For consistency sake, this group should also be labeled as *collocation*.

8.2.3 Conclusion

This chapter investigated unidirectional statistical association measure ΔP by applying it to German adjective-noun combinations. There are two types of asymmetric combinations: *left-predictive* in which the noun acts as predictor of its modifier and *right-predictive* in which the modifier is the predictor of its head. The first type is to a large extent represented by proper names and technical terms, whereas the second one is typical for technical terms and collocations.

Bedeutung eWDG

sehr, unendlich weit

BEISPIELE:

das ist ein **himmelweiter** *(= sehr großer, gewaltiger)* Unterschied

ich sah mich plötzlich **himmelweit** von dem Ziele meiner Hoffnungen gerückt [↗ G. KELLER, *Gr. Heinrich*, 4,273]

Zwei Welten, zwei **himmelweit** verschiedene, scheinbar unüberbrückbare Welten! [↗ NABL, *Ortliebsche Frauen*, 248]

Thesaurus www.openthesaurus.de (11/2022)

Synonymgruppe

↗enorm · ↗exorbitant · ↗extrem · ↗gewaltig · ↗groß · himmelweit (eingeschränkter Gebrauch) · ↗immens · ↗imposant · ↗mächtig · ↗riesig • ↗krass ugs.

Synonymgruppe

himmelweit (eingeschränkter Gebrauch) • sehr weit Hauptform · ↗kilometerweit ugs. · ↗meilenweit ugs.

FIGURE 8.1: DWDS-entry for the adjective *himmelweit*.

Both measures provide additional information about AN combinations and their restrictedness which should be reflected in their dictionary entries and statistical sketches. When it comes to dictionary entries, it is often the case that typical combinations for a given lexical unit are listed. However, it is never stated explicitly which combinations are highly restricted. This is especially important for left-predictive combinations. Consider the pair

hat Adjektivattribut		logDice	Freq.
1. regional	M W A	9.0	4778
2. fein	M W A	8.7	3697
3. gravierend	M W A	8.5	2142
4. wesentlich	M W A	8.4	5269
5. kulturell	M W A	8.3	3676
6. signifikant	M W A	8.2	1249
7. fundamental	M W A	8.1	1304
8. entscheidend	M W A	8.0	4125
9. erheblich	M W A	7.8	3492
10. grosse	M W A	7.5	1571
11. gewaltig	M W A	7.5	1361
12. qualitativ	M W A	7.5	834
13. grundlegend	M W A	7.4	1200
14. genetisch	M W A	7.3	847
15. markant	M W A	7.2	734
16. geschlechtsspezifisch	M W A	7.1	496
17. beträchtlich	M W A	7.0	717
18. deutlich	M W A	6.9	4885
19. kraß	M W A	6.8	552
20. himmelweit	M W A	6.8	418

FIGURE 8.2: Top-20 adjectival collocates for the noun *Unterschied*

himmelweiter Unterschied 'enormous difference': the adjective's entry in the DWDS[11] is presented in figure 8.1

The definition of the adjective's only sense is 'very, endlessly broad' and an only example in which it is used in an attributive function is with the noun *Unterschied*. It does not explicitly say that it is actually the only possible nominal head for it. However, this information is reflected in the "Thesaurus" part of the entry by marking it as having a restricted use (Ger. *eingeschränkter Gebrauch*). Now consider the Wortprofil[12] for the head noun *Unterschied* presented in figure 8.2.

All the possible modifiers can be sorted by the logDice measure (as presented in the figure) or by raw frequency. However, as discussed above, logDice is a bidirectional measure and only indicates that there is mutual attraction between two elements. It might be useful, especially for language learners,

[11] https://www.dwds.de/wb/himmelweit, *last accessed on 02.07.2023.*

[12] https://www.dwds.de/wp/Unterschied, *last accessed on 02.07.2023.*

to include the unidirectional information into the word sketch as well. According to the sketch based only on logDice and raw frequency, there is no way to know that the use of *himmelweit* is much more restricted than the use of the adjective *krass* that has the same logDice value and almost the same raw frequency.

It is not clear what exact ΔP value is *high* enough for a phrase to be considered *restricted*. Obviously, negative values indicate that the presence of the cue, in this particular case, the presence of an adjective reduces the chance of its co-occurrence with the given head. Proximity to 0 shows that the presence of an adjective does not make any difference. ΔP cannot and should not substitute logDice. Both measures are highly useful additional sources of information about the combinability of words.

Correlation with logDice. The correlation ΔP_1 and ΔP_2 with logDice was first computed for the full sample of 36,945 AN combinations. The visualizations are presented in Figures 8.3 and 8.4. For both measures, Pearson and Spearman correlation scores indicate a positive correlation. Spearman correlation is much higher, especially for ΔP_2, which suggests that the relation between logDice and ΔP is not linear, but monotonic. In order to investigate this more closely, the computation was repeated only for the two top-600 lists. The results are visualized in Figures 8.5 and 8.6. Both Spearman and Pearson scores show no correlation between the directional and bi-directional measures. This suggests that after a certain threshold, asymmetric measures such as ΔP_1 and ΔP_2 manage to detect co-occurrences that have a strong directional association that is missed by the logDice measure.

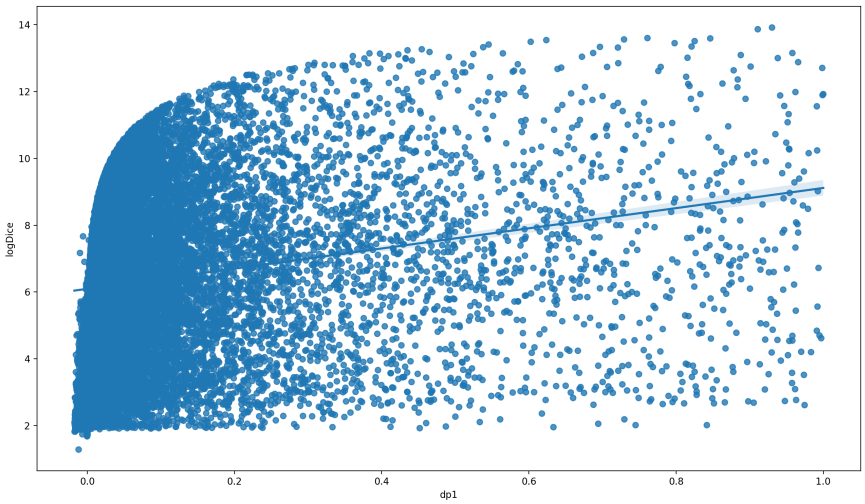


FIGURE 8.3: Correlation between $\Delta P1$ and logDice scores in the AN combinations with frequency > 400. Pearson=0.184, Spearman=0.376

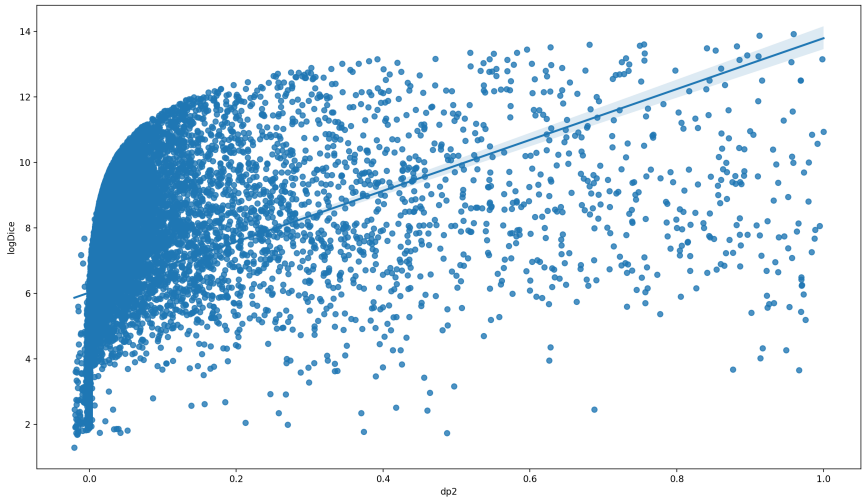


FIGURE 8.4: Correlation between $\Delta P2$ and logDice scores in the AN combinations with frequency > 400. Pearson=0.373, Spearman=0.877

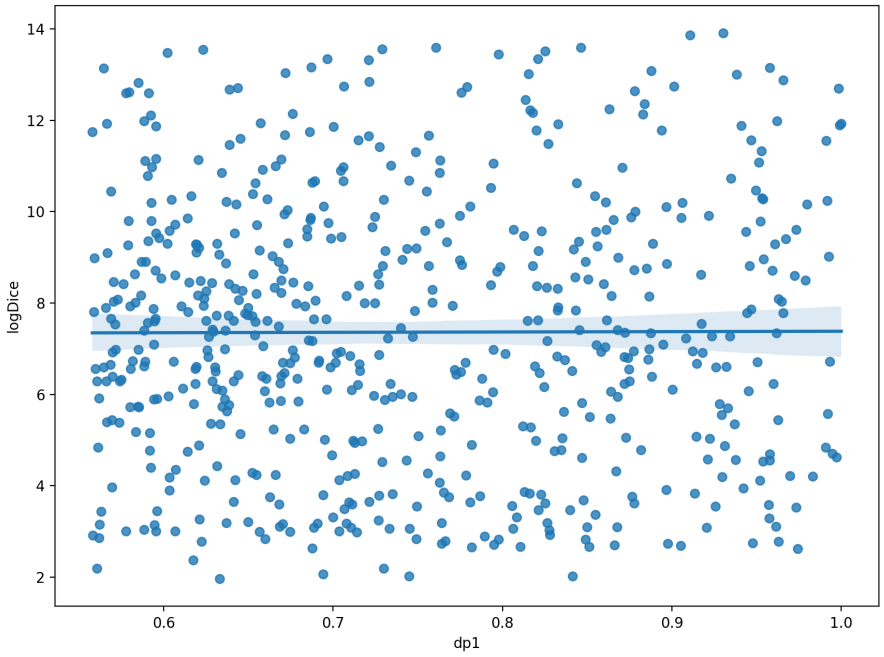


FIGURE 8.5: Correlation between $\Delta P1$ and logDice scores in the 600-best list of left-predictive AN combinations. Pearson=0.003, Spearman=-0.01.

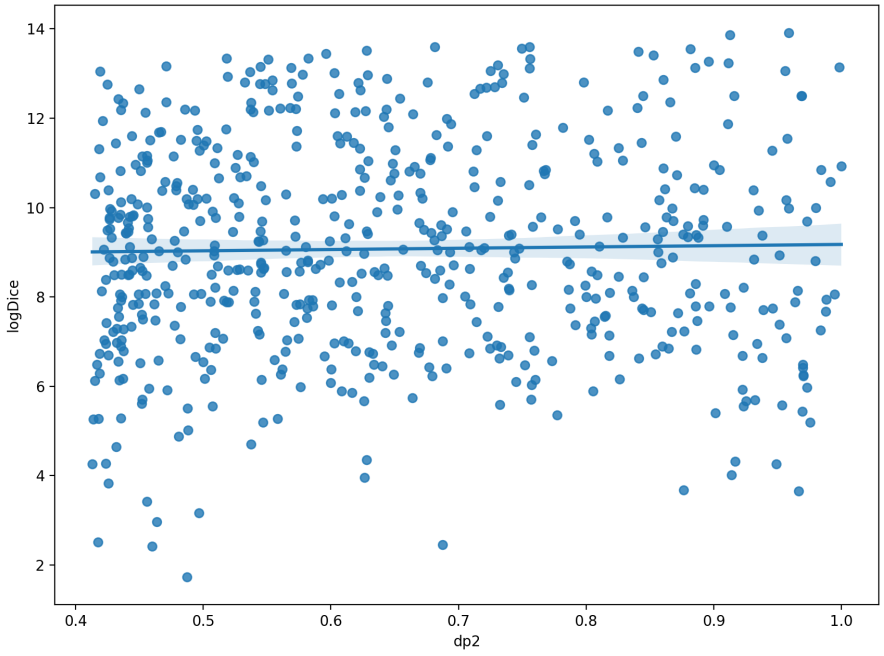


FIGURE 8.6: Correlation between $\Delta P2$ and logDice scores in the 600-best list of right-predictive AN combinations. Pearson=0.021, Spearman=0.023.

8.3 Attributes in asymmetric AN combinations

A sample of free phrases and collocations resulting from the study presented in this chapter provides a perfect opportunity to challenge the annotation scheme of attributes suggested in this thesis. It may be especially informative because this sample also contains participles (present and simple) and pertainyms.

All the collocations and free phrases from both 600-best lists are annotated by the author of this thesis relying on the annotation guidelines and definitions of attributes provided in Appendices C and E. The final subsample contains 220 collocations and 66 free phrases. In total, 52 labels were used to annotate those AN combinations. Table 8.4 presents the distribution of labels and examples from the sample illustrating each attribute.

attribute	num	examples
time	31	***jähes Ende*** 'sudden end', *früher Sonntagmorgen* 'early sunday morning'
intensity	29	***drakonische Strafe*** 'draconian punishment', ***gähnende Leere*** 'gaping emptiness'
evaluation	21	***klangvoller Name*** 'resounding name', *gute Regierungsführung* 'good governance'
social	20	***freie Platzwahl*** 'free seating', *europäische Ethnologie* 'european ethnology'
activity-†	17	***gehütetes Geheimnis*** 'guarded secret', *angegebener Qualifikationspunkt* 'stated qualification point'
manner	12	***strömender Regen*** 'pouring rain', *friedliche Koexistenz* 'peaceful coexistence'
relation	11	***untergeordnete Rolle*** 'minor role', *verschiedene Veranstaltungsform* 'various event types'
cause	10	***aufgeworfene Frage*** 'question raised', ***ungebetener Gast*** 'uninvited guest'
function	10	***lobendes Wort*** 'praising word', ***rausgeschmissenes Geld*** 'wasted money'
other	10	***atomwaffenfreie Welt*** 'world free of nuclear weapons', *vierthöchste Spielklasse* 'fourth highest league'
activity†	10	***aufgehende Sonne*** 'rising sun', ***knurrender Magen*** 'growling stomach'

Continued on the next page

Table 8.4: Continued from previous page

attribute	number	examples
number	8	***erster Profivertrag*** 'first professional contract', *zweite Jahreshälfte* 'second half of the year'
type	8	***darstellende Kunst*** 'performing art', ***alkoholisches Getränk*** 'alcoholic beverage'
epithet	7	***freie Wildbahn*** '(in the) wild', ***weite Ferne*** 'far distance'
health	6	*demenzkranker Mensch* 'person with dementia', *behinderter Mensch* 'disabled person'
quantity	5	***geraume Zeit*** 'fairly long time', *durchschnittliches Jahreseinkommen* 'average annual income'
quality	5	***gesunder Menschenverstand*** 'common sense', *passender Staubsaugerbeutel* 'fitting vacuum cleaner bag'
color	4	***dunkelblondes Haar*** 'dark blond hair', *grüner Klee* 'green clover'
dimension	4	***mundgerechtes Stück*** 'bite-sized piece', *langer Blütenstiel* 'long flower stem'
emotion	4	***brenzlige Situation*** 'hairy situation', ***rabenschwarzer Tag*** 'dismal day'
has-topic*	4	***gymnastische Übung*** 'gymnastic exercise', ***musikalische Früherziehung*** 'early musical education'
importance	4	***integraler Bestandteil*** 'integral element', *wesentliche Vertragspflicht* 'essential contractual obligation'
location	4	*linke Brustseite* 'left chest', *rechte Maustaste* 'right mouse button'
climate	4	*bewölkter Himmel* 'cloudy sky', *trockenster Monat* 'driest month'
certainty	4	***irreparable Schaden*** 'irreparable damage', ***unkalkulierbares Risiko*** 'incalculable risk'
appearance	3	***gebückte Haltung*** 'stooped posture', ***schmerzverzerrtes Gesicht*** 'pain distorted face'
completeness	3	***volle Montur*** 'full gear', *gesamtes Bundesgebiet* 'entire federal terrotory'
duration	3	***ewige Verdammnis*** 'eternal damnation', ***schnellste Rennrunde*** 'fastest lap'

Continued on the next page

Table 8.4: Continued from previous page

attribute	number	examples
has-manner-of-functioning*	2	*elektrischer Fensterheber* 'electric window regulor', ***bildliche Darstellung*** 'visual representation'
stability	2	***festes Ensemblemitglied*** 'permanent ensemble member', ***wechselvolle Geschichte*** 'chequered history'
validity	2	***strafbare Handlung*** 'criminal offence', *zulässiges Gesamtgewicht* 'permissible total weight'
effort	2	***unwegsames Gelände***
material	1	***feinsandiger Strand*** 'fine sand beach'
age	1	*junger Nationaldemokrat* 'young national democrat'
brightness	1	***hellichter Tag*** 'bright day'
consistency	1	***zerlassene Butter*** 'melted butter'
constitution	1	***ausdauernde Pflanze*** 'enduring plant'
has-no-part*†	1	***schnurloses Telefon*** 'cordless phone'
naturalness	1	***streunender Hund*** 'stray dog'
sex	1	*weiblicher Haushaltsvorstand* 'female head of household'
shape	1	***lockiges Haar*** 'curly hair'
sound	1	***schallendes Gelächter*** 'resounding laughter'
has-specialization*	1	***hellseherische Fähigkeit*** 'clairvoyant ability'
speed	1	***kometenhafter Aufstieg*** 'meteoric rise'
temperature	1	***lauwarmes Wasser*** 'lukewarm water'

TABLE 8.4: Attributes in asymmetric AN phrases ("attribute"), number of instances of each attribute in the sample ("num"), examples from the data sample ("examples'): collocations are depicted in bold. The symbol † marks new suggested relations. The symbol ⋆ marks relations borrowed from Dima (2019) originally suggested for NN compounds.

The majority of labels stem from the inventory presented in Appendix E. As the sample contains pertainyms and participles, a few new relations had to be introduced.

Pertainyms. As discussed in previous chapter (see section 6.2.3), some denominal adjectives are not listed in GermaNet under the semantic class PERTAINYM, but under a different suitable semantic class. It was exemplified by the attribute material for the adjective *hölzern* 'wooden'. The new sample provides further examples: *leidvolle Erfahrung* 'painful experience' derived from the noun *Leid* 'suffering' is listed under emotion;

katholische Soziallehre 'catholic social doctrine' – under the semantic class SOCIAL. In such cases, the phrases in the sample were annotated with according attributes. However, there are also examples of pertainyms that could not be assigned to an adjectival semantic class in GermaNet and are directly linked to the nouns they are derived from: e.g. *elektrisch* 'electric' is linked to *Elektrizität* 'electricity' or *musikalisch* 'musical' is linked to *Musik* 'music'. The inventory of relations suggested in this thesis can easily accommodate descriptive adjectives that denote properties, but most pertainyms do not denote properties but rather objects. Section 2.1 suggested that inventories of relations in noun-noun compounds might be applied in such cases. This question is outside of the scope of this thesis. Nevertheless, it would be a wasted opportunity not to try that approach for a small number of AN phrases with pertainyms as modifiers in the sample. The inventory of NN compounds suggested in Dima (2019) is considered and the following relations are borrowed: `has-specialization`, `has-topic`, `has-manner-of-functioning`, `has-no-property`. Example (49) illustrates how such relations are interpreted.

(49) a. *elektrischer Fensterheber* 'electric window regulator': *elektrischer Fensterheber* `has-manner-of-functioning` *Elektrizität*
b. *musikalische Früherziehung* 'early musical education': *musikalische Früherziehung* `has-topic` *Musik*
c. *schnurloses Telefon* 'cordless phone': *schnurloses Telefon* `has-no-part` *Schnur*

Note that the last example (49c) is different from the first two as the adjective actually denotes an absence of a part: a telephone **without** a cord. The inventory in Dima (2019) only has the relation `has-part`. Most likely, the opposite meaning is rarely found in NN compounds and is expressed in German by adjectives termed *privative* in GermaNet, i.e. "denominal adjectives defined by negation".[13] Although the compound *Schnurlostelefon* does exist, its frequency is lower than that of the corresponding adjective according to the DWDS frequency scale. An alternative interpretation could be the relation `manner-of-functioning`: *schnurloses Telefon* `manner-of-functioning` *Schnur*.

Participles. A number of combinations with participles as modifiers were annotated with labels that were underrepresented in the GerCoAt dataset described in chapter 5, namely, the attributes `function` as in example (50a) and `cause` as in example (50b). Already in chapter 4 (section 4.3.2), it has been noted that these meanings are rarely expressed by descriptive adjectives. However, they were kept in the inventory and were successful at accommodating participles.

(50) a. *aufgeworfene Frage* 'question raised': `cause`
b. *rausgeschmissenes Geld* 'wasted money': `function`

[13] `https://uni-tuebingen.de/en/142852`, *last accessed on 12.09.2023.*

A group of participles required introducing new relations. Present participles, as in example (51a), denote an activity performed by their head noun (head noun is a subject): *stomach growls*. Past participles, as in example (51b), denote an activity done to their head noun (head noun is a direct object): *someone guards the secret*. In the first case, relation `activity` was introduced; in the second, the relation `activity-` is used, where the symbol "-" marks that the roles are reversed.

(51) a. *knurrender Magen* 'growling stomach'
b. *gehütetes Geheimnis* 'guarded secret'

A more fine-grained distinction is needed for such examples which should stem from an inventory developed for verb-noun combinations. For instance, the relation in *growling stomach* can be described in terms of the standard Lexical Function `Son` "X emits its characteristic sound" (Mel'čuk, 2015, p. 232). Relations in verb-noun combinations are outside of the scope of this thesis; however, this topic will be shortly discussed in the final chapter (section 9.2).

Summary. This small scale study of attributes in asymmetric AN combinations has shown that the inventory of attributes compiled in this thesis is robust and capable of accommodating examples from various semantic classes. It has also shown a few limitations, more specifically, that relational and participial adjectives require a different approach. The next chapter will continue this discussion in a more general manner by outlining possible research directions that could tackle the identified gaps.

Chapter 9

Conclusions

9.1 Discussion and conclusions

This thesis has focused on adjective-noun collocations and free phrases and their semantic properties, more specifically, on lexical-semantic relations that hold between a head noun and its adjectival modifier.

Classification of AN co-occurrences. As the concept of collocation is extremely fuzzy and collocations are treated very differently in literature, the statistical and semantic criteria for differentiating between adjective-noun collocations, free phrases, idioms, technical terms, proper names, and clichés were defined. Only AN pairs that have high statistical association scores between their elements, in other words, *co-occurrences*, were considered. Guidelines for manual annotation of AN co-occurrences presented in a form of a simple decision tree were developed. A high inter-annotator agreement score shows that the proposed annotation scheme is reliable, and the task is feasible for human annotators. Several iterations of annotation were performed before the guidelines are finalized as presented in Figure H.1 (see Appendix H for more details and examples). The first iteration is presented in Chapter 5, it was a joint effort of Isabel Fuhrmann, Daniela Rossmann, Dr. Alexander Geyken, Dr. Lothar Lemnitzer, Prof. Dr. Erhard Hinrichs, and the author of this thesis. It resulted in compiling a large dataset of German free phrases and collocations GerCo. The second iteration was performed on a different data (the dataset of English AN pairs HeiPLAS) as described in Chapter 7 by Richard Lawrence, PhD and the author of this thesis. Finally, the third iteration presented in Chapter 8 on yet another data sample (asymmetric AN pairs) was performed solely by the author of this thesis.

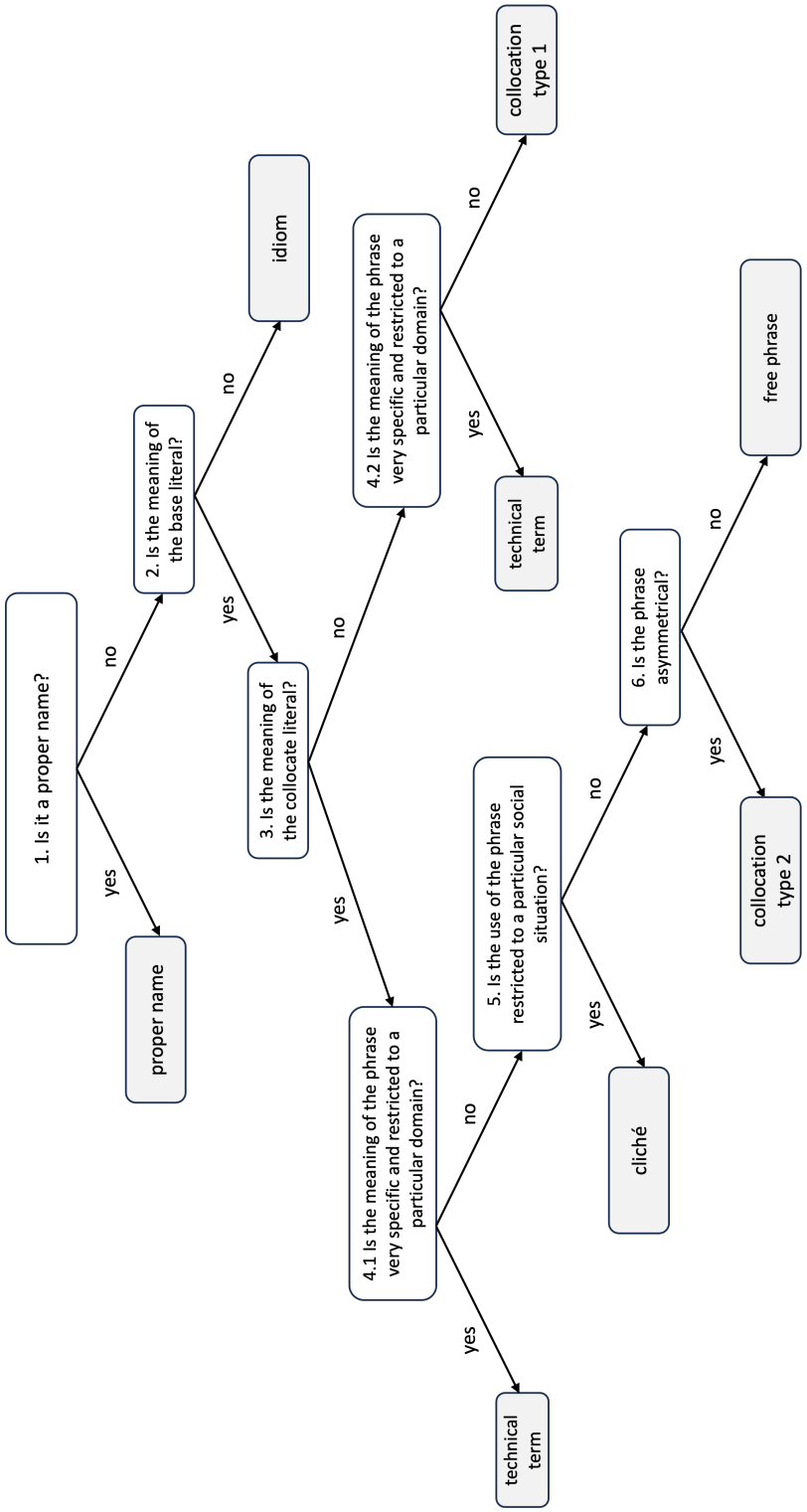


FIGURE 9.1: Finalized decision tree for classifying AN co-occurrences.

Inventory of lexical-semantic relations. A major effort in this thesis was dedicated to describing and classifying the lexical-semantic relations in AN collocations and free phrases. Four theoretical frameworks were compared in a small-scale case study described in Chapter 4. Based on the study's findings, an initial inventory of relations was compiled that takes into account two criteria: semantic coverage and granularity. Such relations are referred to as *attributes* in this thesis. Attributes are abstractions over specific properties of nouns that are usually expressed by adjectives (Hartung et al., 2017). For instance, the attribute `color` is an abstraction over the adjectives *blue, green, white, etc.*. The initial inventory of attributes was largely borrowed from the German lexical-semantic net GermaNet and applied to the constructed collection of collocations and free phrases GerCo. The annotation was performed by two students of linguistics Daniela Rossmann and Eva Huber. Based on the disagreement cases between the annotators, the scheme was further adapted by the author of this thesis with the input from Neele Falk and Prof. Dr. Erhard Hinrichs.

The semantically annotated dataset GerCoAt allows to draw a comparison between free phrases and collocations in terms of semantic relations that hold between their constituents. Application of the developed annotation scheme to a different set of data (HeiPLAS+) performed by the author showed the need to add a few more categories. The final inventory consisting of 54 attributes presented in Appendix E provides a broad semantic coverage and takes into account both concrete and abstract nouns. According to Mel'čuk (1999), only the meanings that are frequently expressed arbitrarily are worth accounting for and regular meanings found predominantly in free phrases are of no interest. This thesis shows that any lexical-semantic relation from the inventory can be found in collocations. There are, indeed, certain tendencies: the meaning `intensity` is often expressed collocationally whereas the meaning of, say, `color` is mostly expressed in free phrases. However, there are also examples that do not reflect these tendencies. The distribution of collocations across attributes in GerCoAt and HeiPLAS+ shows that collocations are ubiquitous, and arbitrary lexical preferences exist for nouns from various semantic fields. For instance, in the collocation *kohlrabenschwarzes Haar* 'jet-black hair', the attribute is `color`.

Attributes and Lexical Functions. The compiled inventory of attributes has only partial overlaps with the theoretical frameworks compared in Chapter 4. Consider the Meaning-Text Theory and the inventory of LFs first. Note that the compiled inventory does not take into consideration polarity that is expressed in the MTT by the LF `Anti`. However, it can be added to the inventory if required. Figure 9.2 presents an intuition on the overlaps beween standard LFs and attributes from the compiled inventory. There are two direct mappings. The label `intensity` is largely inspired by and fully corresponds to the standard LF `Magn`. The second category is `epithet` that has been completely borrowed from the LF inventory. That leaves three standard LFs: `Bon`, `Pos`, and `Ver`.

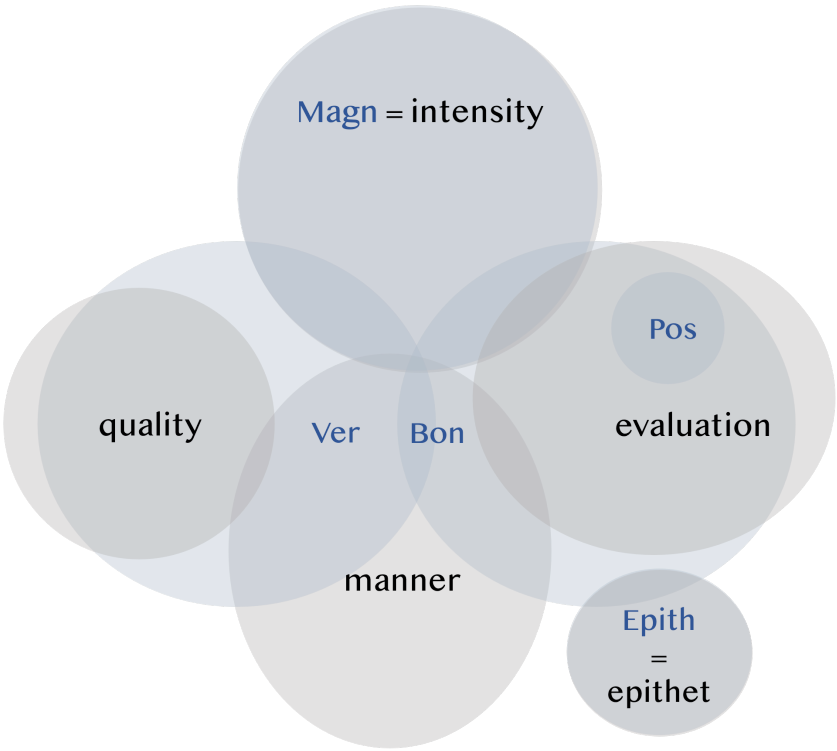


FIGURE 9.2: Intuition for mapping of the meanings covered by Standard Lexical Functions (in blue font) to the attributes from the compiled inventory.

The LF `Bon` overlaps with the attribute `evaluation` only in clear cases of positive evaluation, as in *wonderful weather*; negative evaluation as in *horrible weather* would be described by the complex LF `AntiBon`. However, also phrases such as *playful mood* and *shabby clothes* are annotated as `Bon` and `AntiBon` respectively in the `LexFunc` dataset (Espinosa-Anke et al., 2019). The approach taken in this thesis does not license seeing them as instances of `evaluation`, but rather of `description`. Hence, the LF `Bon` only partially overlaps with the attributes `description` and `evaluation` from the inventory. The highly specific LF `Pos` is fully mapped to `evaluation`.

The objective qualifier `Ver` and its antonym `AntiVer` defined in Mel'čuk (2015) as "as it should (not) be", correspond to a large extent to the label `quality`. For instance, this holds for the following examples from HeiPLAS+: *true believer*, *false testimony/accusation*. The two adjectives carry the meaning of '(not) being in accordance with fact or reality'. Compare them to a few examples of `Ver` from `LexFunc`: *disarming smile*, *shrewd businessman*, or *sharp tool*. Judgements of the type 'a smile *should be* disarming', 'a businessman *should be* shrewd' or 'a tool *should be* sharp' are very subjective. The approach in this thesis tries to avoid this subjectivity and classifies the three examples as follows in terms of attributes:

(52) a. *disarming smile* - `manner`,
b. *shrewd businessman* - `intelligence`,
c. *sharp tool* - `shape`

The analysis presented above does not support the conclusion made earlier in section 4.6 that LFs are extremely fine-grained and are, thus, difficult to distinguish from each other. It is true that, in some cases, the distinction between, say, `Bon` and `Ver` is fuzzy because they express very specific meanings, but there is more to it. The confusion has also to do with meaning partitioning. As examples (52a-c) show, the LF `Ver` combines different meaning: e.g. `shape, intelligence, manner`. However, it cannot be seen as a coarse-grained generalization over these attributes. In fact, it can refer to any property of a noun that the speaker would find relevant for describing how something *should (not) be*. Note that this does not hold true for examples such as *wrong answer* or *true believer* where the adjectives indeed express the meaning of how something should be according to certain rules or criteria. This explains the large overlap of the LF `Ver` with the attribute `quality`. The same holds for the LF `Bon`: the major overlap is with the attribute `evaluation` and other less obvious cases of evaluation such as *playful mood* are accommodated by other categories.

As for the non-standard LFs, they are extremely fine-grained as they are expressed in form of precise definitions. They may overlap with any of the attributes with the exception of those that fully overlap with standard LFs: `intensity` and `epithet`.

Attributes and qualia roles. Now consider the qualia roles from the theory of the Generative Lexicon. Figure 9.3 shows an intuition about the alignment between the four qualia dimensions and the attributes. Attributes referring to physical properties of entities such as `color` and `shape` can be seen as specifications of the FORMAL qualia dimension (Pustejovsky, 1991). As figure 9.3 illustrates, there is a large number of labels that can be aligned with the FORMAL role. In contrast, the CONSTITUTIVE and TELIC dimensions are represented only by three attributes each, and the AGENTIVE dimension has only one. Considering the wide lexical variety of the studied datasets, the intuition is that the pattern ATTRIBUTIVE ADJECTIVE + NOUN is not the typical way of expressing the three qualia dimensions.

Similarly, Sorokin et al. (2015) show that relations in nominal compounds map well to the qualia roles, with FORMAL role having the largest amount of specifications whereas the TELIC and AGENTIVE dimensions remain underrepresented. It was already noted in Chapters 4 (section 4.3.2) and 8 (section 8.3) that the meanings of `function` which belongs to the telic qualia dimension (T) and of `origin` which belongs to the agentive dimension (A) are mostly activities. These meanings are more likely to be expressed by verbs as examples (53a-d) illustrate, or by participles as was illustrated in example (50) in the previous chapter.

(53)
a. A: to **bake** a cake — T: to **eat** a cake
b. A: to **build** a house — T: to **live** in the house
c. A: to **write** a book — T: to **read** a book
d. A: to **brew** beer — T: to **drink** beer

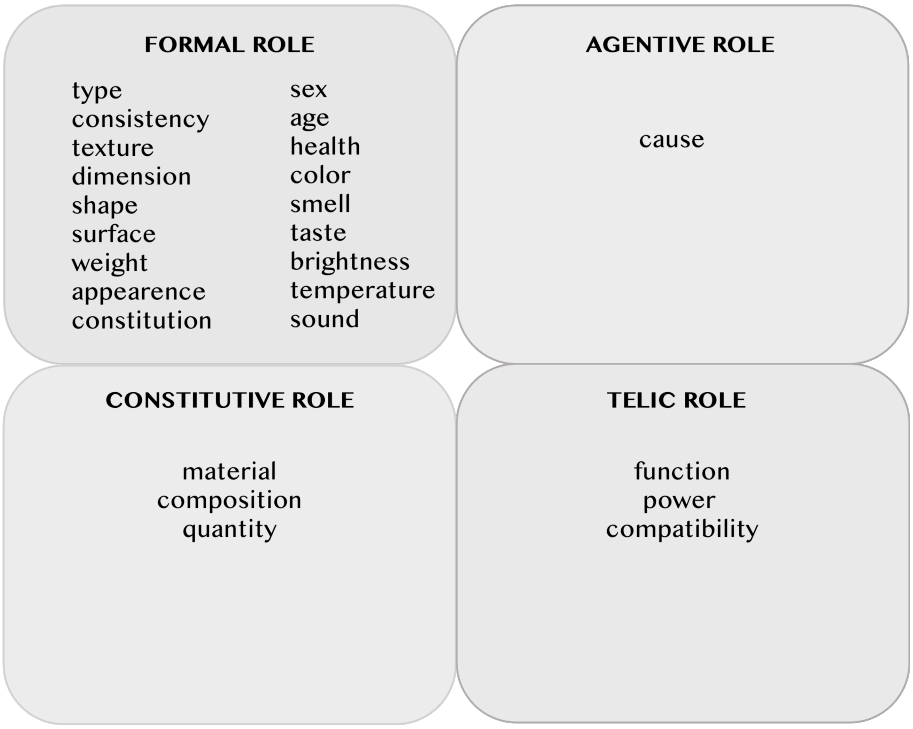


FIGURE 9.3: Mapping of the meanings covered by Qualia Roles to the categories from the GerCoAt inventory.

It is much more difficult to find corresponding adjective-noun phrases. Adjectival taxonomy in GermaNet does contain the nodes *ursachenspezifisch* 'cause-related' and *funktionalitätsspezifisch* 'function-related'. The hierarchy of the former is flat and goes only one level deeper: the node has 44 direct hyponyms such as the denominal adjectives *wetterbedingt* 'due to the weather' and *ernährungsbedingt* 'due to nutrition'. The node 'function-related' has a larger family size: there are 196 hyponyms in total going up to five levels deep. Most of the adjectives under that node are deverbal ones with the suffix *-bar* as in *zusammenklappbar* 'foldable' and *drehbar* 'rotatable'. This shows that the two meanings can be expressed by adjectives, but the group of possible lexical units is very homogeneous as they are either derived from nouns or from verbs.

There is a large number of categories that could not be mapped to any of the four qualia dimensions. The labels `intensity` and `manner/description` describe an inherent property of nouns denoting states, processes and events. These categories together with `character` and `intelligence` resemble specifications of the FORMAL role as they are capable of distinguishing the noun within a larger domain. Consider the attribute `character`: a *shy person* is distinguished from a *confident person*. However, the FORMAL dimension of a *person* can also be represented by adjectives referring to physical properties of a PERSON such as *tall* and *short*; or *blond* and *brunet*; or *old* and *young*. Should there be a distinction between physical FORMAL role and abstract FORMAL role? Apart from that, one attribute stands aside as it either addresses more than one dimension or it is not clear without context which dimension is selected, namely `evaluation`.

Attributes and Frame Elements. Finally, the alignment with Frame Elements (FEs) is partially possible, especially for those attributes that could not be mapped to qualia roles, the ones typically describing abstract nouns such as processes, states, and events. Most of the attributes that do map to qualia roles have their equivalents among FEs: e.g. `color`, `material`, `dimension`, `age`. However, they are found not in all possible frames, e.g. in the frame "Clothing" the FE MATERIAL is present, but not COLOR, or DIMENSION, or COST. It has already been advocated in Gamerschlag et al. (2014) that a combination of Frame Semantics with Concept Frames could provide much richer semantic representations: both the participants of situations (Frame Elements) and attributes are necessary for representing events, concepts, and states of affairs as frames.

Statistical asymmetry. Chapter 8 studied a statistical property of some MWEs that has not yet received much attention in research, namely, *asymmetry*. A corpus study was conducted in order to determine what types of asymmetric AN co-occurrences make their way into the n-best lists based on an asymmetric statistical association measure ΔP. It showed that ΔP provides additional information about restrictedness in AN combinations: it is very effective at extracting the so-called *unilexemes*, i.e. lexical units that have a highly restricted combinability. For instance, the German adjective *himmelweit* 'enormous' has a very strong preference for the noun *Unterschied* 'difference' and is unlikely to be used to describe semantically related nouns such as *Diskrepanz* 'discrepancy' or *Kontrast* 'contrast'. This is reflected in the value of ΔP computed based on corpus frequencies.

9.2 Future work

This thesis proposes an inventory of relations that hold between the elements of adjective-noun combinations. Such relations are referred to as attributes in this thesis. The proposed annotation scheme is robust and capable of accommodating descriptive adjectives from various semantic classes, but it is not exhaustive. In order to obtain a complete picture about the meanings expressed by adjectival modifiers, relational adjectives and participles should be taken into consideration.

Relational adjectives or pertainyms have been only shortly discussed in this thesis. Relations in pertainyms have been studied and several inventories have been proposed (see Warren (1984). Apart from that, as this thesis suggests, relations developed for noun-noun compounds can be applied to AN phrases with a pertainym as modifier. This is especially relevant for languages that do not rely on compounding as much as German and English, for instance Russian. Russian does license compounding as well, but a common way to express the meanings corresponding to compounds in English or German is building phrases with relational adjectives, as well as

prepositional phrases or nominal phrases in Genitive case.[1] A few examples of phrases with relational adjectives in Russian are presented in (54). The proposed relations are borrowed from the inventory of NN compounds proposed in Dima (2019).

(54) a. *karton-n-aja* *korobka*; relation `material`
paper-ADJ-INFL box
'paper box'

b. *vesen-n-ij* *den'*; relation `time`
spring-ADJ-INFL day
'spring day'

c. *dver-n-ajarutchka*; relation `is-part`
door-ADJ-INFL handle
'door handle'

d. *moloch-n-yjkoktejl*; relation `has-ingredient`
milk-ADJ-INFL cocktail
'milk cocktail'

Apart from that, relational adjectives can acquire descriptive readings and form collocations as for instance, *volchij appetit*[2] lit. 'wolf's appetite' which means 'ravenous appetite'.

Participles were shortly discussed in the previous chapter (section 8.3). It was noted that they express meanings that are not present in the developed inventory of relations as they either denote an *activity* performed by their head noun (head noun is a subject) or an activity done to their head noun (head noun is a direct object). A more fine-grained distinction is needed for such examples which should stem from an inventory developed for verb-noun combinations. For verbal collocations, an inventory of about 20 standard LFs has been developed by Mel'čuk (2015). Mel'čuk (2015) distinguishes standard LFs for the following meanings: causation (*to strike fear*), phase (*to open fire*), realisation/fulfillment (*to meet the demands*), involvement (*a car hits*), manifestation (*scorn drips*), proximity (*storm gathers*), preparation (*to load a program*), degradation (*milk turns (sour*), characteristic sound (*wind howls*), obstruction (*economy stagnates*), cessation (*voice breaks*), excessiveness (*engine races*). Those simple LFs can be further combined into complex ones.

Naturally, the question arises whether the inventory of LFs can accommodate the variety of meanings expressed by in verb-noun pairs. Or is its

[1] A comparison of English compounds and Russian relational adjectives is presented in Mezhevich (2002).

[2] Such adjectives, derived mostly from animate nouns are sometimes referred to as *possessive adjectives*. They express the meaning of belonging in their literal sense: *volch'ya lapa* = 'a wolf's paw'

coverage limited, similarly to the adjectival LFs? Assuming that this is the case and there is a large number of meanings not reflected in the LFs, what approach should be used to expand the inventory? Can a wordnet be used following the approach to adjective-noun phrases in this thesis? Both WordNet and GermaNet classify verbs into the following semantic classes:[3]

- Possession
- Location
- Emotion
- Social Verbs
- Body
- Cognition
- Communication
- Competition
- Contact
- natural Phenomenon
- Creation
- Change
- Consumption
- Perception
- Stative Verbs

In order to get a sense of whether and how information about relations in verb-noun pairs is reflected in the wordnets, consider three examples:

- **light-verb construction *to make a decision*:** in WordNet, the verb *to make* has 49 senses one of which corresponds to "perform or carry out" with the examples *to make a decision/a move/advances/a phone call*. Its semantic class is CREATION. The corresponding collocation in German is *Entscheidung treffen* (lit. 'to meet a decision'), the verb's sense is not modelled in GermaNet as it is too specific.

- **causation *to strike fear***: in WordNet, the verb's semantic class is PERCEPTION, whereas its definition is "cause to experience suddenly". Thus, the causation meaning is only reflected in the definition, but not in the semantic class of the verb. The German equivalent is more straightforward: *Angst auslösen* 'to cause fear' where the verb *auslösen* in that sense belongs to the semantic class CREATION.

- **characteristic sound *wind howls***: the verb *to howl* (German *heulen*) is a hyponym of the node *to make noise* in Wordnet, and of the node *allgemeines Geräusch* 'general sound' in GermaNet. However, the semantic fields are different: COMMUNICATION in WordNet versus PERCEPTION in GermaNet. Obviously, with regard to *wind*, the more suitable semantic class is PERCEPTION.

The examples presented above indicate that both wordnets provide relevant semantic information for verbs, but not in a consistent way. The semantic classes of verbs in both WordNet and GermaNet can be considered as potential labels for coarse-grained classification of verbal collocations. For adjectives, this thesis relied on a more fine-grained classification one level below the semantic classes. Following the same path for verbs would be

[3]Further fine-grained classification in the two lexical-semantic nets differs to various degrees.

more challenging due to the complexity of verbal hierarchy. Verbal taxonomy is much deeper than the adjectival one. Moreover, the semantic classes listed above are not part of the verbal taxonomy,[4] and there are verbs that do not inherit the semantic class of their hypernyms. For instance, the synset [*bekommen, erhalten, erfahren*] 'to get, receive' in GermaNet belongs to the semantic class STATIVE VERBS, whereas its direct hyponym *empfangen* 'to receive' belongs to COMMUNICATION. On a more fine-grained level, the categories are not as consistent as the artificial concepts integrated as nodes in the adjectival taxonomy.

Morphosyntactic form. This thesis studied bigram co-occurrences because it focused on the relations that hold between the base and its collocate. Neither a broader context, nor morphosyntactic forms were considered in this work apart from relying on the context to semantically annotate collocations. However, examples from corpora suggest that there are adjective-noun co-occurrences that tend to have a fixed morphosyntactic form. For example:

(55) a. *dicke Freunde* 'close friends': this collocation is mostly used in plural form as opposed to, say, *enger Freund* 'close friend';
b. *in freier Natur/Wildbahn* 'in the wild' is fixed as a prepositional phrase.
c. the pair *gut+Letzt* is only used in the phrase *zu guter Letzt* which stands for 'last but not least'.

Flexibility of form, or, in this case, its rigidity, is one of the defining characteristic features of multi-word expressions. Collocations are considered to be flexible, but does this hold true for all types of collocations? It would be interesting to conduct a study on such tendencies for other collocations and compare them to other MWEs and free phrases.

[4] Remember that in the adjectival taxonomy, semantic classes are the unique beginners and thus are directly integrated into the hierarchy of adjectives (see section 4.5.1)

Bibliography

Lorraine G Allan. 1980. A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, 15(3):147–149.

Margarita Alonso Ramos, Alfonso Nishikawa, and Orsolya Vincze. 2009. Dice in the web: An online spanish collocation dictionary. In *eLexicography in the 21st century : New challenges, new applications. Proceedings of eLex 2009*, pages 369–374.

Margarita Alonso Ramos and Leo Wanner. 2007. Collocation chains: how to deal with them? In *Proceedings of the 3rd International Conference on Meaning-Text Theory*, pages 11–20, Klagenfurt. Wiener Slawischer Almanach, Sonderband 69.

Salvatore Attardo. 2005. The role of affordances at the semantics/pragmatics boundary. In *Proceedings of the Annual Meeting of the Cognitive Science Society, volume 27*, pages 169–174.

R Harald Baayen. 2011. Corpus linguistics and naive discriminative learning. *Revista Brasileira de Linguística Aplicada*, 11(2):295–328.

Timothy Baldwin and Su Nam Kim. 2010. Multiword expressions. Second edition, pages 267–292. CRC Press.

Adrien Barbaresi. 2018. A corpus of German political speeches from the 21st century. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, pages 792–797, Paris, France. European Language Resources Association (ELRA).

Marco Baroni, Raffaella Bernardi, and Roberto Zamparelli. 2014. Frege in space: A program for composition distributional semantics. In *Linguistic Issues in Language Technology, Volume 9, 2014 - Perspectives on Semantic Representations for Textual Inference*. CSLI Publications.

Marco Baroni and Roberto Zamparelli. 2010. Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1183–1193, Cambridge, MA. Association for Computational Linguistics.

Lawrence W. Barsalou. 1992. Frames, concepts, and conceptual fields. In *Frames, fields, and contrasts: New essays in semantic and lexical organization*, pages 21–74. Lawrence Erlbaum Associates, Inc.

Lawrence W Barsalou. 2017. What does semantic tiling of the cortex tell us about semantics? *Neuropsychologia*, (105):18–38.

Sabine Bartsch. 2004. *Structural and functional properties of collocations in English: A corpus study of lexical and pragmatic constraints on lexical co-occurrence*. Gunter Narr Verlag.

Nuria Bel, Federica Busa, Nicoletta Calzolari, Elisabetta Gola, Alessandro Lenci, Monica Monachini, Antoine Ogonowski, Ivonne Peters, Wim Peters, Nilda Ruimy, Marta Villegas, and Antonio Zampolli. 2000. SIMPLE: A general framework for the development of multilingual lexicons. In *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC'00)*, Athens, Greece. European Language Resources Association (ELRA).

Frederico Belcavello, Marcelo Viridiano, Alexandre Diniz da Costa, Ely Edison da Silva Matos, and Tiago Timponi Torrent. 2020. Frame-based annotation of multimodal corpora: Tracking (a)synchronies in meaning construction. In *Proceedings of the International FrameNet Workshop 2020: Towards a Global, Multilingual FrameNet*, pages 23–30, Marseille, France. European Language Resources Association.

Melanie J. Bell and Martin Schäfer. 2013. Semantic transparency: challenges for distributional semantics. In *Proceedings of the IWCS 2013 Workshop Towards a Formal Distributional Semantics*, pages 1–10, Potsdam, Germany. Association for Computational Linguistics.

Morton Benson. 1989. The structure of the collocational dictionary. *International Journal of Lexicography*, 2:1–14.

Morton Benson, Robert Ilson, and Evelyn Benson. 1986. *Lexicographic Description of English.*, volume 14 of *Studies in Language Companion Series (SLCS)*. John Benjamins Publishing Company.

Chris Biemann and Eugenie Giesbrecht. 2011. Distributional semantics and compositionality 2011: Shared task description and results. In *Proceedings of the Workshop on Distributional Semantics and Compositionality*, pages 21–28, Portland, Oregon, USA. Association for Computational Linguistics.

Christopher M Bishop. 2006. *Pattern recognition and machine learning*. Springer Science+ Business Media.

Gemma Boleda. 2006. *Automatic acquisition of semantic classes for adjectives*. Ph.D. thesis, Pompeu Fabra University.

Gerlof Bouma. 2009. Normalized (pointwise) mutual information in collocation extraction. *Proceedings of GSCL*, pages 31–40.

Sabine Brants, Stefanie Dipper, Peter Eisenberg, Silvia Hansen, Esther König, Wolfgang Lezius, Christian Rohrer, George Smith, and Hans Uszkoreit. 2004. TIGER: Linguistic interpretation of a German corpus. *Journal of Language and Computation*, 2:597–620.

Joan Bresnan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Padó, and Manfred Pinkal. 2009. Using FrameNet for the semantic analysis of German: Annotation, representation, and automation. In Hans C. Boas, editor, *Multilingual FrameNets in Computational Lexicography*, pages 209–244. Mouton.

Dietrich Busse. 2012. *Frame-Semantik. Ein Kompendium*. De Gruyter, Berlin, Boston.

Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20:37–46.

Silvio Cordeiro, Aline Villavicencio, Marco Idiart, and Carlos Ramisch. 2019. Unsupervised compositionality prediction of nominal compounds. *Computational Linguistics*, 45(1):1–57.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Jörg Didakowski and Alexander Geyken. 2014. From DWDS corpora to a German word profile – methodological problems and solutions. In *OPAL – Online publizierte Arbeiten zur Linguistik 2/2014*, pages 39–47.

Corina Dima. 2019. *Composition Models for the Representation and Semantic Interpretation of Nominal Compounds*. Ph.D. thesis, Eberhard Karls Universität Tübingen.

David Dowty. 1991. Thematic proto-roles and argument selection. *Language*, 67(3):547–619.

Bradley Efron and Robert J Tibshirani. 1994. *An introduction to the bootstrap*. CRC press.

Peter Eisenberg. 2020. *Grundriss der deutschen Grammatik. Das Wort*. J.B. Metzler.

Peter Eisenberg and Rolf Schöneich. 2020. *Grundriss der deutschen Grammatik. Der Satz*. J.B. Metzler.

Nick C. Ellis. 2006. Language acquisition as rational contingency learning. *Applied Linguistics*, 27(1):1–24.

Luis Espinosa-Anke, Steven Schockaert, and Leo Wanner. 2019. Collocation classification with unsupervised relation vectors. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5765–5772, Florence, Italy. Association for Computational Linguistics.

Stefan Evert. 2004. *The statistics of word cooccurrences: word pairs and collocations*. Ph.D. thesis, Institut für maschinelle Sprachverarbeitung, University of Stuttgart.

Stefan Evert. 2008. A lexicographic evaluation of german adjective-noun collocations. *Proceedings of the LREC Workshop Towards a Shared Task for Multiword Expressions (MWE 2008), Marrakech, Morocco*.

Stefan Evert and Hannah Kermes. 2003. Experiments on candidate data for collocation extraction. In *10th Conference of the European Chapter of the Association for Computational Linguistics*, Budapest, Hungary. Association for Computational Linguistics.

Stefan Evert, Peter Uhrig, Sabine Bartsch, and Thomas Proisl. 2017. E-view-alation - a large-scale evaluation study of association measures for collocation identification. *Proceedings of eLex 2017âElectronic lexicography in the 21st century: Lexicography from Scratch*, pages 531–549.

Neele Falk, Yana Strakatova, Eva Huber, and Erhard Hinrichs. 2021. Automatic classification of attributes in german adjective-noun phrases. In *Proceedings of the 14th International Conference on Computational Semantics (IWCS)*, pages 239–249, Groningen, The Netherlands (online). Association for Computational Linguistics.

Alvan R. Feinstein and Domenic V. Cicchetti. 1990. High agreement but low kappa: I. The problems of two paradoxes. *Journal of clinical epidemiology*, 43(6):543–549.

Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. MIT Press, Cambrige, MA.

Guangchao Charles Feng. 2013. Factors affecting intercoder reliability: a Monte Carlo experiment. *Quality & Quantity*, 47(5):2959–2982.

Charles J. Fillmore. 1968. The case for case. In Emmon Bach and R. Harms, editors, *Universals in Linguistic Theory*. Holt, Rinehart, and Winston.

Charles J. Fillmore. 1982. Frame semantics. In *Linguistics in the Morning Calm*, pages 111–137. Hanshin Publishing Co., Seoul, South Korea.

Charles J. Fillmore. 1985. Frames and the semantics of understanding. In *Quaderni di Semantica*, volume 6, pages 222–254.

Charles J. Fillmore, Christopher R. Johnson, and Miriam R.L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography*, 16(3):235–250.

Irene Fioravanti, Marco Silvio Giuseppe Senaldi, Alessandro Lenci, and Anna Siyanova-Chanturia. 2021. Lexical fixedness and compositionality in L1 speakers' and L2 learners' intuitions about word combinations: Evidence from Italian. *Second Language Research*, 37(2):291–322.

John R Firth. 1957. *Papers in linguistics: 1934 - 1951*. London: Oxford University Press.

Joseph L. Fleiss. 1971. Measuring nominal scale agreement among many raters. volume 76, pages 378–382.

Thomas Gamerschlag, Doris Gerland, Rainer Osswald, and Wiebke Petersen. 2014. *General Introduction*. Springer International Publishing, Cham.

Marcos Garcia, Marcos García Salido, and Margarita Alonso-Ramos. 2019. A comparison of statistical association measures for identifying dependency-based collocations in various languages. In *Proceedings of the Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)*, pages 49–59, Florence, Italy. Association for Computational Linguistics.

Dirk Geeraerts. 2009. *Theories of Lexical Semantics*. Oxford University Press.

Alexander Geyken, Jörg Didakowski, and Alexander Siebert. 2008. Generation of word profiles on the basis of a large and balanced german corpus. In *Proceedings of the 13th EURALEX International Congress*, pages 371–383, Barcelona, Spain. Institut Universitari de Linguistica Aplicada, Universitat Pompeu Fabra.

James J. Gibson. 1979. The theory of affordances. In *The Ecological Approach to Visual Perception*, pages 127–137. Houghton Mifflin Harcourt (HMH).

Daniel Gildea and Martha Palmer. 2002. The necessity of parsing for predicate argument recognition. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 239–246, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

Wendy van Ginkel and Ton Dijkstra. 2020. The tug of war between an idiom's figurative and literal meanings: Evidence from native and bilingual speakers. *Bilingualism: Language and Cognition*, 23(1):131–147.

Stefan Th Gries. 2013. 50-something years of work on collocations. *International Journal of Corpus Linguistics*, 18(1):137–166.

Stefan Th. Gries and Philip Durrant. 2021. Analyzing co-occurrence data. In Magali Paquot and Stefan Th. Gries, editors, *A practical handbook of corpus linguistics*, pages 141–160. Springer, Berlin New York.

Jeffrey S. Gruber. 1965. *Studies in lexical relations*. Ph.D. thesis.

Emiliano Guevara. 2010. A regression model of adjective-noun compositionality in distributional semantics. In *Proceedings of the 2010 Workshop on GEometrical Models of Natural Language Semantics*, pages 33–37, Uppsala, Sweden. Association for Computational Linguistics.

Kilem Li Gwet. 2008. Computing inter-rater reliability and its variance in the presence of high agreement. 61(Pt 1):29–48.

Hessel Haagsma, Johan Bos, and Malvina Nissim. 2020. MAGPIE: A large corpus of potentially idiomatic expressions. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 279–287, Marseille, France. European Language Resources Association.

Sepand Haghighi, Masoomeh Jasemi, Shaahin Hessabi, and Alireza Zolanvari. 2018. PyCM: Multiclass confusion matrix library in Python. *Journal of Open Source Software*, 3(25):729.

M.A.K. Halliday. 1966. Beginning the study of lexis. In R.H. Robin C. E. Bazell, J.C. Catford, M.A.K. Halliday, editor, *In Memory of J.R. Firth*, Longman edition, pages 148–430. London.

Birgit Hamp and Helmut Feldweg. 1997. GermaNet - a lexical-semantic net for german. In *Proceedings of the ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*, Madrid.

Matthias Hartung. 2015. *Distributional Semantic Models of Attribute Meaning in Adjectives and Nouns*. Ph.D. thesis, Heidelberg University, Germany.

Matthias Hartung, Fabian Kaupmann, Soufian Jebbara, and Philipp Cimiano. 2017. Learning compositionality functions on word embeddings for modelling attribute meaning in adjective-noun phrases. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 54–64, Valencia, Spain. Association for Computational Linguistics.

Franz Josef Hausmann. 1984. Wortschatzlernen Ist Kollokationslernen. Zum Lehren Und Lernen Französischer Wortverbindungen. *Praxis des neusprachlichen Unterrichts*, 31:395–407.

Franz Josef Hausmann. 1985. Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. In *Lexikographie und Grammatik*, pages 118 – 129. De Gruyter, Berlin, Boston.

Verena Henrich and Erhard Hinrichs. 2010. GernEdiT - the GermaNet editing tool. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

Verena Henrich, Erhard Hinrichs, and Reinhild Barkey. 2014. Aligning word senses in GermaNet and the DWDS dictionary of the German language. In *Proceedings of the Seventh Global Wordnet Conference*, pages 63–70, Tartu, Estonia. University of Tartu Press.

Marie Hinrichs, Richard Lawrence, and Erhard Hinrichs. 2020. Exploring and visualizing Wordnet data with GermaNet Rover. In *Proceedings of the CLARIN Annual Conference 2020*, pages 32–36.

Franz Hundsnurscher and Jochen Splett. 1982. *Semantik der Adjektive des Deutschen: Analyse der semantischen Relationen*. Westdeutscher Verlag, Opladen.

Anne-Laure Jousse. 2007. Formalizing non-standard lexical functions. In *Proceedings of the 3rd International Conference on Meaning-Text Theory*, pages 219–228, Klagenfurt. Wiener Slawischer Almanach, Sonderband 69.

Hans Kamp and Barbara Partee. 1995. Prototype theory and compositionality. *Cognition*, 57(2):129–191.

Adam Kilgarriff. 2006. Collocationality (and how to measure it). In *Proceedings of the 12th EURALEX International Congress*, pages 997–1004, Torino, Italy. Edizioni dell'Orso.

Adam Kilgarriff, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. The Sketch Engine: ten years on. *Lexicography*, pages 7–36.

Adam Kilgarriff, Pavel Rychlý, Pavel Smrž, and David Tugwell. 2004. The Sketch Engine. *Information Technology*.

Paul Kingsbury and Martha Palmer. 2002. From TreeBank to PropBank. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'02)*, Las Palmas, Canary Islands - Spain. European Language Resources Association (ELRA).

Göran Kjellmer. 1991. A mint of phrases. In *English Corpus Linguistics*. Longman, London.

Philipp Koehn. 2005. Europarl: A parallel corpus for statistical machine translation. In *Proceedings of Machine Translation Summit X: Papers*, pages 79–86, Phuket, Thailand.

Daniël de Kok and Sebastian Pütz. 2019. *Stylebook for the Tübingen Treebank of Dependency-parsed German (TüBa-D/DP)*. Seminar für Sprachwissenschaft, Universität Tübingen, Tübingen, Germany.

Menelaos Konstantinidis, Lisa. W. Le, and Xin Gao. 2022. An empirical comparative assessment of inter-rater agreement of binary outcomes and multiple raters. *Symmetry*, 14(2).

Maximilian Köper and Sabine Schulte im Walde. 2016. Automatically generated affective norms of abstractness, arousal, imageability and valence for 350 000 German lemmas. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2595–2598, Portorož, Slovenia. European Language Resources Association (ELRA).

Iztok Kosem, Simon Krek, and Polona Gantar. 2020. Defining collocation for slovenian lexical resources. *Slovenščina 2.0: empirical, applied and interdisciplinary research*, 8(2):1–27.

Brigitte Krenn. 2000. CDB – a database of lexical collocations. In *Proceedings of the 2nd International Conference on Language Resources & Evaluation*, Athens, Greece.

J. Richard Landis and Gary G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174.

Andrea Lehr. 1996. *Kollokationen und maschinenlesbare Korpora, volume 168 of Germanistische Linguistik*. Niemeyer, Tubingen.

Alessandro Lenci, Federica Busa, Nilda Ruimy, Elisabetta Gola, Monica Monachini, Nicoletta Calzolari, Antonio Zampolli, Emilie Guimier, Gaëlle Recourcé, Lee Humphreys, Ursula Von Rekovsky, Antoine Ogonowski, Clare McCauley, Wim Peters, Ivonne Peters, Robert Gaizauskas, and Marta Villegas. 2000. *SIMPLE Work Package 2-Linguistic Specifications. Deliverable D2. 1*.

Gary Libben, Martha Gibson, Yeo Bom Yoon, and Dominiek Sandra. 2003. Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language*, 84(1):50 – 64. Brain and Language Special Issue.

An Van Linden. 2012. *Modal Adjectives: English Deontic and Evaluative Constructions in Diachrony and Synchrony*. De Gruyter Mouton.

Sebastian Löbner. 2014. Evidence for frames from human language. In Thomas Gamerschlag, Doris Gerland, Rainer Osswald, and Wiebke Petersen, editors, *Frames and Concept Types: Applications in Language and Philosophy*, pages 23–67. Springer International Publishing, Cham.

Sebastian Löbner. 2021. Frames at the interface of language and cognition. *Annual Review of Linguistics*, 7(1):261–284.

John Lyons. 1977. *Semantics*. Cambridge University Press.

Christopher D Manning and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA, USA.

Diana McCarthy, Sriram Venkatapathy, and Aravind Joshi. 2007. Detecting compositionality of verb-object combinations using selectional preferences. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 369–379, Prague, Czech Republic. Association for Computational Linguistics.

Colin McIntosh, Ben Francis, and Richard Poole. 2009. *Oxford Collocations Dictionary for students of English*. Oxford University Press.

Igor Mel'čuk. 1996. Lexical functions: a tool for the description of lexical relations in a lexicon. *Lexical functions in lexicography and natural language processing*, 31:37–102.

Igor Mel'čuk. 1998. Collocations and lexical functions. In A.P.Cowie, editor, *Phraseology. Theory, Analysis, and Applications*, pages 23–55. Clarendon press, Oxford.

Igor Mel'čuk. 1999. *Opyt teorii lingvističeskich modelej «Smysl <=> Tekst»*. Shkola "Yaziki russkoi kultury", Moskva.

Igor Mel'čuk. 2012a. Phraseology in the language, in the dictionary, and in the computer. *Yearbook of Phraseology*, 3(1).

Igor Mel'čuk. 2012b. *Yazik: ot smisla k tekstu*. Languages of Slavic Cultures, Moscow.

Igor Mel'čuk. 2015. *Semantics: From Meaning to Text*, volume 3. John Benjamins, Amsterdam.

Igor Mel'čuk. 2023. *General phraseology. Theory and practice*. John Benjamins Publishing Company.

Igor Mel'čuk, Nadia Arbatchewsky-Jumarie, Lidija Iordanskaja, Suzanne Mantha, and Polguère Alain. 2007. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques IV*. Les Presses de l'Université de Montréal, Montréal.

Igor Mel'čuk and Leo Wanner. 1996. Lexical functions and lexical inheritance for emotion lexemes in german. volume 31 of *Studies in Language Companion Series*, pages 209–278. John Benjamins Publishing Company.

Igor Mel'čuk and Alexander Zholkovsky. 1984. Explanatory combinatorial dictionary of modern Russian. *Wiener Slawistischer Almanach*.

A. Meyers, R. Reeves, C. Macleod, R. Szekely, V. Zielinska, B. Young, and R. Grishman. 2004. The nombank project: An interim report. In *HLT-NAACL 2004 Workshop: Frontiers in Corpus Annotation*, pages 24–31, Boston, Massachusetts, USA. Association for Computational Linguistics.

Ilana Mezhevich. 2002. English compounds and russian relational adjectives. In *Proceedings of the North West Linguistics Conference 2002 (NWCL'2002)*, pages 95–114, Canada. Simon Fraser University Linguistics Graduate Student Association.

Lukas Michelbacher, Stefan Evert, and Hinrich Schütze. 2007. Asymmetric association measures. *Proceedings of the Recent Advances in Natural Language Processing (RANLP 2007)*.

Lukas Michelbacher, Stefan Evert, and Hinrich Schütze. 2011. Asymmetry in corpus-derived and human word associations. *Corpus Linguistics and Linguistic Theory*, 7(2):245 – 276.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems 26*, pages 3111–3119, Lake Tahoe, Nevada, USA. Curran Associates, Inc.

Katherine J. Miller. 1998. Modifiers in WordNet. In Christiane Fellbaum, editor, *WordNet: An Electronic Lexical Database*. The MIT Press.

Marvin Minsky. 1974. A framework for representing knowledge. *MIT-AI Laboratory Memo 306*.

Rosamund Moon. 1998. *Fixed Expressions and Idioms in English: A corpus-based approach*. Clarendon Press, Oxford.

Julius M. Moravcsik. 1991. What Makes Reality Intelligible? Reflections on Aristotle's Theory of Aitia. In Lindsay Judson, editor, *Aristotle's Physics: A Collection of Essays*, pages 31–48. Oxford University Press.

Daniela Naumann, Diego Frassinelli, and Sabine Schulte im Walde. 2018. Quantitative semantic variation in the contexts of concrete and abstract words. In *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics*, pages 76–85, New Orleans, Louisiana. Association for Computational Linguistics.

Roberto Navigli and Simone Paolo Ponzetto. 2012. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217–250.

Douglas L. Nelson, Cathy L. McEvoy, and Thomas A. Schreiber. 1998. The university of south florida word association, rhyme, and word fragment norms.

Nadja Nesselhauf. 2003. The use of collocations by advanced learners of english and some implications for teaching. *Applied linguistics*, 24:223–242.

Geoffrey Nunberg, Ivan A Sag, and Thomas Wasow. 1994. Idioms. volume 70, pages 491–538. Johns Hopkins University Press.

Sebastian Padó. 2007. *Cross-Lingual Annotation Projection Models for Role-Semantic Information*. Ph.D. thesis, Saarland University.

Irene Pagliai. 2023. Bridging the gap: Creation of a lexicon of 150 pairs of English and Italian idioms including normed variables for the exploration of idiomatic ambiguity. *Journal of Open Humanities Data*.

H.E. Palmer. 1938. *A Grammar of English Words*. Longmans, Green.

Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106.

Ellie Pavlick and Chris Callison-Burch. 2016. So-called non-subsective adjectives. In *Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics*, pages 114–119, Berlin, Germany. Association for Computational Linguistics.

Pavel Pecina. 2008. *Lexical Association Measures: Collocation Extraction*. Ph.D. thesis.

Pavel Pecina and Pavel Schlesinger. 2006. Combining association measures for collocation extraction. In *Proceedings of the COLING/ACL on Main conference poster sessions*, pages 651–658. Association for Computational Linguistics.

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12(85):2825–2830.

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar. Association for Computational Linguistics.

Ivonne Peters and Wim Peters. 2000. The treatment of adjectives in SIMPLE: Theoretical observations. In *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC'00)*, Athens, Greece. European Language Resources Association (ELRA).

Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 2227–2237. Association for Computational Linguistics.

Gill Philip. 2011. *Colouring Meaning : Collocation and Connotation in Figurative Language.*, volume 43 of *Studies in Corpus Linguistics*. John Benjamins Publishing Co.

Carl Pollard and Ivan A. Sag. 1994. *Head-driven phrase structure grammar.* Chicago/London:The University of Chicago Press.

Sameer Pradhan, Julia Bonn, Skatje Myers, Kathryn Conger, Tim O'gorman, James Gung, Kristin Wright-bettner, and Martha Palmer. 2022. PropBank comes of Age—Larger, smarter, and more diverse. In *Proceedings of the 11th Joint Conference on Lexical and Computational Semantics*, pages 278–288, Seattle, Washington. Association for Computational Linguistics.

Pragglejaz Group. 2007. Mip: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, 22(1):1–39.

James Pustejovsky. 1991. The Generative Lexicon. *Computational Linguistics*, 17(4):409–441.

James Pustejovsky. 1995. *The Generative Lexicon.* Language, Speech, and Communication. A Bradford Book.

James Pustejovsky and Olga Batiukova. 2019. *The Lexicon*. Cambridge Textbooks in Linguistics. Cambridge University Press.

Uwe Quasthoff. 2011. *Wörterbuch der Kollokationen im Deutschen*. De Gruyter, Berlin, Boston.

Victor Raskin and Sergei Nirenburg. 1995. Lexical semantics of adjectives. *New Mexico State University, Computing Research Laboratory Technical Report, MCCS-95-288*.

Victor Raskin and Sergei Nirenburg. 1998. An applied ontological semantic microtheory of adjective meaning for natural language processing. In *Machine Translation*, pages 135–227.

Siva Reddy, Diana McCarthy, and Suresh Manandhar. 2011. An empirical study on compositionality in compound nouns. In *Proceedings of 5th International Joint Conference on Natural Language Processing*, pages 210–218, Chiang Mai, Thailand. Asian Federation of Natural Language Processing.

Tilmann Reuther. 1996. On dictionary entries for support verbs: The cases of Russian VESTI, PROVODIT' and PROIZVODIT'. volume v. 31 of *Studies in Language Companion Series*, pages 181–208. John Benjamins Publishing Company.

Nilda Ruimy, Elisabetta Gola, and Monica Monachini. 2001. Lexicography informs lexical semantics: The SIMPLE experience. In Pierrette Bouillon and Federica Busa, editors, *The language of word meaning*. Cambridge University Press.

Michael Rundell, editor. 2010. *Macmillan Collocations Dictionary for Learners of English*. Macmillan Education.

Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, Colin Baker, and Jan Scheffczyk. 2016. *FrameNet II: Extended Theory and Practice*.

Pavel Rychly. 2008. A lexicographer-friendly association score. In *Sojka, Petr /Horák, Aleš (Hg.): Proceedings of the 2nd Workshop on Recent Advances in Slavonic Natural Languages Processing, RASLAN 2008*, pages 6–9, Brno.

Ivan A. Sag, Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger. 2002. Multiword expressions: A pain in the neck for NLP. In *Computational Linguistics and Intelligent Text Processing*, pages 1–15, Berlin, Heidelberg. Springer Berlin Heidelberg.

Manfred Sailer. 2021. Idioms. In *Head-Driven Phrase Structure Grammar*, pages 777–809. Language Science Press.

Agata Savary, Marie Candito, Verginica Barbu Mititelu, Eduard Bejček, Fabienne Cap, Slavomír Čéplö, Silvio Ricardo Cordeiro, Gülşen Eryiğit, Voula Giouli, Maarten van Gompel, Yaakov HaCohen-Kerner, Jolanta Kovalevskaitė, Simon Krek, Chaya Liebeskind, Johanna Monti, Carla Parra Escartín, Lonneke van der Plas, Behrang QasemiZadeh, Carlos Ramisch, Federico Sangati, Ivelina Stoyanova, and Veronika Vincze. 2018. PARSEME multilingual corpus of verbal multiword expressions. In Agata Savary Veronika Vincze Stella Markantonatou, Carlos Ramisch, editor, *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*, pages 87–147. Language Science Press, Berlin.

Dietmar Schabus and Marcin Skowron. 2018. Academic-industrial perspective on the development and deployment of a moderation system for a newspaper website. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC)*, pages 1602–1605, Miyazaki, Japan.

Dietmar Schabus, Marcin Skowron, and Martin Trapp. 2017. One million posts: A data set of German online discussions. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, pages 1241–1244, Tokyo, Japan.

Roland Schäfer. 2015. Processing and querying large web corpora with the COW14 architecture. In *Proceedings of Challenges in the Management of Large Corpora 3 (CMLC-3)*, Lancaster. UCREL, IDS.

Roland Schäfer and Felix Bildhauer. 2012. Building large corpora from the web using a new efficient tool chain. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, pages 486–493, Istanbul, Turkey. European Language Resources Association (ELRA).

Ulrike Schneider. 2020. ΔP as a measure of collocation strength. *Corpus Linguistics and Linguistic Theory*, 16(2):249–274.

Patrick Schone and Daniel Jurafsky. 2001. Is knowledge-free induction of multiword unit dictionary headwords a solved problem? In *Proceedings of the 2001 Conference on Empirical Methods in Natural Language Processing*.

Karin Kipper Schuler. 2006. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. Ph.D. thesis, University of Pennsylvania.

Vered Shwartz and Ido Dagan. 2019. Still a pain in the neck: Evaluating text representations on lexical composition. *Transactions of the Association for Computational Linguistics*, 7:403–419.

J. McH. Sinclair. 1966. Lexis as a linguistic level. In R.H. Robin C. E. Bazell, J.C. Catford, M.A.K. Halliday, editor, *In Memory of J.R. Firth*, Longman edition, pages 410–461. London.

John Sinclair. 1991. *Corpus, concordance, collocation*. Oxford University Press.

Pia Sommerauer, Antske Fokkens, and Piek Vossen. 2020. Would you describe a leopard as yellow? Evaluating crowd-annotations with justified and informative disagreement. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 4798–4809, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Daniil Sorokin, Corina Dima, and Erhard Hinrichs. 2015. Classifying semantic relations in German nominal compounds using a hybrid annotation scheme. In *Journal of Cognitive Science 16-3*, pages 261–286. Institute for Cognitive Science, Seoul National University.

Caroline Sporleder, Linlin Li, Philip Gorinski, and Xaver Koch. 2010. Idioms in context: The IDIX corpus. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta. European Language Resources Association (ELRA).

Yana Strakatova, Neele Falk, Isabel Fuhrmann, Erhard Hinrichs, and Daniela Rossmann. 2020. All that glitters is not gold: A gold standard of adjective-noun collocations for German. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 4368–4378, Marseille, France. European Language Resources Association.

Yana Strakatova and Erhard Hinrichs. 2019. Semantic modelling of adjective-noun collocations using FrameNet. In *Proceedings of the Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)*, pages 104–113, Florence, Italy. Association for Computational Linguistics.

Michael Stubbs. 2001. *Words and phrases: Corpus studies of lexical semantics*. Blackwell Publishers Oxford.

Jörg Tiedemann. 2012. Parallel Data, Tools and Interfaces in OPUS. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*, pages 2214–2218, Istanbul, Turkey. European Language Resources Association (ELRA).

Amos Tversky. 1977. Features of similarity. *Psychological review*, 84:327.

Ton van der Wouden. 1997. *Negative Contexts. Collocation, polarity, and multiple negation*. Routledge.

Orsolya Vincze, Estela Mosqueira, and Margarita Alonso Ramos. 2011. An online collocation dictionary of Spanish. In *Proceedings of the 5th International Conference on Meaning-Text Theory. Barcelona*, pages 275–286.

Wiltrud Wagner. 2020. *Idioms and Ambiguity in Context: Phrasal and Compositional Readings of Idiomatic Expressions*. De Gruyter.

Leo Wanner. 1996. *Lexical Functions in Lexicography and Natural Language Processing*. Number v. 31 in Studies in Language Companion Series. John Benjamins Publishing Company.

Beatrice Warren. 1984. *Classifying adjectives*. Acta Universitatis Gothoburgensis, Göteborg, Sweden.

Beatrice Warren. 2003. The role of links and/or qualia in modifier-head constructions. In Brigitte Nerlich, editor, *Polysemy : Flexible Patterns of Meaning in Mind and Language*, pages 233–252. De Gruyter Mouton.

Markus Werning. 2008. The "complex first" paradox: Why do semantically thick concepts so early lexicalize as nouns? *Interaction Studies*, 9(1):67–83.

Michael Wilson. 1988. MRC Psycholinguistic Database: Machine-usable dictionary, version 2.00. *Behavior Research Methods, Instruments, & Computers*, 20(1):6–10.

Nahathai Wongpakaran, Tinakon Wongpakaran, Danny Wedding, and Kilem L Gwet. 2013. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Medical Research Methodology*, 13(1):61.

Jayden Ziegler and Jesse Snedeker. 2018. How broad are thematic roles? evidence from structural priming. *Cognition*, 179:221–240.

Gisela Zifonun, Ludger Hoffmann, and Bruno Strecker. 1997. *Grammatik der deutschen Sprache*. De Gruyter, Berlin, New York.

Pienie Zwitserlood. 1994. The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes*, 9(3):341–368.

Appendix A

Most frequent frame elements from FrameNet: top-50

frame element	frequency		frame element	frequency
Time	818		Domain	73
Place	685		Containing_event	71
Manner	663		Name	60
Degree	479		Period_of_iterations	58
Means	450		Event_description	54
Purpose	402		Role	54
Explanation	380		Event	53
Duration	273		Cognizer	51
Depictive	266		Speed	51
Circumstances	263		Distance	50
Agent	209		Iterations	50
Frequency	185		Subregion	49
Descriptor	161		Ground	47
Result	157		Direction	40
Instrument	132		Speaker	40
Entity	112		Experiencer	39
Source	102		Addressee	38
Goal	92		Attribute	36
Path	92		Patient	36
Theme	90		Area	34
Medium	82		Protagonist	34
Topic	81		Activity	33
Cause	80		Beneficiary	33
Type	79		Material	33
Particular_iteration	77		Re-encoding	33

TABLE A.1: Frame Elements sorted by their frequency in FrameNet. Frequency stands for the number of frames in which a given frame element is found in FrameNet

Appendix B

Literal senses of the adjectives in the GerCo dataset

Adjective	Definition
alt	gibt das Alter, die Lebensjahre an
arm	Gegenwort zu reich; ohne Geld, mittellos
barock	entsprechend der Bedeutung von Barock
bitter	Geschmacksbezeichnung (herb schmeckend)
blank	glänzend, strahlend
blind	ohne Sehvermögen
deftig	kräftig u. nahrhaft
dick	gibt den Abstand zwischen zwei gegenüberliegenden Seiten eines Körpers an
dumm	ohne Verstand, töricht
dunkel	ohne Licht, lichtarm, finster; Gegenteil zu hell
frech	es jmdm. gegenüber in herausfordernder Weise an der nötigen Achtung fehlen lassend, (ein wenig) respektlos
frei	unabhängig
frischgebacken	eben erst gebacken
grob	Gegenteil von fein; rau, derb
hart	Gegenteil von weich; fest, keinem Druck nachgebend
hell	Gegenwort zu dunkel
herrlich	prachtvoll, hervorragend, ausgezeichnet
historisch	die Geschichte betreffend, geschichtlich
hölzern	aus Holz
karg	gering, nicht reichlich
knapp	nicht reichlich, kaum ausreichend
leicht	Gegenwort von schwer, von geringem Gewicht
mächtig	Macht besitzend, einflussreich
mild	sanft, weich
offen	nicht mit einem Schlüssel zugeschlossen, nicht verschlossen
prall	dick und straff gespannt
rau	uneben, nicht glatt
reich	wohlhabend, vermögend, begütert, viele Werte besitzend

rund	kugelförmig, kreisförmig, ringförmig, ohne Ecken und Kanten
sanft	ausgeglichen und friedfertig
scharf	gut schneidend; Gegenteil von stumpf
schlau	mit viel Verstand und List
schwarz	von völlig dunkler Färbung, (nahezu) alle Lichtstrahlen absorbierend, kein Licht reflektierend
sicher	nicht von Gefahr bedroht, ungefährdet
spät	am Ende eines Zeitabschnitts
stark	viel Kraft besitzend (physisch, einflussreich)
starr	unbeweglich, steif
steil	oft bis zur Senkrechten nach oben ansteigend oder nach unten abfallend
steinig	reich an Steinen
stürmisch	stark windig, mit Sturm
süß	wie Zucker, Honig schmeckend
teuer	hoch im Preis, viel Geld kostend, kostspielig
tief	von oben (weit) nach unten reichend, sich weit nach unten erstreckend
tot	gestorben, nicht mehr am Leben
wild	nicht gezähmt, freilebend;
windig	mit Wind, voller Luftbewegung
zäh	schwerflüssig, zähflüssig
zart	fein, weich, dünn (Material)

Appendix C

Initial guidelines

In the following task you will be presented with a list of adjective-noun collocations. A context sentence is given for each collocation to help avoid ambiguity. You will be asked to perform the following subtasks (see next page for an example annotation):

1. **Choose the suitable attribute for a given noun phrase,** e.g. in *hohe Stimme* the adjective *hoch* refers to **`Ton`**, whereas in *hohe Miete* the attribute is **`Quantität`**. Altogether there are 56 attributes available for the given dataset. You will find the list of all the attributes with their definitions and examples/comments after the example.

   For each adjective, a list of possible attributes will be given (the blue table in the upper left corner). However, there may be cases where none of the suggested attributes is suitable, such cases should be marked as "unknown". In the column "comments", you can suggest an additional attribute from the list of all attributes or, if none of them is suitable, your own suggestion or a paraphrase (see example for *hohe Jagd*).

2. **Disambiguate the adjective's sense given a list of GermaNet senses and their lexical id numbers.** The definitions of the senses and the id numbers are provided for each adjective in the green table on the right. You need to enter the corresponding adjective's id number for each example.

3. **If a noun has more than one sense, disambiguate the noun's sense and assign the corresponding lexical id number.** The nouns that have more than one sense in GermaNet should be disambiguated: the number of senses is given in round brackets next to the noun, e.g. "Stimme(4)". In such cases, you are required to look up the noun in GermaNet and choose the correct lexical id number and the corresponding semantic class. For example, for *hohe Stimme*: NN lex unit=40408, NN sem class=Kommunikation. Please, make sure that you enter the lexical unit id number and not the synset id.

Annotation example for the adjective "hoch". Everything written in red - what the annotator is expected to enter, the rest is done (semi-)automatically.

attributes
Dimension
Intensität
Ton
Quantität

Adjective's GermaNet senses and lex_unit_id
bezeichnet die Ausdehnung von unten nach oben senkrecht zum Boden: 3478
von beträchtlichem Ausmaß: 3247
mit großer Schwingungszahl, von Tönen: 145154
intensiv, hochgradig: 145123

adj	nn	attribute	ADJ_lex_unit	context	NN_lex_unit	NN_sem_class	Comments/suggestions
hoch	Stimme(4)	Ton	145154	Sie redete mit einer affektiert **hohen Stimme** auf ihn ein.	40408	Kommunikation	
hoch	Summe(2)	Quantität	3247	Die **hohe Summe** kam zustande, weil es 17 Wochen lang in dieser Klasse keinen Gewinner gegeben hatte.	20289	Besitz	
hoch	Miete(4)	Quantität	3247	In Wahrheit litt die große Masse der Arbeiter schwer unter Wohnungsnot, **hohen Mieten** und hohen Lebensmittelpreisen.	21302	Besitz	
hoch	Interesse(3)	Intensität	145123	Es besteht ein **hohes Interesse** an neuen Entwicklungen und möglichen Verbesserungen	19591	Attribut	
hoch	Druck(5)	Intensität	145123	Grund dafür sei der **höhere Druck**, der auf Beschäftigungslose ausgeübt werde.	19273	Attribut	
hoch	Mehrwertsteuer*	Quantität	3247	Ein Problem bei der Gesamtkalkulation sei die **hohe Mehrwertsteuer**, argumentieren sie.	21097	Besitz	
hoch	Jagd(2)	special	non	Rot-, Dam-, Reh- und Schwarzwild, zur **hohen Jagd** gehörig (Schal-, Hufwild); Hase, Kaninchen zur niederen Jagd.	23989	Geschehen	type? (das Jagen von Hochwild)

* if a noun has only one sense in GermaNet, its lexical unit id and the corresponding semantic class are assigned automatically

	Attribut	**Definition**	***Beispiele***
1.	Alter	Anzahl der durchlebten Jahre (der Lebewesen oder auch Gegenstände)	*junge Frau, neues Auto*
2.	Art	eine Gruppe aus einer Gesamtheit, die sich durch besondere Merkmale auszeichnet, eine Sorte von etwas, könnte als hyponym des Basis-Substantives betrachtet werden	*"grüner Tee" ist eine "Sorte" von Tee*
3.	Aufwand	die benötigte Energie oder Arbeit, um eine Handlung durchzuführen oder etwas herzustellen/einzurichten	*schwerer Beruf, spartanische Einrichtung*
4.	Aussehen	das äußere Erscheinungsbild	*unrasierter Mann, schicker Anzug*
5.	Belebtheit	Zustand eines Lebewesens; am Leben oder nicht	*lebloser Körper*
6.	Beliebtheit	allgemein gern gesehen und geschätzt	*angesehene Familie, beliebter Lehrer*
7.	Beschaffenheit	die Struktur, Textur einer Sache	*körniger Sand, seidenes Kleid*
8.	Bewegung	das Bewegen, die Ortsveränderung, Lageveränderung	*hektisches Blättern, fließendes Wasser*
9.	Bewertung	Vorgang, der den Wert von etwas abschätzt	*positiver Einfluss, falsche Antwort*
10.	Beziehung	Verhältnis, in dem Menschen oder Organisationen zueinander stehen	*volksfeindliche Politik, umweltfreundliche Produkte*
11.	Bezug	wechselseitiges Verhältnis zwischen beliebigen Objekten	*anschauliche Darstellung, zufällige Entdeckung*
12.	Charakter	Gesamtheit der individuellen geistig-seelischen Eigenschaften eines Menschen	*aggressiver Hund, faule Menschen*
13.	Chemie	chemische Zusammensetzung und Eigenschaften der Stoffe	*organische Säure, jodhaltige Salbe*
14.	Dauer	zeitliche Erstreckung von etwas	*langjährige Zusammenarbeit*
15.	Dimension	Ausdehnung, Abmessung eines Körpers nach Länge, Breite und Höhe	*breite Straße, lange Hosen*
16.	Ertrag	Menge der erzeugten Güter, Gewinn	*einträgliches Geschäft, ergiebige Grube*
17.	Existenz	Dasein, Vorhandensein	*anwesende Person, reale Welt*

18.	Farbe	durch Lichtstrahlen bestimmter Wellenlänge hervorgerufene Erscheinung vor dem Auge	*rotes Kleid, blaue Wand*
19.	Feuchtigkeit	Vorhandensein von Wasser (oder anderen Flüssigkeiten) in geringen Mengen	*nasses Tuch, knochentrockener Boden*
20.	Form	sichtbare Art der äußeren Gestaltung, Erscheinungsart	*eckiger Tisch, herzförmiges Blatt*
21.	Funktionalität	die Fähigkeit eines Produktes oder einer Komponente, bestimmte Aufgaben zu lösen	*unbrauchbares Gerät, kaputtes Fahrrad*
22.	Gefühl_Auslöser	Gefühl oder Emotion, die durch externe Faktoren ausgelöst wird	*trauriger Film, angenehme Nachricht*
23.	Gefühl_Empfindung	Gefühl oder Emotion, die eine Person selbst erlebt	*glückliches Kind, trauriger Mann*
24.	Genauigkeit	mit einem Muster, Vorbild, einer Vergleichsgröße [bis in die Einzelheiten] übereinstimmend; einwandfrei stimmend, exakt	*genaue Zahl, exakte Messung*
25.	Geruch	Art, wie etwas riecht	*aromatischer Duft, muffige Turnhalle*
26.	Geschicklichkeit	die Fähigkeit, eine Sache schnell und ohne großen Aufwand zu erledigen	*geschickter Kletterer, unbeholfene Kinder*
27.	Geschlecht	Gesamtheit biologischer Merkmale, die ein Lebewesen als weiblichen, männlichen oder intersexuellen Typus charakterisiert	*männliches Tier, transgender Person*
28.	Geschmack	Art, wie etwas schmeckt	*leckerer Kuchen, salzige Brühe*
29.	Geschwindigkeit	Schnelligkeit, Tempo	*schnelle Bewegung, langsamer Prozess*
30.	Gesellschaft_allgemein	Gesamtheit der Menschen, die zusammen unter bestimmten politischen, wirtschaftlichen und sozialen Verhältnissen leben	*traditionelles Fest, einheimische Bevölkerung*
31.	Gesellschaft_Land	betrifft ein souveränes, von Grenzen umgebenes Gebiet, Staatsgebiet, Staat	*europäischer Markt, deutsche Firma*
32.	Gesellschaft_politisch	politische Regulierungen, Beziehung zwischen Staat und Bevölkerung,	*demokratische Abstimmung, private Schule*

33.	Gesellschaft_Region	betrrifft ein Bereich beziehungsweise ein Gebiet, das eine Einheit bildet und mehrere Städte und Gemeinden umfasst	*städtische Infrastruktur, bundesweites Verbot*
34.	Gesellschaft_Religion	betrifft einen durch Lehre und Satzungen festgelegten Glaube	*gläubige Familie, christliche Lehre*
35.	Gesellschaft_Status	beschreibt den Wohlstand von Personen/ deren Stand in einer Gesellschaft	*arbeitslose Akademiker, adlige Familie*
36.	Gesundheit	körperliches Befinden	*gesundes Kind, schwangere Frau*
37.	Gewicht	durch Wiegen ermittelte Masse, Last eines Körpers	*schwerer Koffer*
38.	Gewissheit	ohne Zweifel; Klarheit, sicheres Wissen	*voraussehbarer Ausgang, eindeutige Absage*
39.	Gewohnheit	Über längere Zeiträume ähnlich ablaufenden Ereignissen (nicht) entsprechend	*typische Symptome, unübliche Methode*
40.	Gültigkeit	rechtlich, gesetzlich o. ä. anerkannt und entsprechend wirksam	*gültiger Vertrag, zulässiges Gesamtgewicht*
41.	Helligkeit	Vorhandensein von Licht oder wenig intensive Färbung	*grelle Blitze, schummrige Bar*
42.	Herkunft	woher etwas oder jemand kommt	*einheimischer Vogel, fremde Person*
43.	Intelligenz	geistige Fähigkeit, Klugheit	*kluges Kind, intelligente Frau*
44.	Intensität	Grad, Stärke. Normalerweise wird dabei eine schon bestehende Eigenschaft des Substantives verstärkt/abgeschwächt	*große Angst, schwacher Trost*
45.	Klima	typische Art und wiederkehrender Ablauf der Witterung	*trockener Sommer, klarer Himmel*
46.	Konsistenz	Beschaffenheit der Dichte, Festigkeit	*elastische Schnur, flüssige Treibstoffe*
47.	Konstitution	körperliche Verfassung eines Lebewesens	*kräftiger Mann, schwaches Individuum*
48.	Körpergefühl	Wahrnehmung des Körpers	*hungriges Kind, müde Katze*
49.	Körperzustand	Art und Weise, wie ein Körper zu einem bestimmten Zeitpunkt ist	*schwangere Frau, geschlechtsreifes Tier*
50.	Kosten	die Ausgaben, der Aufwand an Geld	*billiges Kleid, kostenlose Beratung*

51.	Landschaft	Teil der Erdoberfläche, der durch Bodengestalt, Bewachsung, Besiedlung sein besonderes Gepräge erhalten hat	*hügelige Gegend, flaches Land*
52.	Lokalisation	Ortsbestimmung, Zuordnung zu einer bestimmten Stelle	*äußere Wand, ferne Insel*
53.	Neigung	Veranlagung, Hang zu etw.	*gesprächiger Mensch, konsumfreudiges Publikum*
54.	Norm	Eigenschaft, die den Erwartungen (nicht) entspricht	*sonderbares Gefühl, groteske Figur*
55.	Oberfläche	flache Außenseite eines Körpers	*glatter Marmor, rutschiges Parkett*
56.	Ordnung	aufgeräumter oder organisierter Zustand	*unaufgeräumtes Kinderzimmer, strukturierter Tagesablauf*
57.	Pertonym	abgeleitet vom Substantiv und gehört zu der entsprechenden semantischen Klasse	*theatralische Inszenierung, sandiger Feldweg*
58.	Physik	durch die Gesetze der Physik beschrieben	*elektrische Nähmaschine, magnetische Kopplung*
59.	Quantität	Menge, Anzahl, Umfang	*zahllose Anekdoten, ausreichende Geldmittel*
60.	Raum_Zustand	der Zustand eines dreidimensionalen, geschlossenen oder offenen Raumes	*leere Straße, ausgebuchter Konferenzraum*
61.	Reife von Nahrung	Vollendung eines Wachstumsprozesses	*reife Trauben, frisches Gemüse*
62.	Richtung	Seite, nach der jmd., etw. gerichtet ist	*südliche Grenze, rechtes Ufer*
63.	Sauberkeit	sauberer Zustand	*schmutzige Hände*
64.	Sicherheit	Sache, Person oder Situation, die die Sicherheit von Personen gewährleistet/gefährdet	*gefährlicher Verbrecher, harmloser Virus*
65.	Sicht	Fähigkeit, mit den Augen Lichtspektren zu empfangen, diese ans Gehirn weiterzusenden und dann zu verarbeiten	*sichtbare Konturen, undeutliche Schrift*
66.	Stabilität	Haltbarkeit gegenüber großer Beanspruchung	*tragfähige Brücke, wacklige Konstruktion*
67.	Strenge	sorgfältige, genau der Vorschrift entsprechende Regelung	*rigide Sparpolitik, flexible Regelung*
68.	Tätigkeit	betrifft Tätigkeiten, Aufgaben, Verantwortungen	*verantwortlicher Mitarbeiter, disponible Arbeitskräfte*
69.	Temperatur	messbarer Wärmegrad eines Stoffes	*warme Suppe, kalter Boden*
70.	Ton	mit dem Ohr wahrnehmbare akustische Schwingungen	*leise Musik, stimmhafter Laut*

71.	Übereinstimmung	eine Gleichheit bestimmter Werte oder Fakten	*gleiche Höhe, diverse Projekte*
72.	Ursache	vorhergehender Sachverhalt, der einen momentanen Sachverhalt wesentlich bestimmt hat	*wetterbedingte Absage, hausgemachte Pasta*
73.	Verhalten_Tier	beschreibt die Art, Aussehen oder Merkmale eines Tieres	*zahmes Eichhörnchen, nachtaktive Eule*
74.	Verknüpfung	ein engerer Zusammenhang (zwischen mehreren Dingen, Tätigkeiten, Sachverhalten o. Ä.)	*trennbare Einheiten, wechselseitige Sympathie*
75.	Verträglichkeit	positive oder negative Aufnahme und Verarbeitung von Nahrungsmitteln, Substanzen und Situationen	*schädliche Stoffe, unbedenkliche Situation*
76.	Vollständigkeit	ein Bestand ohne Lücken, das Vorhandensein alles Dazugehörenden	*fertiger Kuchen, abgeschlossenes Studium*
77.	Wichtigkeit	Eigenschaft, bedeutsam, wichtig zu sein	*bedeutsame Entdeckung, brennende Frage*
78.	Wirkung	von einer Ursache, einem Verursacher ausgehende Beeinflussung, hervorgebrachte Folgen, erzieltes Ergebnis	*manipulativer Trick, wirksame Therapie*
79.	Wissen/Erfahrung	Kenntnis, die jemand durch wiederholte Praxis in einem bestimmten Gebiet bekommt	*unwissendes Kind, begabte Sängerin*
80.	Zeit	bestimmter Moment im Ablauf von Zeiteinheiten	*plötzliche Wende, früher Morgen*
81.	Zustand	zu einem bestimmten Zeitpunkt gegebene Lage, Verfassung, in der sich jmd., etw. befindet	*rohes Ei, geschlossene Tür*

Appendix D

Annotation templates for semantic classes

FIGURE D.1: Annotation template: semantic class ANIMAL.

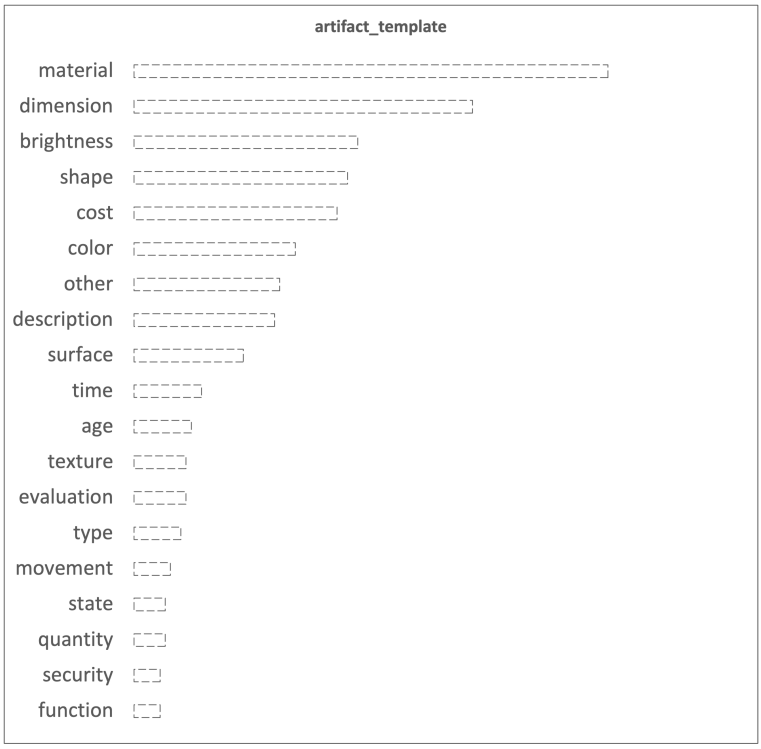


FIGURE D.2: Annotation template: semantic class ARTIFACT.

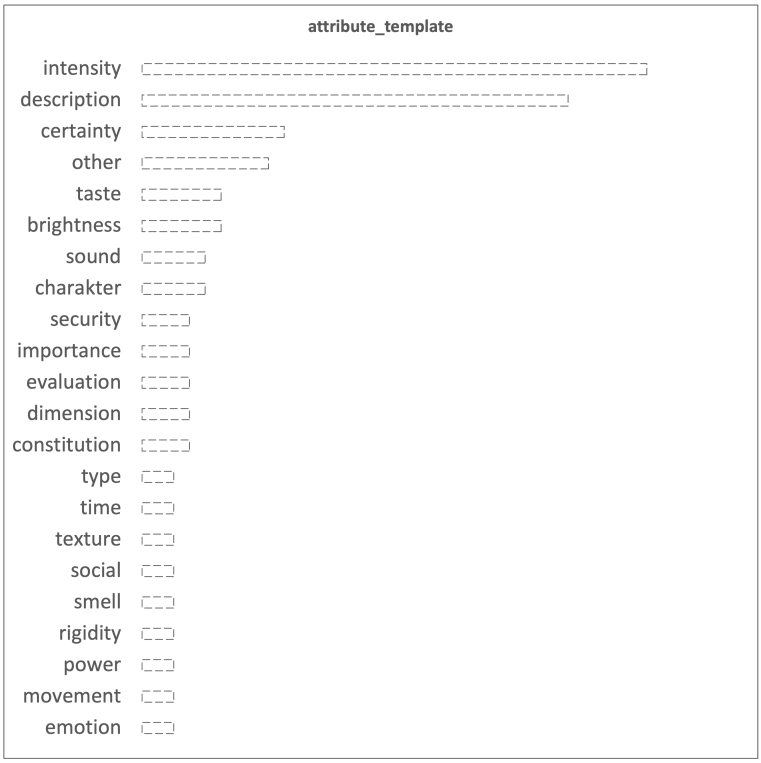


FIGURE D.3: Annotation template: semantic class ATTRIBUTE.

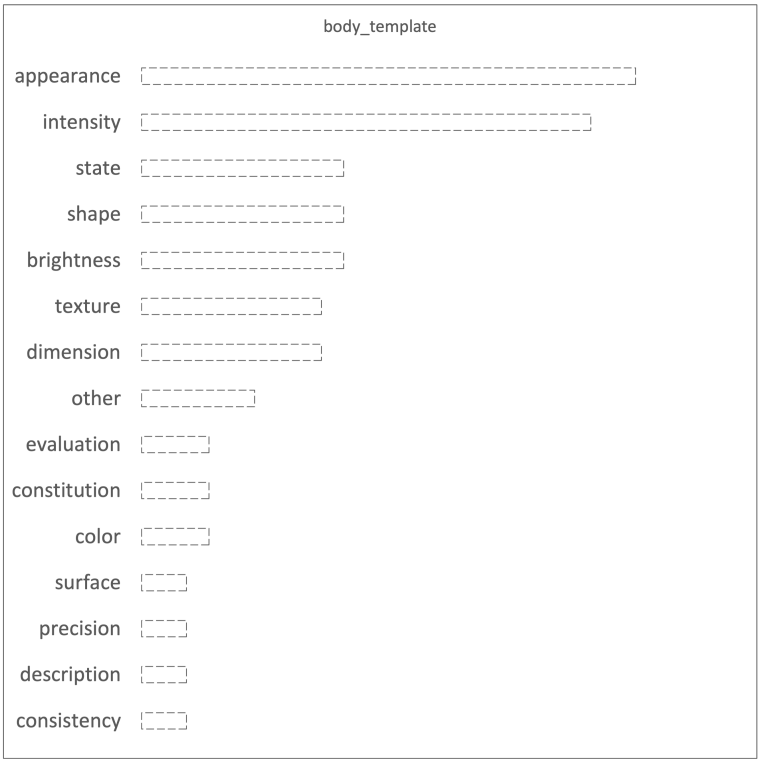


FIGURE D.4: Annotation template: semantic class BODY.

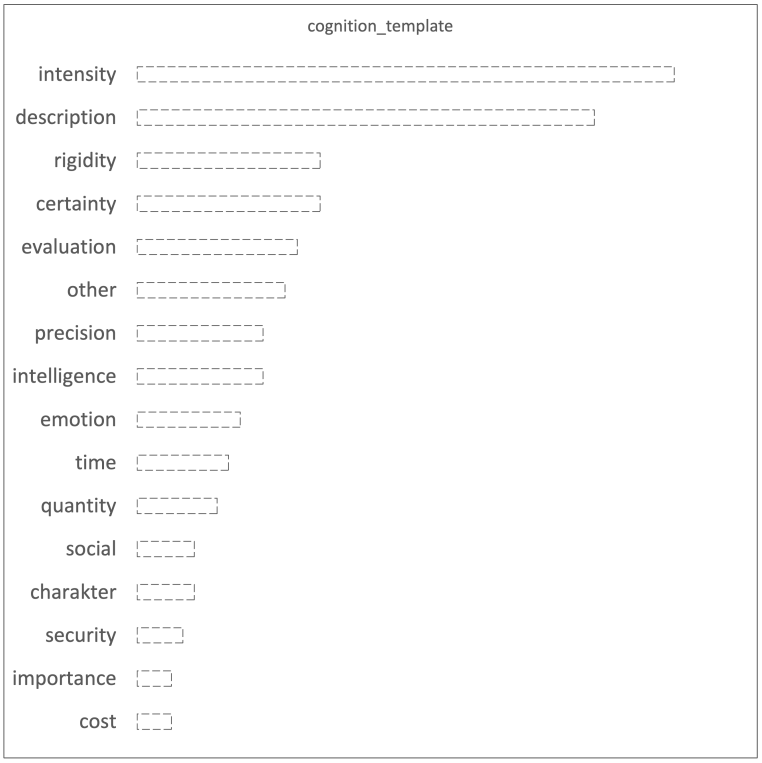


FIGURE D.5: Annotation template: semantic class COGNITION.

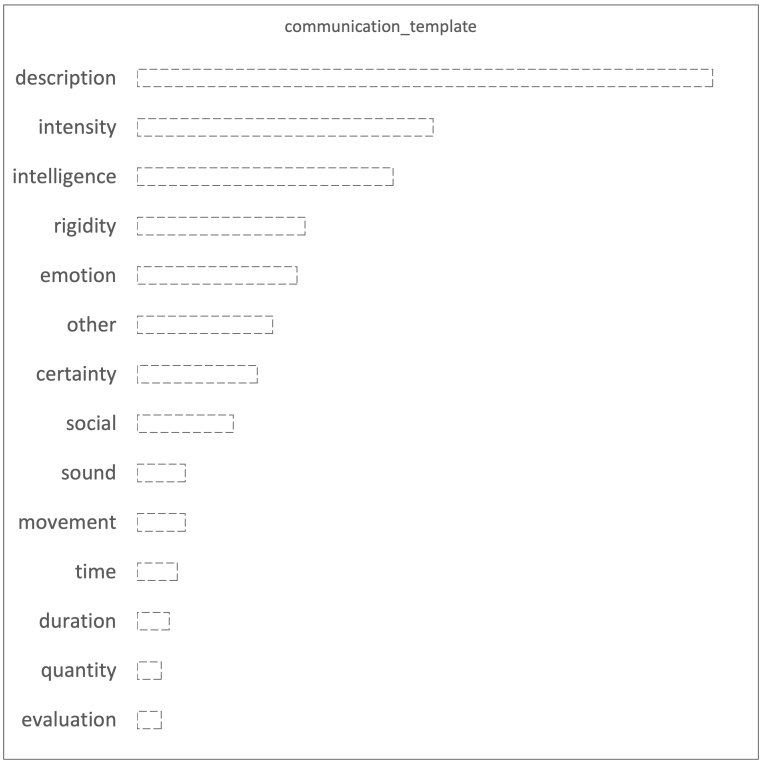


FIGURE D.6: Annotation template: semantic class COMMUNICATION.

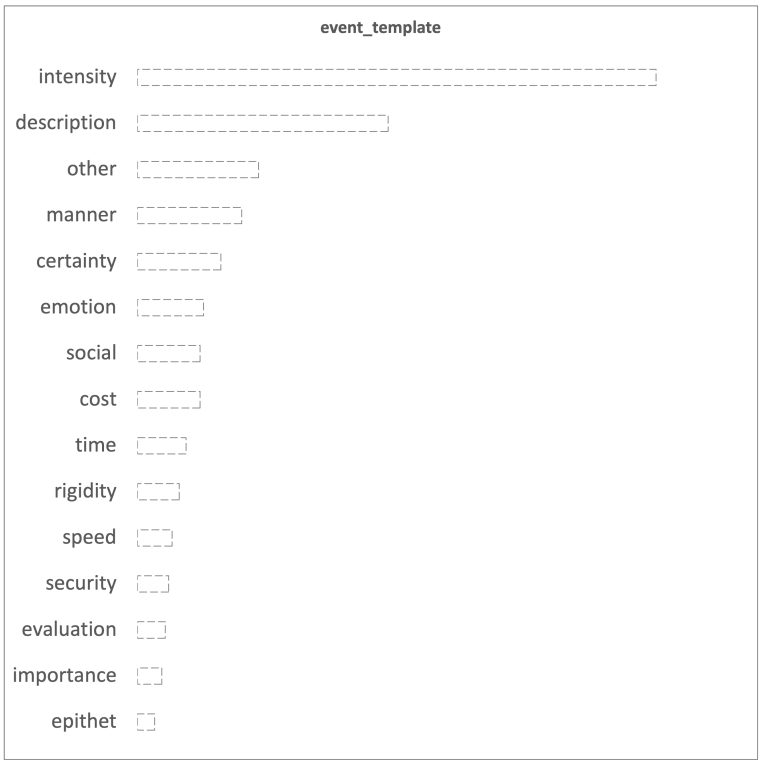


FIGURE D.7: Annotation template: semantic class EVENT.

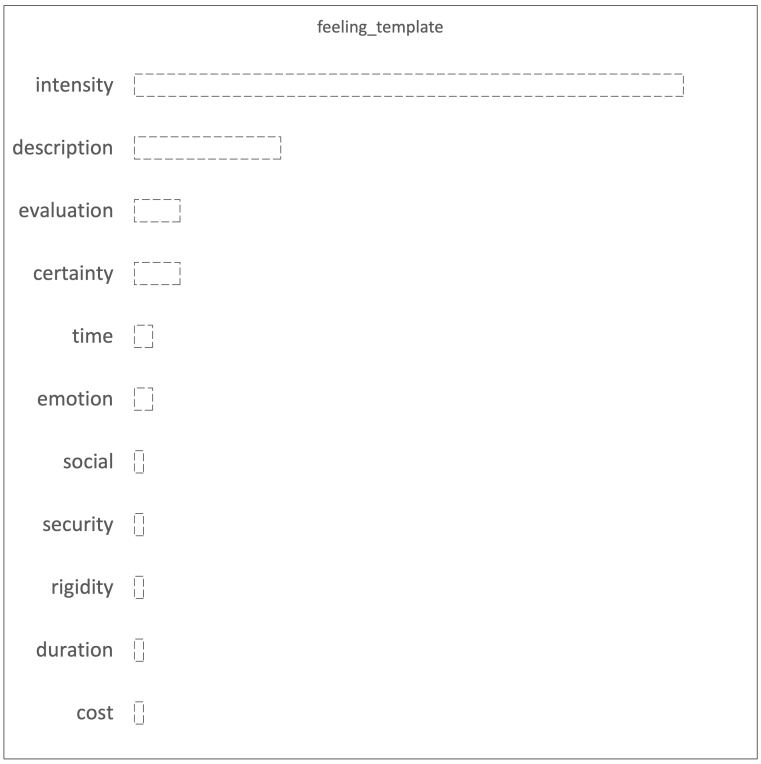


FIGURE D.8: Annotation template: semantic class FEELING.

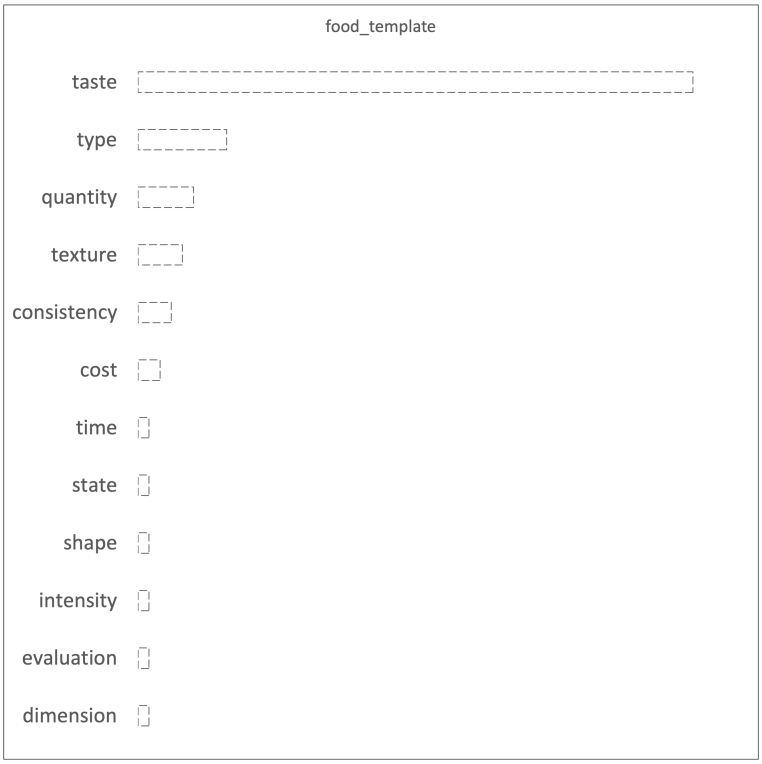


FIGURE D.9: Annotation template: semantic class FOOD.

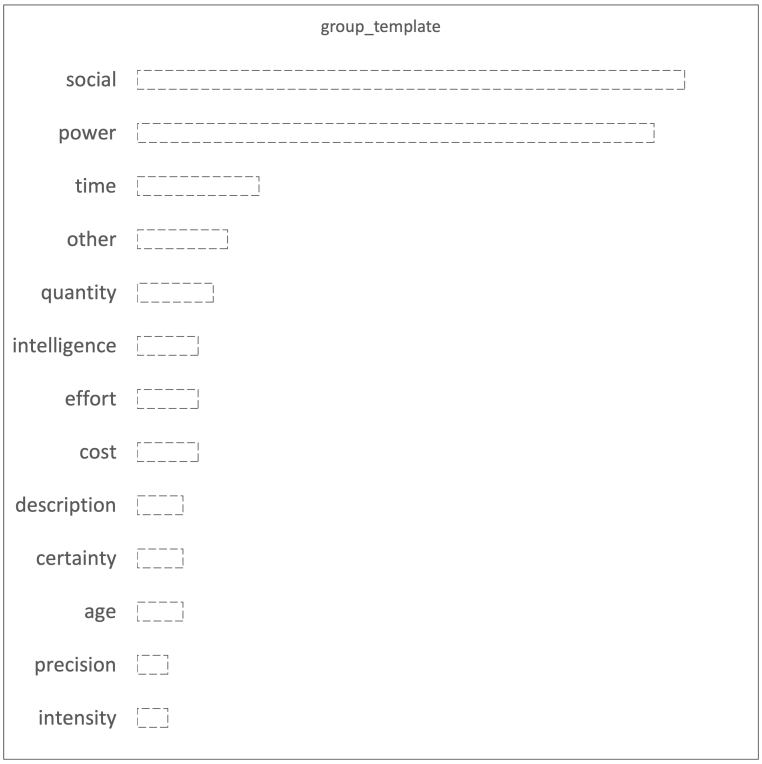


FIGURE D.10: Annotation template: semantic class GROUP.

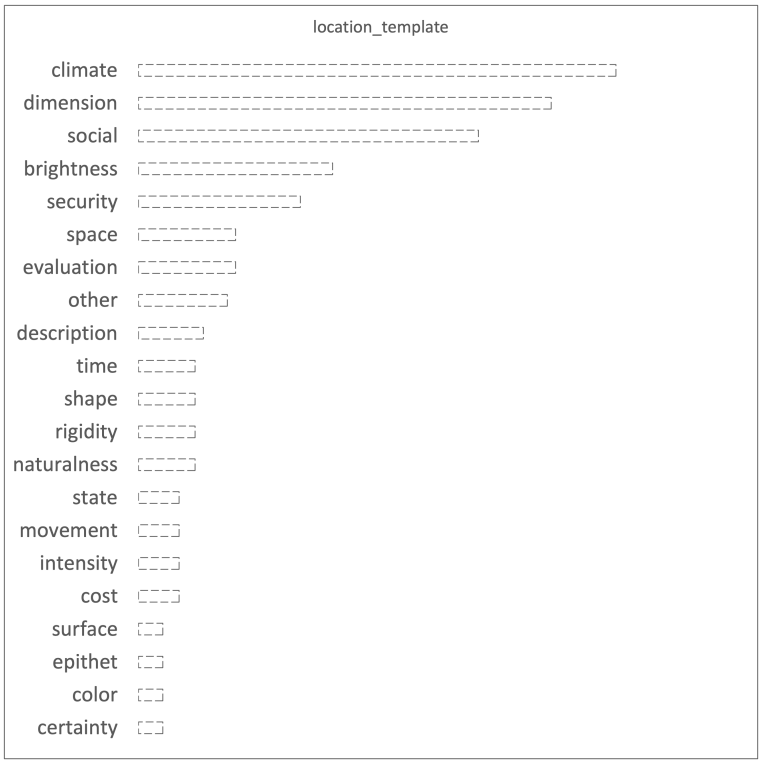


FIGURE D.11: Annotation template: semantic class LOCATION.

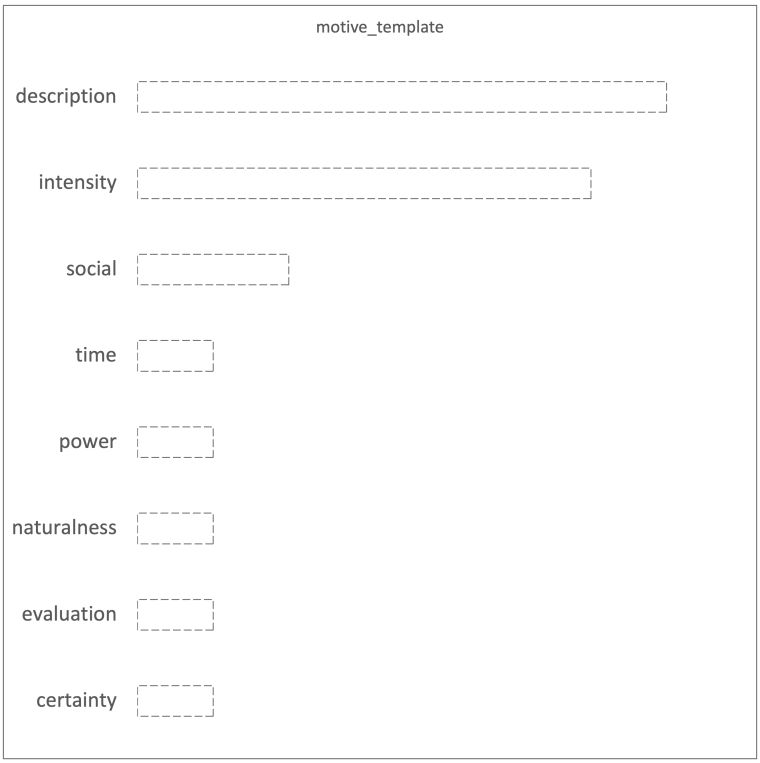


FIGURE D.12: Annotation template: semantic class MOTIVE.

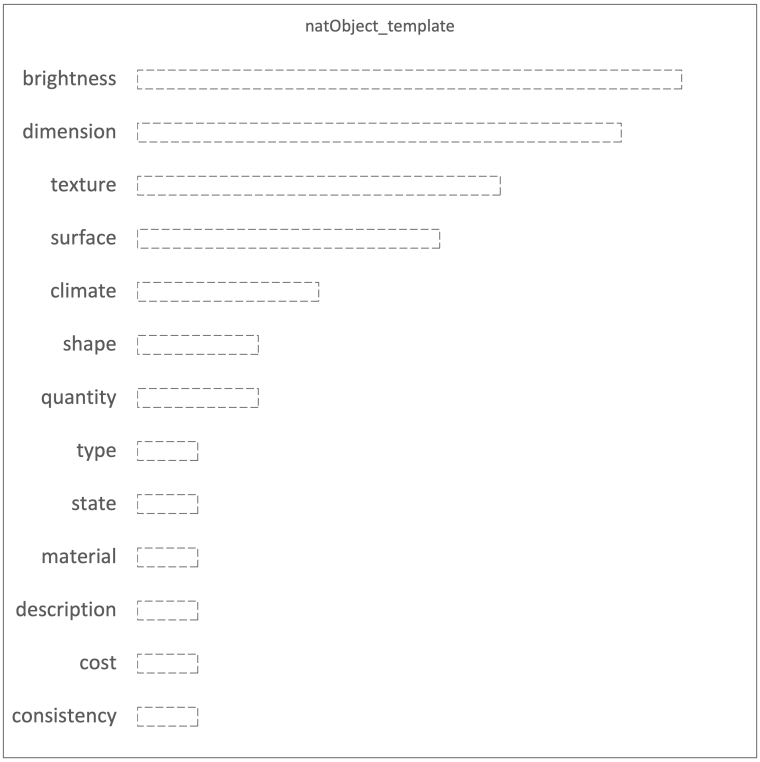


FIGURE D.13: Annotation template: semantic class NATOBJECT.

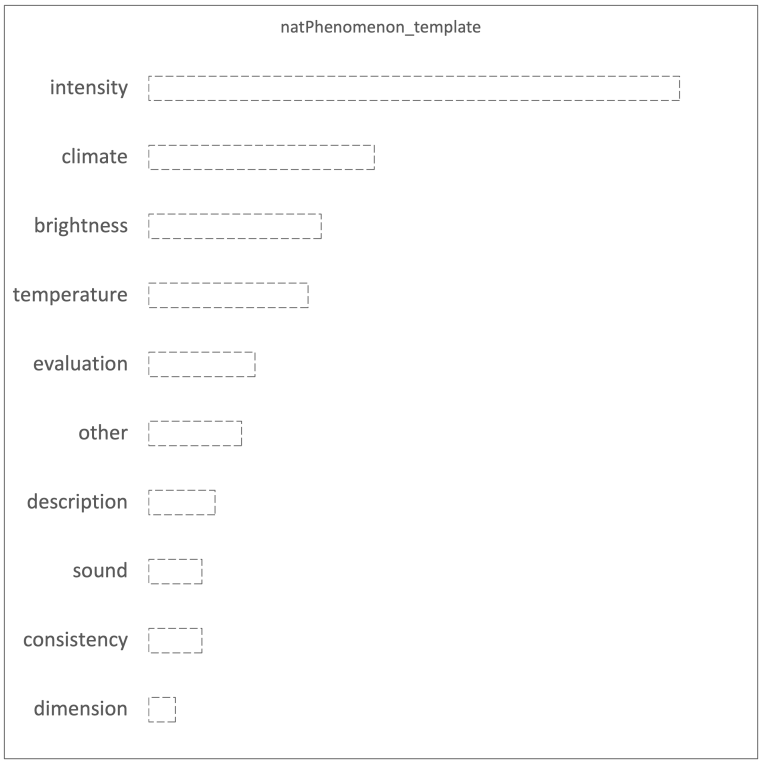


FIGURE D.14: Annotation template: semantic class NATPHENOMENON.

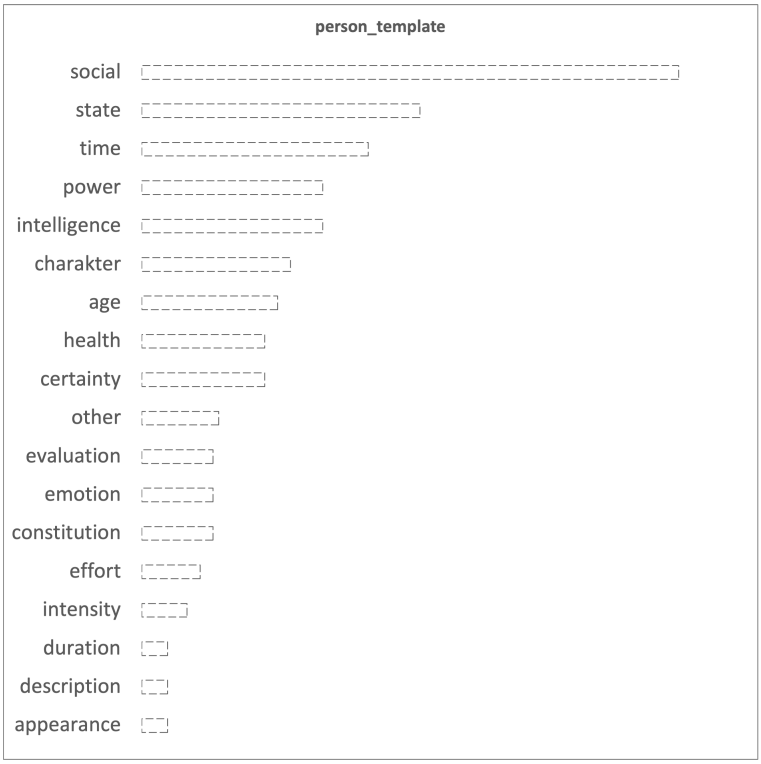


FIGURE D.15: Annotation template: semantic class PERSON.

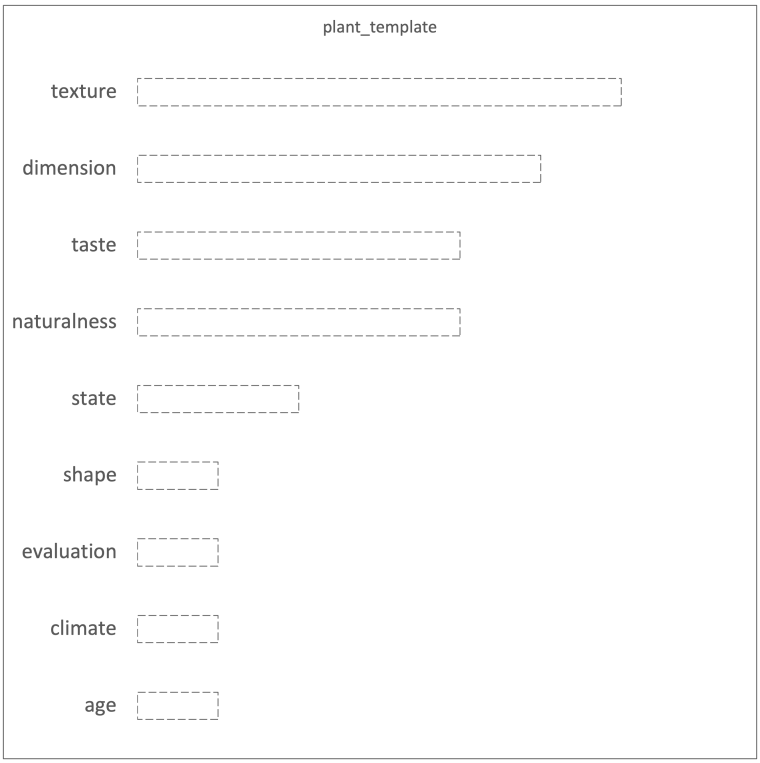


FIGURE D.16: Annotation template: semantic class PLANT.

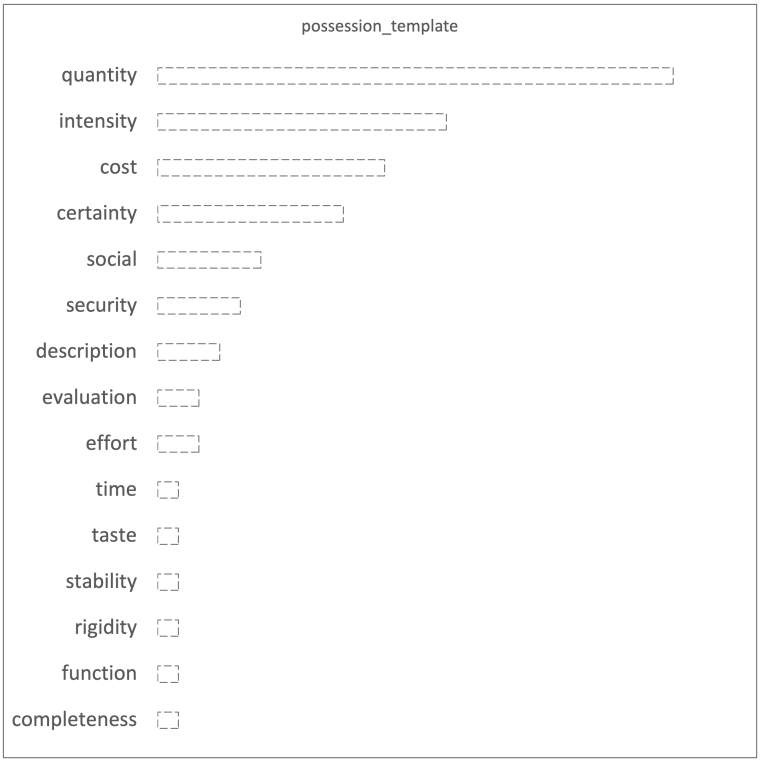


FIGURE D.17: Annotation template: semantic class POSSESSION.

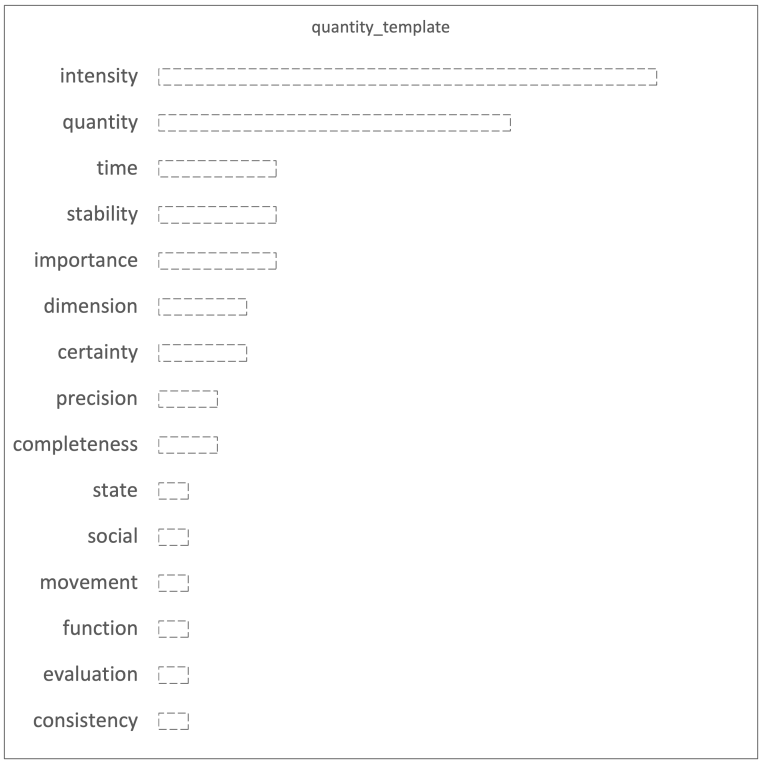


FIGURE D.18: Annotation template: semantic class QUANTITY.

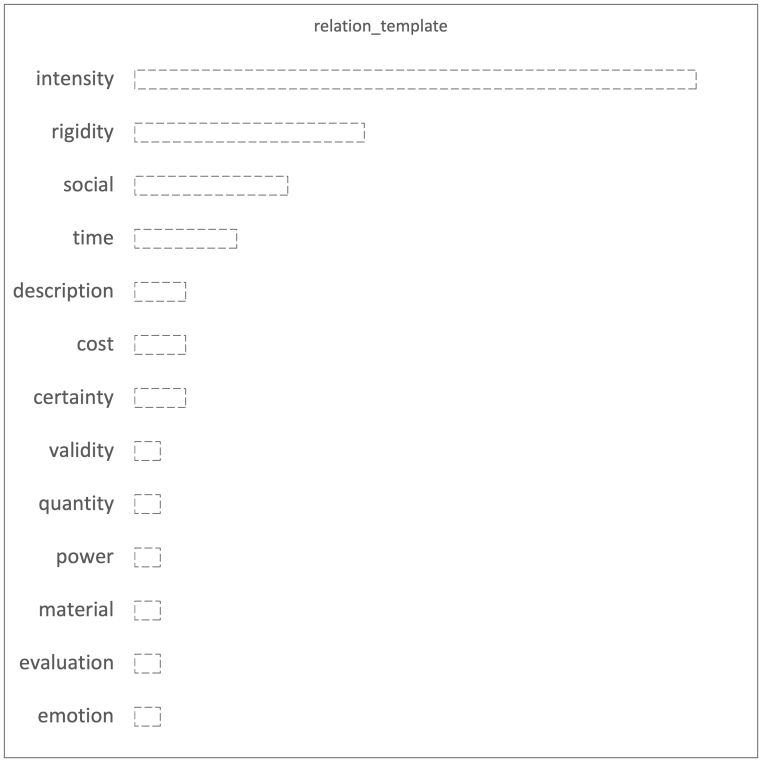


FIGURE D.19: Annotation template: semantic class RELATION.

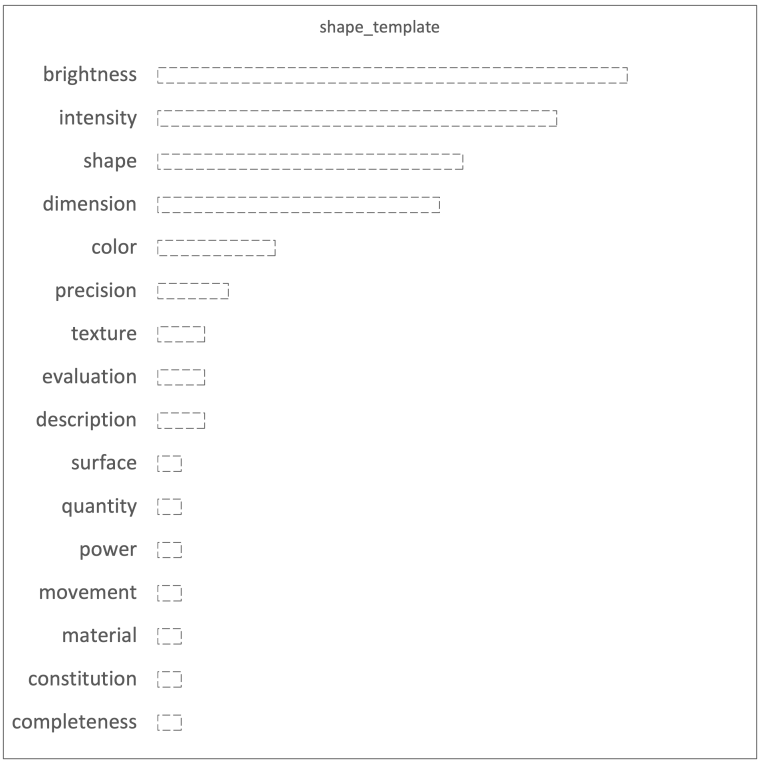


FIGURE D.20: Annotation template: semantic class SHAPE.

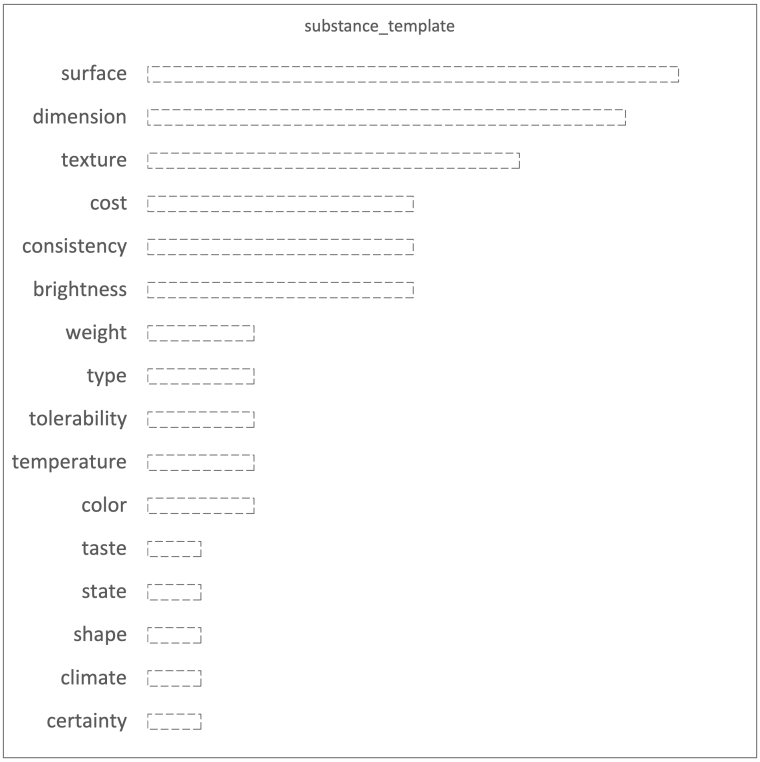


FIGURE D.21: Annotation template: semantic class SUBSTANCE.

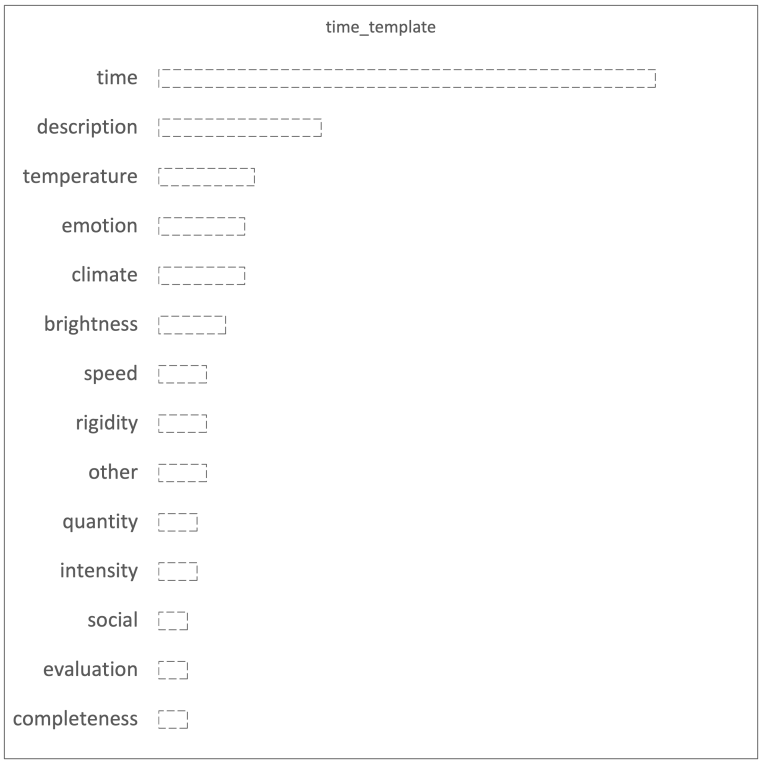


FIGURE D.22: Annotation template: semantic class TIME.

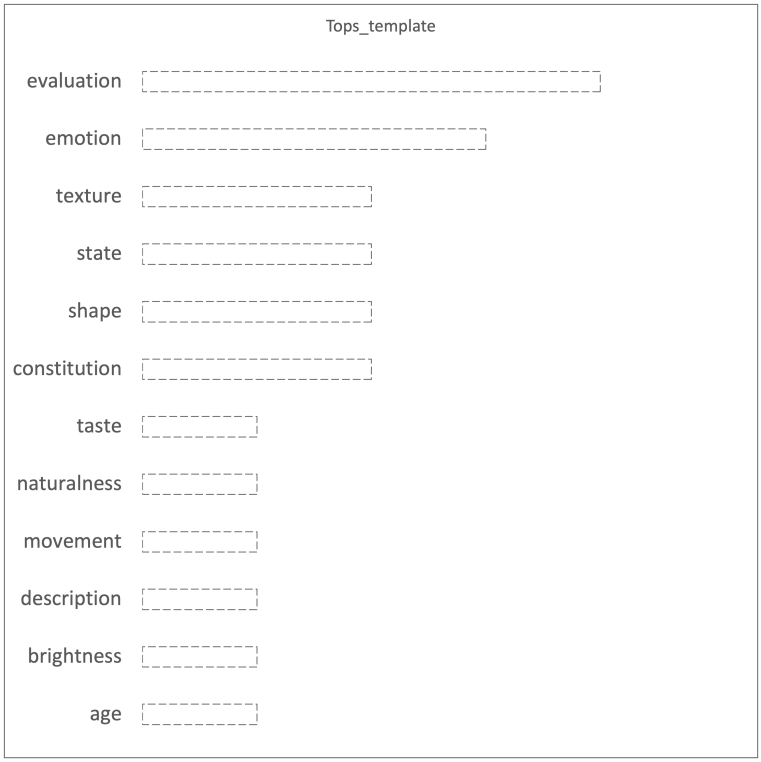


FIGURE D.23: Annotation template: semantic class TOPS.

Appendix E

Inventory of relations in adjective-noun phrases

The next pages contain the entire inventory of attributes created based on the CerCoAt and HeiPLAS+ datasets. The definitions of the attributes in part stem from the glosses in the Princeton WordNet[1], they are marked as (WN).

[1] http://wordnetweb.princeton.edu/perl/webwn

	attribute	definition	examples	comment
1	age	number of lived years	*old lady, new car*	for living beings and inanimate objects
2	appearance	the external appearance of a living being	*slim child, blond student*	
3	brightness	presence of light or intensity of coloring	*dark room, light-colored dress*	
4	cause	refers to the origin of an entity, the factors responsible for bringing it about	*unnatural death, intentional damage*	similar to the agentive dimension in the Qualia Structure
5	certainty	clarity, certain knowledge, without doubt	*certain death, obvious error*	Epistemic modality: expresses how confident the speaker is about the existence of the noun or of some salient aspect of the noun's meaning.
6	character	characteristic property that defines the apparent individual nature of someone (WN) or their behaviour	*fierce dog, obedient child*	also includes temporal properties that refer to someone's behaviour
7	climate	weather conditions	*sultry summer, windy day*	
8	color	a visual attribute of things that results from the light they emit or transmit or reflect (WN)	*red dress, blue wall*	
9	completeness	the state of being complete and entire; having everything that is needed (WN)	*comprehensive coverage, open account*	
10	composition	refers to the chemical composition of an entity	*contaminated water, pure oxygen*	similar to the constitutive dimension in the Qualia Structure
11	consistency	density, viscosity	*thick mass, thin oil*	
12	constitution	physical makeup of a living being or an inanimate object	*delicate figure, weak pillar*	
13	continuity	steadiness and/or regularity of a process or event;	*constant pain, incessant noise*	
14	cost	the property of having material worth (WN)	*expensive car, free entrance*	
15	dimension	magnitude of an object inluding length, width, height, angle, size	*steep mountain, broad river*	
16	duration	temporal extension	*old friend, short life*	
17	effort	the energy or work required to perform or accomplish something	*easy answer, tough opponent*	semclass person: a *tough opponent* implies that dealing with the opponent requires a lot of energy or work

18	emotion	affective and emotional state	*poor animal, sad person, blue day*	the emotion can be either **triggered** by the head noun as in *poor animal: animal* triggers compassion expressed by the adjective *poor;* the emotion can be **experienced** by the head noun as in *sad person;* or have a different relation to the head noun as in *blue day: sadness* referred to by the adjective *blue* describes someone's perception of the *day*
19	epithet	clichéd expressive modifier	*humble opinion, proud parent*	the semantic load of the modifier is very low and it contributes almost nothing to the meaning of the nominal head
20	evaluation	subjective positive or negative asessment of something	*beautiful theory, sweet child*	
21	function	the use or purpose of an entity	*potent liquor, strong engine*	
22	health	condition of body and mind of a living being	*healthy person, infertile couple*	
23	importance	the quality of being important and worthy of note (WN)	*critical element, historic decision*	
24	intelligence	the ability to comprehend; to understand and profit from experience (WN)	*smart child, unintelligent employee*	
25	intensity	degree	*bad headache, light rain*	characterizes lexical units that already have a gradability element in their meaning
26	location	refers to spacial characteristics of an entity having to do with its location such as distance, direction, position;	*distant sound, exterior paint*	
27	manner	the way in which an action is carried out/performed or an entity is described	*wild dance, rough poetry*	the *how* of the nominal head
28	material	substance that goes into the makeup of a physical object (WN)	*wooden ceiling*	the property of being **made of** a material
29	movement	change of position and/or location	*rigid face, waving flag*	
30	naturalness	the property of being affected by humans such as being domesticated for animals	*untouched nature, wild animal*	

31	power	possession of controlling influence (WN)	*weak president, powerful family*	
32	precision	accuracy	*accurate measurement, rough estimation*	
33	quality	an objective or subjective estimation of whether something is correct or proper	*wrong answer, appropriate apology*	Objective qualifiers are assessments that can be verified. Subjective qualifiers are those that rather refer to some standards or expectations of how something should be
34	quantity	amount	*tight finances, generous donation*	
35	relation	puts the modified noun in context of other entities, a general comparison in terms of norm	*standard size, different approach*	
36	rigidity	property of being able to change and adjust, flexibility	*strict regulation, rigid guidelines*	
37	security	being free from danger or risk	*safe distance, dangerous situation*	
38	sex	property that distinguishes organisms on the basis of their reproductive roles	*female character, male dog*	
39	shape	spatial form, contour, outline	*sharp knife, round desk*	
40	smell	the sensation that results when olfactory receptors in the nose are stimulated by particular chemicals in gaseous form (WN)	*odorless flower, sweet smell*	
41	social	properties relating to human society such as ethnicity, nationality, occupation, popularity, wealth, etc	*unpopular idea, rich relatives*	
42	sound	the sensation of hearing something	*small voice, loud child*	
43	space	area (usually bounded in some way between things) (WN)	*empty room, panoramic view*	
44	speed	a rate at which something happens (WN) or moves	*meteoric rise, rapid development*	

45	`stability`	the quality of being enduring and free from change or variation (WN)	*frozen price, strong currency*	
46	`state`	the way something is with respect to its main attributes (WN)	*dead animal, open door*	
47	`surface`	the two-dimensional outer boundary of a three-dimensional object (WN)	*rough stone, shiny aluminium*	includes both touch and vision
48	`taste`	a sensation obtained from a substance in the mouth	*sweet bun, savoury soup*	
49	`temperature`	the degree of hotness or coldness of a body or environment (WN)	*mild evening, hotwater*	
50	`texture`	the physical composition of something (especially with respect to the size and shape of the small constituents of a substance) (WN)	*coarse sand, fine grain*	
51	`time`	temporal aspect	*late evening, old name*	
52	`type`	not a property, but a hyponymic relation between the phrase as a whole and its nominal base	*black tea, dark chocolate*	black tea ***is a type of*** tea; dark chocolate **is a type of** chocolate. Such phrases are always collocations (or compounds) as they are highly lexicalized
53	`validity`	the property of being legal	*dirty money, unlawful measure*	
54	`weight`	the property of being heavy or light	*heavy load, lightweight luggage*	

Appendix F

Frames for the bases in the sample

F.1 Aggregate

Definition: This frame contains nouns denoting Aggregates of Individuals. The Aggregates may be described by an Aggregate_property.

Core FEs:

- **Aggregate** - This FE identifies the collection of entities that are the Aggregate.
- **Individuals** - This FE identifies the people or things that form the Aggregate.

Non-Core:

- **Aggregate_property** - This FE describes a property of the Aggregate rather than the Individuals and is used to tag the adjective that modifies the aggregate noun when it is unambiguous.
- **Container_possessor** - The location or political locale that the Organization is contained within.
- **Domain** - The relevant Domain as it pertains to the Aggregate.
- **Name** - The term used to refer to the Aggregate.

F.2 Animals

Definition: The LUs in this frame refer to biological entities other than human beings and plants, which are labeled by the FE Animal.

Core FEs:

- **Animal** - The living organism belonging to the kingdom Animalia.

Non-Core:

- **Characteristic** - The living organism belonging to the kingdom Animalia.

- **Descriptor**- A characterization of the Animal not covered by another frame element.
- **Origin** - The location from which the Animal originated.

F.3 Biological area

Definition: This frame contains words that denote large ecological areas as well as smaller locations characterized by the type of life present–in other words, geography locations as defined by biota.

Core FEs:

- **Locale** - This FE identifies a stable bounded area, and is typically the designation of the nouns of Locale-derived frames.

Non-Core:

- **Constituent_parts** - Constituent_parts make up a Locale.
- **Container_possessor** - Container_possessor is the location of which the Locale is a part.
- **Descriptor** - Any description, characteristic, or property of the Locale which is not covered by more specific FEs.
- **Formational_cause** - Formational_cause brings about the features of the Locale.
- **Name** - This FE indicates the Names of Locales.
- **Relative_location** - Relative_location identifies a place in respect to which a Locale is located.

F.4 Body parts

Definition: This frame covers words for Body_parts belonging to a Possessor, which may be characterized by a Descriptor. The location of the Body_part may be identified in terms of its Attachment or its Orientational_location. A Subregion of a Body_part may also be indicated.

Core FEs:

- **Body_part** - This Frame Element is the Body_part.
- **Possessor** - This Frame Element denotes the Possessor of the Body_part.

Non-Core:

- **Attachment** - The Attachment is the larger body part to which the BP is attached.
- **Orientational_location** - This Frame Element describes the location of the Body_part with respect to the inherent orientation of the body (or the larger body part) the Body_part is attached to.
- **Descriptor** - This Frame Element is used for a characteristic or description of the Body_part. Note that Descriptor is not an FE of the same character as other FEs. Specifically, rather than filling a slot, it identifies a modifier of its head noun.
- **Subregion** - This FEidentifies the Subregion of the Body_part referred to by the target.

F.5 Calendric unit

Definition: Words in this frame name the different parts of the calendric cycle, both man-made and natural. The Unit (e.g. Tuesday) specifies some time period as part of a specific larger temporal Whole (Tuesday of next week), or may be resolved to an exact time span by a Relative_time (next Tuesday).[1]

Core FEs:

- **Relative_time** - Relative_time is used for the word or words that locate the time with respect to an identifiable reference point.
- **Unit** - The calendric unit that is denoted. LUs in this frame all denote the Unit.
- **Whole** - Whole is used for the larger unit of which the target word names a part.

Non-Core:

- **Count** - The number of intervals of the specified type.
- **Landmark_period** - The time-expression that is used to establish when the Trajector_event occurred.
- **Name** - Name is used for the name of a day (month, etc.). Note that only the word (or words) that is the name is (are) tagged, not the target word.
- **Salient_event** - An event saliently associated with a particular calendric unit.
- **Trajector_event** - An event that is located by the Calendric_unit expression.

[1]The definition in FrameNet is much longer. For the complete definition, consult FrameNet.

F.6 Clothing

Definition: This frame refers to clothing and its characteristics, including anything that people conventionally wear.

Core FEs:

- **Garment** - This FE identifes the clothing worn.

Non-Core:

- **Body_location** - Body_location identifies the place on the body where the clothing is worn.
- **Creator** - The individual who created the Garment.
- **Descriptor** - This FE identifes any description of the clothing besides Style and Material.
- **Material** - This FE identifes the material from which the clothing is made.
- **Name** - The name given to a specific Garment. Note this is not the name of a general class of clothing items, but rather an individual item.
- **Style** - This FE identifes the design, or cut, of the clothing.
- **Subregion** - This FE identifies a part of the clothing.
- **Time_of_creation** - The time at which the Creator makes the Garment.
- **Use** - This FE identifes the Use for which a Garment is worn or was made.
- **Wearer** - The Wearer is the person wearing the clothes, or for whom the clothes are intended.

F.7 Commerce scenario

Definition: Commerce is a situation in which a Buyer and a Seller have agreed upon an exchange of Money and Goods (possibly after a negotiation), and then perform the exchange, optionally carrying it out with various kinds of direct payment or financing or the giving of change.

Core FEs:

- **Buyer** - The Buyer has the Money and wants the Goods.
- **Goods** - Goods is anything including labor or time, for example, which is exchanged for Money in a transaction.

- **Money** - Money is given in exchange for Goods in a transaction.
- **Seller** - The Seller has the Goods and wants the Money.

Non-Core:

- **Manner** - Manner of performing an action
- **Means** - The means by which a commercial transaction occurs.
- **Purpose** - A state of affairs that the agent intends to bring about as a result of participating in the Commercial Transaction.
- **Rate** - In some cases, price or payment is described per unit of Goods.
- **Unit** - This FE is any unit in which goods or services can be measured. Generally, it occurs in a by-PP.

F.8 Emotion directed

Definition: The adjectives and nouns in this frame describe an Experiencer who is feeling or experiencing a particular emotional response to a Stimulus or about a Topic. There can also be a Circumstances under which the response occurs or a Reason that the Stimulus evokes the particular response in the Experiencer.

Core FEs:

- **Event** - The Event is the occasion or happening that Experiencers in a certain emotional state participate in .
- **Experiencer** -The Experiencer is the person or sentient entity that experiences or feels the emotions.
- **Expressor** - The Frame Element Expressor marks expressions that indicate a body part, gesture or other expression of the Experiencer that reflects his or her emotional state. They describe a presentation of the experience or emotion denoted by the adjective or noun.
- **Reason** - The Reason is the explanation for why the Stimulus evokes a certain emotional response.
- **State** - The State is the abstract noun that describes a more lasting experience by the Experiencer.
- **Stimulus** - The Stimulus is the person, event, or state of affairs that evokes the emotional response in the Experiencer.
- **Topic** - The Topic is the general area in which the emotion occurs. It indicates a range of possible Stimulus.

Non-Core:

- **Circumstances** - The Circumstances is the condition(s) under which the Stimulus evokes its response. In some cases it may appear without an explicit Stimulus.
- **Degree** - The Degree is the degree to which the Experiencer feels the emotion.
- **Empathy_target** - The Empathy_target is the individual or individuals with which the Experiencer identifies emotionally and thus shares their emotional response.
- **Frequency**
- **Manner** - The Manner is the way in which the Experiencer experiences the Stimulus.
- **Parameter** - The Parameter is a domain in which the Experiencer experiences the Stimulus.

F.9 Fear

Definition: An Experiencer, Expressor, or State can be described as characterized as having an emotion of fear concerning a particular Topic or as evoked by a Stimulus.

Core FEs:

- **Experiencer** - The Experiencer is the person or sentient entity that experiences or feels the emotions.
- **Expressor** - The body part, gesture, or other expression of the Experiencer that reflects his or her emotional state. They describe a presentation of the experience or emotion denoted by the adjective or noun.
- **State** - The State is the abstract noun that describes a more lasting experience by the Experiencer.
- **Stimulus** - The Stimulus is the person, event, or state of affairs that evokes the emotional response in the Experiencer.
- **Topic** - The Topic is the general area in which the emotion occurs. It indicates a range of possible Stimulus.

Non-Core:

- **Circumstances** - The Circumstances is the condition(s) under which the Stimulus evokes its response. In some cases it may appear without an explicit Stimulus. Quite often in such cases, the Stimulus can be inferred from the Circumstances

- **Degree** - The extent to which the Experiencer's emotion deviates from the norm for the emotion.
- **Empathy_target** - The Empathy_target is the individual or individuals with which the Experiencer identifies emotionally and thus shares their emotional response.
- **Explanation** - The Explanation is the explanation for why the Stimulus evokes a certain emotional response.
- **Manner** - Any description of the way in which the Experiencer experiences the Stimulus which is not covered by more specific FEs, including secondary effects (quietly, loudly), and general descriptions comparing events (the same way). Manner may also describe a state of the Experiencer that affects the details of the emotional experience.
- **Parameter** - The Parameter is a domain in which the Experiencer experiences the Stimulus.
- **Time** - The Time when the Experiencer, Expressor, or State can be described as having said emotion.

F.10 Food

Definition: This frame contains words referring to items of food.

Core FEs:

- **Food**

Non-Core:

- **Constituent_parts** - A part of the Food.
- **Descriptor** - This FE indicates a characteristic or description of the food.
- **Type** - This FE identifies a particular Type of the food item.

F.11 Entity

Definition: This frame is for words that denote highly schematic entities.

Core FEs:

- **Entity** - A thing (either abstract or physical) that exists with some degree of permanence.

Non-Core:

- **Constituent_parts** - Salient parts that make up the Entity.
- **Formational_cause** - Formational_cause indicates the action that brings about the features of an Entity.
- **Name** - This FE is used for the Name of an Entity.
- **Type** - This FE indicates what subtype of the general category the Entity belongs to.

F.12 Intoxicants

Definition: An Intoxicant is ingested in order to achieve an altered state of consciousness.

Core FEs:

- **Intoxicant** - The Intoxicant is ingested.

Non-Core:

- **Country_of_origin** - Country_of_origin identifies where the Intoxicant is grown or originates.
- **Descriptor** - Descriptor describes a quality of the Intoxicant.
- **Type** - Type identifies the kind or sort of Intoxicant.

F.13 Judgment of intensity

Definition: A speaker characterizes the intensity of an Attribute of an Entity by a Degree.

Core FEs:

- **Degree** - A modifier expressing the deviation of the intensity of the Attribute from the norm or from the value for another Entity. This FE is incorporated for each LU in this frame.
- **Entity** - The Entity for which a particular gradable Attribute is appropriate and under consideration.

Core Unexpressed:

- **Attribute** - The feature of an Entity which is under discussion.

Non-Core:

- **Circumstances** - Some specification of the circumstances under which the Entity has a particular Value for the Attribute.

- **Place** - The location where the Entity is has the Attribute to a given Degree.
- **Time** - The Time during which the Entity is in the state of having a particular Value for the Attribute.

F.14 Natural features

Definition: The Locale is a geographical location as defined by shape. This frame includes natural geographic features, including land/ice forms and bodies of water.

Core FEs:

- **Locale** - A stable bounded area. It is typically the designation of the nouns of Locale-derived frames.

Non-Core:

- **Constituent_parts** - Salient parts of the Locale.
- **Container_possessor** - The location that the Locale is a part of.
- **Descriptor** - Any description, characteristic, or property of the Locale which is not covered by more specific FEs.
- **Formational_cause** - Indicates the action (or causer) which brings the features of the Locale about.
- **Name** - This FE is used for the Names of Locales.
- **Relative_location** - A place that a Locale is located with respect to.
- **Type** - The Type is the subtype of natural feature.

F.15 Personal relationship

Definition: The words in this frame have to do with people and the personal Relationships they are or can be a part of.

Core FEs:

- **Partner_1** - That partner in the Relationship who is realized as the subject of verbs in active form sentences, or of adjectives in predicative uses.
- **Partner_2** - That partner in the Relationship who is not expressed as the external argument.
- **Partners** - The joint construal of both Partners in the Relationship.

Non-Core:

- **Degree** - Degree to which event occurs
- **Depictive** - Depictive phrase describing the Partners.
- **Duration** - The length of the relationship.
- **Manner** - Manner of performing an action
- **Means** - An act whereby a focal participant achieves an action indicated by the target.
- **Relationship** - The Relationship between Partners.
- **Source_of_relationship** - The source of the relationship.

F.16 Plants

Definition: The LUs in this frame refer to biological entities other than human beings and animals, which are labeled by the FE Plant.

Core FEs:

- **Plant** - The living organism belonging to the kingdom Plantae.

Non-Core:

- **Characteristic** - A quality of the Plant.
- **Descriptor** - A characterization of the Plant not covered by another frame element.
- **Origin** - The location from which the Plant originated.

F.17 Point of dispute

Definition: The answer to a Question is under discussion in a Group, which still has a difference of opinion among its members. The prominence of the Question relative to others can be indicated by a Status expression.

Core FEs:

- **Question** - The Question is an open proposition the answer to which is under dispute.

Non-Core:

- **Context** - A state or event within which the Question arises as a problem for the Group.

- **Descriptor** - A characterization of the Question, often with regard to its complexity or whether it has been resolved or not.
- **Domain** - The area of human experience which the Question concerns.
- **Group** - The Group are the persons or organizations who have different points of view on the Question.
- **Status** - The Frame Element Status is used for expressions that indicate the relative importance of a Question with respect to the other topics that are disputed.
- **Time** - The time interval during which the Question exists as a problem for the Group. The issue at the time was whether or not to relocate the Rennes-based part of the team to California.
- **Viewpoint** - An individual or individuals from whose point of view the Question is a point of dispute within the Group.

F.18 Precipitation

Definition: Water in some solid or liquid form (the Precipitation) falls from the sky at a particular Place and Time, lasting for a particular Duration. The Rate or Quantity of precipitation may also be indicated.

Core FEs:

- **Place** - The area which experiences precipitation.
- **Precipitation** - The water in liquid or solid form that falls from the sky is the Precipitation. This FE is incorporated in many LUs in this frame.
- **Time** - When the event occurs.

Non-Core:

- **Duration** - The amount of time for which the precipitation event lasts.
- **Explanation** - The Explanation is the reason for the precipitation event.
- **Frequency** - This frame element is defined as the number of times an event occurs per some unit of time.
- **Manner** - The way in which the precipitation falls.
- **Quantity** - The Quantity is the amount of Precipitation that has fallen in a precipitation event or events. It is usually measured as the height that the Precipitation attains if prevented from further movement.
- **Rate** - The Quantity of Precipitation that falls within a certain Duration.
- **Temperature** - This FE identifies the temperature of the precipitation.

F.19 Purpose

Definition: An Agent wants to achieve a Goal, or an object. A Means has been created or is used to allow the creator or a user to achieve a Goal. The Goal is a state of the world that does not currently hold but which the Agent wants to realize and is planning and/or working towards.
Core FEs:

- **Agent** - The volitional actor who intends to achieve a goal.
- **Attribute** - This FE marks expressions that indicate an Attribute for which it is the Agent's goal that it have or reach a particular Value .
- **Goal** - The action or state of affairs that the Agent wants to achieve.
- **Means** - A Means is an object or action that is used or designed to achieve a particular Goal.
- **Value** - This FE marks expressions that indicate which Value the Agent wants the Attribute to reach.

Non-Core:

- **Domain** - This frame element expresses a sphere of activity within which an Agent has a Goal.
- **Restrictor** - This FE marks expressions that indicate a spatial or temporal or individual framework relative to which the Goal is said to hold. More than one Restrictor can be mentioned.
- **Time** - This Frame Element expresses a time period during which an Agent had a particular Goal.

F.20 Quantity

Definition: This frame contains nouns denoting Quantities of a specified Entity and their numerical Values

Core FEs:

- **Entity** - Entity identifies the concept or substance being quantified by a Quantity.
- **Quantity** - Identifies the Quantity-denoting TARGET word.
- **Value** - Value identifies the numerical amount of the Entity.

Non-Core:

- **Descriptor** - A characterization of the Quantity.

F.21 Rewards and punishments

Definition: An Agent (the punisher or rewarder) performs a Response_action on an Evaluee for a Reason, the Evaluee's actions or beliefs. Means and Instrument may also be indicated. The goal of the punishment/reward is to discourage/encourage the actions or beliefs. Words in this frame presuppose that a judgment of the Evaluee has occurred and that the Evaluee is (or becomes) aware of the judgment. This judgment was performed by a cognizer which is either the same as the Agent, or, minimally, a representative of the same institution.

Core FEs:

- **Agent** - The Agent is the person doing the rewarding or punishing.
- **Evaluee** - Evaluee is the person or thing about whom/which a judgment has made and to whom reward/punishment is dealt. With verbs, the Evaluee is typically expressed as Object.
- **Reason** - Typically, there is a constituent expressing the Reason for the Agent's judgment.

Core Unexpressed:

- **Response_action** - The reward or punishment given to the Evaluee by the Agent.

Non-Core:

- **Degree** - Degree of reward or punishment.
- **Depictive** - Depictive phrase describing the Agent of the reward or punishment.
- **Instrument** -The Instrument with which the reward or punishment is carried out.
- **Manner** - Manner of performing the reward or punishment.
- **Means** -The action that is taken that results in punishment/reward.
- **Place** - Where the event takes place.
- **Purpose** -The Purpose of the reward or punishment.
- **Result** - Result of the reward or punishment.
- **Time** - When the event occurs.

F.22 Sensation

Definition: This frame contains nouns that refer to sensations in different modalities.

Core FEs:

- **Body_part** - This FE is assigned to phrases expressing the body part in which a sensation is located.
- **Ground** - This FE is the perceptual background against which the Phenomenon is experienced by the Perceiver.
- **Perceiver_passive** - This FE is only expressed with certain nouns in this frame, and then only with the help of a support verb, such as have: e.g. **I** have a tingling SENSATION in my hands.
- **Percept** - The FE Percept is used for phrases that express the characteristic property of a sensation. Typically this characteristic property is described with reference to the entity or phenomenon that produces the sensation.
- **Source** - This is the entity or phenomenon which gives rise to the sensation. In cases of veridical sensation, this FE is not typically distinguished from Percept.

Non-Core:

- **Degree** - This FE describes how strong the sensation is.
- **Descriptor** - A characterization of the sensation not covered by a more specific FE.
- **Time** - The Time when the sensation is perceived.

F.23 Shapes

Definition: This frame contains words which describe the dimensional extent and Shape of a Substance. These words do not make reference to a Part-whole relationship. There can also be a descriptor of the Shape, Shape_prop.

Core FEs:

- **Shape** - This is the Shape that the Substance is in.
- **Substance** - A Substance of which pieces can be extracted.

Non-Core:

- **Count** - The number of units of the measured Substance.
- **Shape_prop** - This describes a property of the Shape, e.g. it's size.

F.24 Similarity

Definition: Two or more distinct entities, which may be concrete or abstract objects or types, are characterized as being similar to each other. Depending on figure/ground relations, the entities may be expressed in two distinct frame elements and constituents , Entity_1 and Entity_2, or jointly as a single frame element and constituent, Entities. The similarity may be based on appearance, physical properties, or other characteristics of the two entities.

Core FEs:

- **Differentiating_fact** - A fact about Entity_1 or the Entities that reveals how Entity_1 is the same or different from other entities. (Note the contrast with 'as to'.)
- **Dimension** - This FE marks constituents which express a property in respect to which the similarity of the entities is assessed.
- **Entities**- This FE marks constituents that express the set of objects or types whose similarity is at issue.
- **Entity_1** - When there is an asymmetry, Entity_1 is the entity characterized by its similarity to Entity_2, whose characteristics are assumed to be known. Entity_1 is often an external argument.
- **Entity_2** - When the Entities are expressed separately, Entity_2 is the one whose characteristics are assumed to be known; it serves as a basis for establishing characteristics of Entity_1.

Non-Core:

- **Circumstances** - Circumstances describe the state of the world (at a particular time and place) which is specifically independent of the event itself and any of its participants.
- **Degree** - The extent to which entities are similar to each other, in general or with respect to some Dimension(s) The twins are very SIMILAR.
- **Depictive** - This FE is used for any Depictive phrase describing the state of the Entities
- **Explanation** - An inanimate entity or process that causes the similarity.
- **Manner** - This FE identifies the Manner in which Dimension is similar or different
- **Place** - The Place is the general area in which similarity occurs or exists
- **Time** -This FE identifies the Time when the similarity occurs or exists.

Appendix G

Word sketches from Wortprofil

ist Adjektivattribut von		logDice	Freq.
1. Lage	M W A	8.9	4357
2. Situation	M W A	6.9	1132
3. Finanzlage		6.7	57
4. Umstand		6.5	265
5. Wirtschaftslage		6.1	41
6. Wetterbedingung		5.7	18
7. Wetterverhältnis		5.4	12
8. Witterungsverhältnis		5.0	9
9. Tabellenlage		5.0	8
10. Witterung		4.7	13
11. Zustand		4.6	124
12. Witterungsbedingung		4.4	6
13. Finanzsituation		4.3	6
14. Bedingung		4.0	87
15. Verhältnis		3.7	106
16. Ausgangslage		3.6	6
17. Wetter		3.5	31
18. Tatsache		3.4	16
19. Sache		3.4	58
20. Lebensbedingung		3.4	5
21. Folge		3.0	58
22. Angelegenheit		3.0	25
23. Erfahrung		2.7	53
24. Konsequenz		2.5	18
25. Mission		1.8	5
26. Entwicklung		0.8	23
27. Gefühl		0.7	7
28. Position		0.5	9

FIGURE G.1: A screenshot from the Wortprofil: available nominal collocates for the adjective *misslich* 'difficult' sorted by the logDice values. The columns in the table are the following: the nominal head, logDice value, frequency in the corpus.

ist Adjektivattribut von		logDice	Freq.
1. Wohnraum	M W A	8.5	343
2. Preis	M W A	7.3	1713
3. Miete		6.9	142
4. Luxus		6.8	53
5. Mietwohnung		6.7	35
6. Wohnung	M W A	6.0	328
7. Eintrittspreis		5.9	20
8. Wohneigentum		5.9	17
9. Mietpreis		5.5	13
10. Unterkunft		5.4	32
11. Eigenheim		5.4	13
12. Bleibe		5.2	15
13. Bauland		5.2	11
14. Elektroauto		5.1	12
15. Gesundheitsversorgung		5.0	9
16. Synthesizer		5.0	8
17. Farbdrucker		4.9	7
18. Bodenpreis		4.6	6
19. Einfamilienhaus		4.5	9
20. Kleinwagen		4.5	8
21. Medikament		4.4	36
22. Sampler		4.4	5
23. Krankenversicherung		4.4	21
24. Digitalkamera		4.3	5
25. Alternative	M W A	4.1	42
26. Wohnen		4.1	15
27. Eigentumswohnung		4.1	6
28. Euro		3.3	21
29. Kondition		3.3	9
30. Vergnügen		3.2	7
31. Auto		2.8	39
32. Immobilie		2.8	6
33. Lebensmittel		2.4	6
34. Modell		2.2	27
35. Gebühr		2.2	5
36. Kredit		2.2	10
37. Summe		2.1	12
38. Kosten		2.1	22
39. Objekt		1.9	12
40. Ausgabe		1.6	18
41. Angebot		1.6	23
42. Energie		1.6	16
43. Kunst		1.4	17
44. Produkt		1.4	16
45. Rahmen		1.3	6
46. Bedingung		1.1	12
47. Gerät		1.0	7
48. Niveau		0.9	8
49. Haus		0.8	20
50. Bereich		0.4	12

FIGURE G.2: A screenshot from the Wortprofil: available nominal collocates for the adjective *erschwinglich* 'reasonable' sorted by the logDice values. The columns in the table are the following: the nominal head, logDice value, frequency in the corpus.

Appendix H

Classification of co-occurrences

The decision tree presented in Figure H.1 offers a simple procedure for classifying AN co-occurrences manually. An AN pair is considered a **co-occurrence** if the statistical association between its components is strong. There are two ways of defining what *strong* means in this case which are both rather arbitrary: either to determine a particular numeric cut-off threshold of an association score or to determine a cut-off on the number of word pairs in an n-best list, i.e. to take into consideration the top-n word pairs ranked based on the chosen association measure. In AN co-occurrences, the head (noun) is considered the *base* and its modifier (adjective), the *collocate*. The decision tree is to be applied to a pre-selected list of co-occurrences.

Apart from using a conventional association metric such as logDice, it is also recommended to compute a directional association metric such as ΔP or D_{KL}. Step 6 in the decision tree inquires whether a given phrase is asymmetric. In order to make this decision, a numeric cut-off threshold has to be set.

Definitions of each label for adjective-noun combinations are presented below. This classification scheme is not exhaustive and reflects only the findings from this thesis.

- **Proper name** is an adjective-noun combination used as a name or a part of the name of an organization, location, holiday or as a title of a song, a book, a movie, or similar. Examples: *(der) Gestiefelte Kater* 'Puss in Boots', *(das) Schwarze Meer* 'the Black Sea'

- **Idiom** is a recurrent binary word combination in which either only the nominal base has a non-literal meaning or the meaning of both constituents is non-literal. Examples: *dicke Luft* 'tense atmosphere', *schwarzes Schaf* 'black sheep'

- **Collocation type 1** is a recurrent binary word combination consisting of a base (noun) and a collocate (adjective) where the base is used in its literal sense and the collocate's meaning is non-literal. Examples: *helle Freude* 'pure joy', *schwarzer Kaffee* 'black coffee'

- **Technical term** is an adjective-noun combination the use of which is restricted to a particular domain. The meaning of the noun is literal. Examples: *vorläufige Vollstreckbarkeit* 'provisional enforceability', *spätes Geschäft* 'late trading'

- **Cliché** is an adjective-noun combination the use of which is restricted to a particular social situation. Examples: *gutes Gelingen* 'good luck', *herzlichen Glückwunsch* 'congratulations'
- **Free phrase** is a binary word combination consisting of a base (noun) and a modifier (adjective) where both the base and the modifier are used in unrestricted senses. Examples: *grünes Kleid* 'green dress', *herrliches Wetter* 'wonderful weather'
- **Collocation type 2** is a recurrent binary word combination consisting of a base (noun) and a collocate (adjective) where both elements retain their literal senses, but the use of the collocate is contingent on the base. Examples: *himmelweit Unterschied* 'stark difference', *geraume Zeit* 'considerable time'

Consider an example. Asymmetry is measured by ΔP. The cut-off is set to 0.7. The phrase under consideration is *geraume Zeit* 'considerable time':

1. Is it a proper name? – no
2. Is the meaning of the base *Zeit* literal? – yes
3. Is the meaning of the collocate *geraum* literal? – yes
4. Is the meaning of the phrase very specific and restricted to a particular domain? – no
5. Is the use of the phrase restricted to a particular social situation? – no
6. Is the phrase asymmetric? – yes, $\Delta P2 = 0.97$
7. Classify as **collocation type 2**

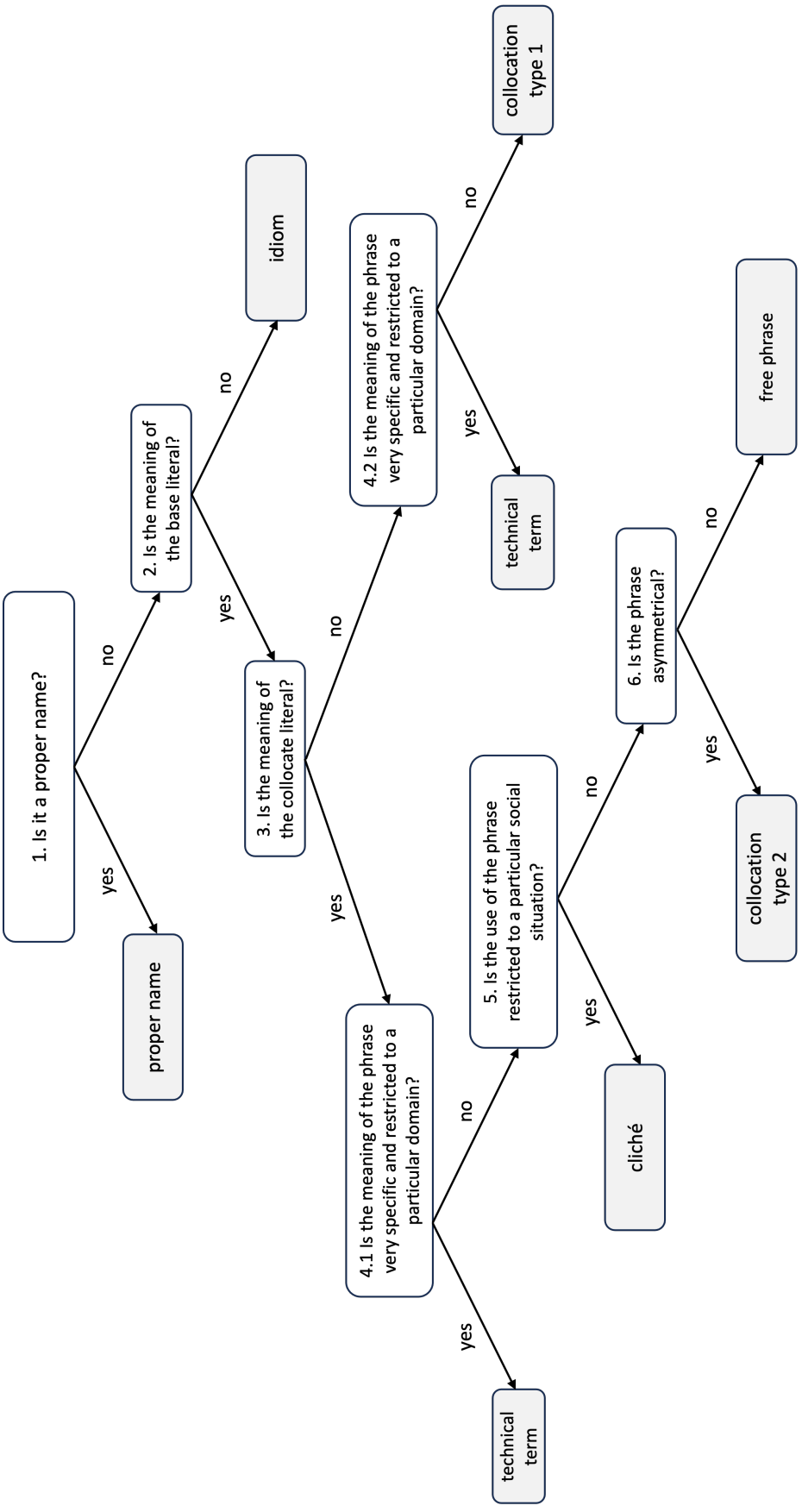


FIGURE H.1: Decision tree for classifying AN co-occurrences.

Consider a second example. The phrase under consideration is *vorläufige Vollstreckbarkeit* 'provisional enforceability':

1. Is it a proper name? – no
2. Is the meaning of the base *Vollstreckbarkeit* literal? – yes
3. Is the meaning of the collocate *vorläufig* literal? – yes
4. Is the meaning of the phrase very specific and restricted to a particular domain? – yes
5. Classify as **term**

Consider a third example. The phrase under consideration is *reife Frucht* 'ripe fruit':

1. Is it a proper name? – no
2. Is the meaning of the base *Frucht* literal? – yes
3. Is the meaning of the collocate *reif* literal? – yes
4. Is the meaning of the phrase very specific and restricted to a particular domain? – no
5. Is the use of the phrase restricted to a particular social situation? – no
6. Is the phrase asymmetric? – no, $\Delta P1 = 0.03$, $\Delta P2 = 0.12$
7. Classify as **free phrase**

Consider a fourth example. The phrase under consideration is *dicke Luft* 'tense atmosphere' (lit.'thick air'):

1. Is it a proper name? – no
2. Is the meaning of the base *Luft* literal? – no
3. Classify as **idiom**

Appendix I

Sense definitions of adjectives

Table I.1 presents sense definitions for each adjective in the GerCoAt dataset and the corresponding attributes. The sense definitions are based on the paraphrases in GermaNet and definitions in the DWDS and German Wiktionary. The notation in column "art. concept GN" is the following:

- artificial concept, e.g. steigerungsspezifisch: the concept comes directly from the adjectival modeling in GermaNet;
- artificial concept, artificial concept, e.g. bewertungsspezifisch, charakterspezifisch: both concepts are available in the modeling of the given adjective;
- artificial concept*, e.g. steigerungsspezifisch*: the concept comes from GermaNet, but it is not avaiable in the GN modeling of the given adjective AND the given sense is not reflected in the modeling of the adjective in GN;
- artificial concept (artificial concept), e.g. zustandsspezifisch (belebtheitsspezifisch): the artificial concept in brackets is the one from GN modeling of the given adjective sense, the first artificial concept NOT in brackets is the one selected from the full inventory in GN and judged more suitable by the annotators;
- —- (artificial concept): the attribute does not have an equivalent in GN, but the given adjective sense is modeled in GN under a different artificial concept (the one in brackets);
- —-: the attribute does not have an equivalent in GN and the given adjective sense is not modeled in GN;

adjective	definition	art. concept GN	example NN	attribute	status
alt	bejahrt, reich an Jahren 'aged, rich in years'	altersspezifisch	*Dame* 'lady'	age	free
alt	schon lange bestehend 'long existing'	dauerspezifisch	*Freund* 'friend'	duration	collocation
alt	vor langer Zeit gemacht oder geschehen, vorherig 'made or happened a long time ago, previously'	zeitspezifisch	*Regierung* 'government'	time	collocation
arm	ohne Geld, mittellos 'without money, destitute'	schichtenspezifisch	*Bevölkerung* 'population'	social	free
arm	bedauernswert, unglücklich 'unfortunate, unhappy'	Gefühl erweckend	*Kreatur* 'creature'	emotion	collocation
bitter	Geschmacksbezeichnung, gleich dem Geschmack der Galle 'taste description, like the taste of bile'	geschmacksspezifisch	*Mandel* 'almond'	taste	free
bitter	freudelos, traurig, übertragen: unangenehm, schmerzlich 'joyless, sad; fig.: unpleasant, painful'	Gefühl erweckend/ empfindend	*Moment* 'moment'	emotion	collocation
bitter	drückt große Verstärkung aus 'expresses great amplification'	steigerungsspezifisch	*Kälte* 'cold'	intensity	collocation
blank	glänzend, strahlend 'shiny, radiant'	oberflächenspezifisch	*Boden* 'floor'	surface	free
blank	unbedeckt, unbekleidet 'uncovered, unclothed'	aussehensspezifisch	*Oberkörper* 'torso'	appearance	collocation
blank	offensichtlich, rein, klar erkennbar 'obvious, pure, clearly recognizable'	gewissheitsspezifisch	*Lüge* 'lie'	certainty	collocation
blind	ohne Sehvermögen 'without sight'	krankheitsspezifisch	*Hund* 'dog'	health	free
blind	maßlos 'immoderate'	steigerungsspezifisch (privativ)	*Wut* 'rage'	intensity	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
blind	angelaufen, trübe, undurchsichtig'tarnished, cloudy, opaque'	oberflächenspezifisch	*Spiegel* 'mirror'	surface	collocation
blind	kritiklos 'without criticism'	—-	*Liebe* 'love'	description	collocation
blind	ohne Berechtigung 'without authorization'	geltungsspezifisch*	*Passagier* 'passanger'	validity	collocation
deftig	herzhaft 'savory, hearty'	geschmacksspezifisch	Wurst	taste	free
deftig	unhöflich, derb 'rude, rough'	—- (gesellschaftsspezifisch)	*Wortwahl* 'choice of words'	description	collocation
deftig	drückt große Verstärkung aus 'expresses great amplification'	steigerungsspezifisch*	*Niederlage* 'defeat'	intensity	collocation
dick	Übergewicht haben, zuviel wiegen 'Being overweight, weighing too much'	aussehensspezifisch	*Kind* 'child'	appearance	collocation
dick	gibt den Abstand zwischen zwei gegenüberliegenden Seiten eines Körpers an 'indicates the distance between two opposite sides of a body'	dimensionsspezifisch*	*Stapel* 'stack, pile'	dimension	free
dick	zähflüssig; so, dass es zwar nicht fest ist, aber zäh fließt 'viscous; such that it is not solid but flows viscously'	konsistenzspezifisch	*Nebel* 'fog'	consistency	collocation
dick	undurchdringlich 'impenetrable'	beschaffenheitsspezifisch	*Jacke* 'jacket'	texture	collocation
dick	eng, nahe 'narrow, close'	steigerungsspezifisch (beziehungsspezifisch)	*Freunde* 'friends'	intensity	collocation
dumm	ohne Verstand 'without understanding'	intelligenzspezifisch	*Hund* 'dog'	intelligence	free
dumm	mit nachteiligen Folgen, unangenehm, übel 'with adverse consequences, unpleasant, nasty'	bewertungsspezifisch (Gefühl erweckend)	*Fehler* 'mistake'	evaluation	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
dunkel	ohne oder wenig Licht, lichtlos; im Farbton mehr oder weniger kräftig 'without or little light, lightless; more or less strong in color tone'	helligkeitsspezifisch	*Zimmer* 'room'	brightness	free
dunkel	verdächtig, zwielichtig, ungewiss 'suspicious, dubious, uncertain'	gewissheitsspezifisch (bewertungsspezifisch)	*Geschäft* 'business'	certainty	collocation
dunkel	düster, unerfreulich, trostlos 'gloomy, unpleasant, desolate'	Gefühl erweckend	*Stunde* 'hour'	emotion	collocation
dunkel	negativ, schlecht 'negative, bad'	bewertungsspezifisch	*Seite* 'side'	evaluation	collocation
dunkel	unbestimmt, unklar 'indeterminate, unclear'	steigerungsspezifisch*	*Erinnerung* 'memory'	intensity	collocation
dunkel	tief tönend 'deep sounding'	geräuschspezifisch	*Stimme* 'voice'	sound	collocation
dunkel	eine bestimmte Art von .. 'a certain type of . . . '	—-	*Schokolade* 'chocolate'	type	collocation
frech	(ein wenig) respektlos '(a little) disrespectful'	charakterspezifisch	*Bursche* 'fellow'	character	free
frech	anmaßend oder unverschämt ist 'presumptuous or impertinent'	—- (charakterspezifisch)	*Lied* 'song'	description	collocation
frei	unbedeckt, unbekleidet 'uncovered, unclothed'	aussehensspezifisch*	*Oberkörper* 'torso'	appearance	collocation
frei	nicht mit Kosten verbunden; ohne Entgelt 'not associated with costs; free of charge'	kostenspezifisch	*Eintritt* 'entry'	cost	collocation
frei	ohne Hilfsmittel 'without aid'	—- (aufwandsspezifisch)	*Rede* 'speech'	description	collocation
frei	in der freien Natur 'in the great outdoors'	—-	*Natur* 'nature'	epithet	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
frei	unabhängig 'independent'	gesellschaftsspezifisch	*Journalistin* 'journalist'	social	free
frei	nicht besetzt 'not occupied'	raumverteilungsspezifisch, steigerungsspezifisch, quantitätsspezifisch	*Fläche* 'area'	space	collocation
frischgebacken	eben erst gebacken 'just baked'	zeitspezifisch	*Brot* 'bread'	time	free
frischgebacken	erst vor Kurzem (zu etw.) geworden, ernannt, berufen 'only recently become (sth.), appointed'	zeitspezifisch*	*Mutter* 'mother'	time	collocation
grob	bezogen auf Materialien: unfein, unbehauen, unbearbeitet 'related to materials: unrefined, unhewn, unprocessed'	beschaffenheitsspezifisch	*Leinen* 'linen'	texture	free
grob	grob, unhöflich, barsch 'rude, impolite, harsh'	description (charakterspezifisch)	*Scherz* 'joke'	description	collocation
grob	im hohen Maße 'to a high degree'	steigerungsspezifisch	*Unsinn* 'nonsense'	intensity	collocation
grob	ungenau 'imprecise'	genauigkeitsspezifisch	*Schätzung* 'estimation'	precision	collocation
hart	wie ein Festkörper; nicht gasförmig oder flüssig; 'like a solid body; not gaseous or liquid;'	konsistenzspezifisch	*Gegenstand* 'object'	consistency	free
hart	stabil, belastungsfähig, standhaft, (von Werten) ständig, gleichbleibend; 'stable, resilient, steadfast, (of values) constant, unchanging;'	sicherheitsspezifisch	*Währung* 'currency'	stability	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
hart	eine bestimmte Art von .. 'a certain type of . . . '	—-	*Droge* 'drug'	type	collocation
hart	von Fakten: lückenlos und über alle Zweifel erhaben bewiesen 'of facts: proven without gaps and beyond all doubt'	gewissheitsspezifisch (sicherheitsspezifisch)	*Fakt* 'fact'	certainty	collocation
hart	übertragen: keinem Druck nachgebend 'fig.: not yielding to any pressure'	—-	*Kurs* 'course'	rigidity	collocation
hart	stark, intensiv 'strong, intense'	steigerungsspezifisch	*Kontrast* 'contrast'	intensity	collocation
hart	Anstrengung erfordernd, mühsam 'requiring effort, laborious'	aufwandsspezifisch	*Konkurrenz* 'competition'	effort	collocation
hart	schwer erträglich, heftig, derb 'hard to bear, violent, coarse'	—-	*Realität* 'reality'	description	collocation
hart	abgehärtet, widerstandsfähig 'hardened, resistant'	konstitutionsspezifisch	*Bursche* 'fellow'	constitution	collocation
hart	fest,energisch, entschieden 'firm, energetic, decisive'	charakterspezifisch	*Kerl* 'chap'	character	collocation
hell	Eine Farbe mit hohem Weißanteil; leuchtend, nicht dunkel 'A color with a high white content; bright, not dark'	helligkeitsspezifisch	*Fenster* 'window'	brightness	free
hell	glücklich, positiv 'happy, positive'	bewertungsspezifisch	*Zukunft* 'future'	evaluation	collocation
hell	stark, intensiv 'strong, intense'	steigerungsspezifisch*	*Freude* 'joy'	intensity	collocation
hell	drückt ein gewaltiges Ausmaß aus 'expresses an enormous scale'	quantitätsspezifisch*	*Schar* 'flock, crowd'	quantity	collocation
hell	hoch tönend 'high pitched'	geräuschspezifisch	*Stimme* 'voice'	sound	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
hell	eine bestimmte Art von ... 'a certain type of ...'	—-	*Bier* 'beer'	type	collocation
herrlich	prachtvoll, hervorragend, ausgezeichnet 'splendid, outstanding, excellent'	bewertungsspezifisch	*Farbe* 'color'	evaluation	free
herrlich	gut gelungen 'well done'	bewertungsspezifisch	*Tor* 'goal (in sports)'	evaluation	collocation
historisch	einer anderen Geschichtsepoche angehörend 'belonging to a different historical epoch'	zeitspezifisch	*Gebäude* 'building'	time	collocation
historisch	Epoche machend, epochal, epochemachend, prägend 'epoch-making, epochal, epoch-making, formative'	bedarfsspezifisch	*Entscheidung* 'decision'	importance	collocation
hölzern	aus Holz 'made of wood'	beschaffenheitsspezifisch	*Boot* 'boat'	material	free
hölzern	übertragen: steif, wie aus Holz 'Fig.: stiff, as if made of wood'	— (bewegungsspezifisch)	*Dialog* 'dialogue'	description	collocation
karg	knapp, gering bemessen; 'scarce, small'	quantitätsspezifisch (aufwandsspezifisch)	*Gehalt* 'salary'	quantity	free
karg	auf das Notwendigste reduzierte, genügsame Lebensweise 'a frugal lifestyle reduced to the bare essentials'	—- (aufwandsspezifisch)	*Arbeitszimmer* 'study (room)'	description	collocation
karg	ohne (größere) Vegetation; Landwirtschaft: ertragsarm 'without (major) vegetation; agriculture: low yield'	klimatisch, landschaftsspezifisch	*Steppe* 'steppe'	climate	collocation
knapp	nicht reichlich, kaum ausreichend 'not plentiful, barely sufficient'	quantitätsspezifisch (steigerungsspezifisch)	*Zeit* 'time'	quantity	free

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
knapp	eng, kaum noch passend 'tight, barely fitting'	dimensionsspezifisch	*T-Shirt* 't-shirt'	dimension	collocation
knapp	gerade noch zureichend 'just enough'	gewissheitsspezifisch (steigerungsspezifisch)	*Resultat* 'result'	certainty	collocation
knapp	frei von Überflüssigem 'free from superfluous'	—-	*Geste* 'gesture'	description	collocation
knapp	etwas weniger als die genannte Zeitspanne 'slightly less than the specified period'	zeitspezifisch	*Frist* 'deadline'	time	collocation
leicht	ein geringes Gewicht habend 'having a low weight'	gewichtsspezifisch	*Gepäck* 'luggage'	weight	free
leicht	eine bestimmte Art von ... 'a certain type of ...'	—-	*Waffe* 'weapon'	type	collocation
leicht	unschädlich, verträglich, zuträglich 'harmless, compatible, beneficial'	zuträglichkeitsspezifisch	*Küche* 'cuisine'	tolerability	collocation
leicht	nicht sehr groß, gering 'not very large, low'	steigerungsspezifisch	*Veränderung* 'change'	intensity	collocation
leicht	ohne große Anstrengung, einfach, mühelos 'without great effort, simply, effortlessly'	aufwandsspezifisch	*Gegner* 'opponent'	effort	collocation
leicht	geistig anspruchslos, heiter, nur unterhaltend 'intellectually undemanding, cheerful, only entertaining'	—-	*Lektüre* 'reading'	description	collocation
mächtig	Macht besitzend, einflussreich 'Possessing power, influential'	—- (verhaltensspezifisch, schichtenspezifisch)	*Organisation* 'organisation'	power	free

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
mächtig	von beträchtlicher Größe, Ausdehnung, umfangreich 'of considerable size, extent, extensive'	dimensionsspezifisch (wirkungsspezifisch)	*Mauer* 'wall'	dimension	collocation
mächtig	von hohem Grad, sehr stark 'of high degree, very strong'	steigerungsspezifisch	*Auftrieb* 'boost'	intensity	collocation
mild	ohne Kraftaufwand; zart; 'without effort; delicate;'	steigerungsspezifisch (erziehungsspezifisch)	*Behandlung* 'treatment'	intensity	free
mild	nicht sehr stark 'not very strong'	steigerungsspezifisch (erziehungsspezifisch)	*Strafe* 'punishment'	intensity	collocation
mild	ohne Strenge 'without rigor'	—- (erziehungsspezifisch)	*Richter* 'judge'	description	collocation
mild	nicht stark im Geschmack 'not strong in taste'	geschmacksspezifisch	*Kaffee* 'coffee'	taste	collocation
mild	von einer mittleren Temperatur zwischen kühl und warm 'from a medium temperature between cool and warm'	temperaturspezifisch (klimatisch)	*Nacht* 'night'	temperature	collocation
mild	verträglich, zuträglich 'compatible, beneficial'	zuträglichkeitsspezifisch*	*Seife* 'soap'	tolerability	collocation
offen	nicht geschlossen, nicht zugemacht 'not closed'	zustandsspezifisch	*Fenster* 'window'	state	free
offen	unvollendet 'unfinished'	vollständigkeitsspezifisch	*Ende* 'end'	completeness	collocation
offen	ehrlich, aufrichtig 'honest, sincere'	—- (charakterspezifisch)	*Atmosphäre* 'atmosphere'	description	collocation
offen	vor den Augen aller; öffentlich 'before the eyes of all; publicly'	institutionsspezifisch	*Brief* 'letter'	social	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
offen	zugänglich; geöffnet; ungehindert; uneingeschränkt 'accessible; open; unhindered; unrestricted'	raumverteilungsspezifisch (aufwandspezifisch)	*Strecke* 'route'	space	collocation
offen	eine bestimmte Art von ... 'a certain type of ...'	—-	*Wunde* 'wound'	type	collocation
prall	ganz ausgefüllt, geweitet, gedehnt 'completely filled, widened, stretched'	dimensionsspezifisch (steigerungsspezifisch, quantitätsspezifisch)	*Traube* 'grape'	dimension	free
prall	ganz ausgefüllt 'completely filled'	quantitätsspezifisch	*Kasse* 'cash'	quantity	collocation
prall	mit voller Intensität 'with full intensity'	steigerungsspezifisch (bezugspezifisch)	*Hitze* heat'	intensity	collocation
rau	uneben, nicht glatt 'uneven, not smooth'	oberflächenspezifisch	*Fassade* 'facade'	surface	free
rau	unfreundlich, grob 'unfriendly, rude'	charakterspezifisch	*Bursche* 'fellow'	character	collocation
rau	unwirtlich, karg 'inhospitable, barren'	klimatisch, landschaftsspezifisch	*Norden* 'north'	climate	collocation
rau	grob, derb, unprätentiös 'rough, coarse, unpretentious'	—-	*Humor* 'humour'	description	collocation
rau	übertragen: nicht flüssig 'transfer: not liquid'	—-	*Lauf* 'run'	manner	collocation
rau	bewegt 'moved'	bewegungsspezifisch	*See* 'lake'	movement	collocation
rau	misstönend 'discordant'	geräuschspezifisch*	*Gesang* 'singing'	sound	collocation
reich	wohlhabend, vermögend 'wealthy'	schichtenspezifisch	*Eltern* 'parents'	social	free
reich	groß, umfangreich, eine Fülle von etwas aufweisend 'large, extensive, having an abundance of something'	quantitätsspezifisch	*Ölvorkommen* 'oil resources'	quantity	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
reich	hervorragend ausgestattet, prächtig, üppig 'excellently equipped, splendid, opulent'	—- (bewertungsspezifisch, aussehensspezifisch)	*Ausstattung* 'furniture'	description	
rund	kreisförmig, ringförmig 'circular, ring-shaped'	formspezifisch	*Holztisch* 'wooden table'	shape	free
rund	dicklich, rundlich 'chubby, roundish'	aussehensspezifisch	*Kinn* 'chin'	appearance	collocation
rund	gelungen, ohne Mängel, Störungen o.Ä.'successful, without defects, faults or similar'	bewertungsspezifisch	*Sache* 'matter'	evaluation	collocation
rund	voll, ganz 'full, whole'	vollständigkeitsspezifisch*	*Summe* 'sum'	completeness	collocation
sanft	ausgeglichen und friedfertig 'balanced and peaceful'	charakterspezifisch (erziehungsspezifisch)	*Seele* 'soul'	character	free
sanft	beruhigend und angenehm 'soothing and pleasant'	—-	*Melodie* 'melody'	description	collocation
sanft	ohne Spitze, nicht spitz endend 'without a sharp peak, not pointed'	dimensionsspezifisch*	*Hügel* 'hill'	dimension	collocation
sanft	nicht sehr stark 'not very strong'	steigerungsspezifisch (erziehungsspezifisch)	*Wind* 'wind'	intensity	collocation
sanft	ohne Kraftaufwand; zart 'without force; delicate'	steigerungsspezifisch*	*Massage* 'massage'	manner	collocation
scharf	gut schneidend 'that cuts well'	formspezifisch	*Messer* 'knife'	shape	free
scharf	bissig, aggressiv 'snappy, aggressive'	charakterspezifisch*	*Hund* 'dog'	character	collocation
scharf	im strengen, herrischen Tonfall 'in a stern, bossy tone'	—- (erziehungsspezifisch)	*Formulierung* 'wording'	description	collocation
scharf	anstrengend, alle Kräfte erfordernd	aufwandsspezifisch*	*Konkurrenz* 'competition'	effort	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
scharf	Munition, die besonders bei kriegerischen Auseinandersetzungen zur Bekämpfung des Gegners eingesetzt wird 'Ammunition used to combat the enemy, particularly in armed conflicts'	funktionalitätsspezifisch*	*Munition* 'munition'	function	collocation
scharf	mit voller Intensität 'with full intensity'	steigerungsspezifisch	*Kontrast* 'contrast'	intensity	collocation
scharf	genau 'precise'	genauigkeitsspezifisch	*Trennung* 'division'	precision	collocation
scharf	energisch durchgreifend, streng, hart 'vigorously assertive, strict, harsh, unyielding'	—- (erziehungsspezifisch)	*Regel* 'rule'	rigidity	collocation
scharf	den Geruchssinn betreffend: in der Nase beißend 'concerning the sense of smell: biting in the nose'	geruchsspezifisch	*Geruch* 'smell'	smell	collocation
scharf	plötzlich 'sudden'	zeitspezifisch	*Kurve* 'turn'	time	collocation
schlau	mit viel Verstand und List 'with a lot of sense and cunning'	intelligenzspezifisch	Kater	intelligence	collocation
schlau	ein Mobiltelefon mit umfangreichen Computer-Funktionen 'a cell phone with extensive computer functions'	funktionalitätsspezifisch*	*Telefon* 'phone'	function	collocation
schwarz	Farbe, die im RGB-Farbraum Werte von Rot, Blau und Grün nahe Null hat 'Color that has values of red, blue and green close to zero in the RGB color space'	farbspezifisch, helligkeitsspezifisch	*Hemd* 'shirt'	color	free

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
schwarz	unglücklich, unheilbringend 'unfortunate, calamitous'	Gefühl erweckend	*Tag* 'day'	emotion	collocation
schwarz	dunkle Hautfarbe habend 'having a dark skin color'	gesellschaftsspezifisch	*Mehrheit* 'majority'	social	collocation
schwarz	eine bestimmte Art von .. 'a certain type of'	—-	Tee	type	collocation
sicher	vor Gefahren geschützt 'Protected from dangers'	sicherheitsspezifisch	*Ort* 'place'	security	free
sicher	mit sehr hoher Wahrscheinlichkeit; ohne Zweifel, gewiss 'with very high probability; without doubt, certainly'	gewissheitsspezifisch	*Beweis* 'proof'	certainty	collocation
spät	in der Zeit schon weit fortgeschritten 'already far advanced in time'	zeitspezifisch	*Abend* 'evening'	time	free
spät	zukünftig, nachfolgend 'future, subsequent'	zeitspezifisch	*Präsident* 'president'	time	collocation
stark	körperlich kräftig 'physically strong'	konstitutionsspezifisch	*Mann* 'man'	constitution	free
stark	charakterfest, willensstark 'with strong character, strong-willed'	charakterspezifisch	*Frau* 'woman'	character	collocation
stark	große Anstrengung erfordernd 'requiring great effort'	aufwandsspezifisch*	*Konkurrent* 'competitor'	effort	collocation
stark	von großer Leistungsfähigkeit 'of great efficiency'	funktionalitätsspezifisch*	*Motor* 'engine'	function	collocation
stark	drückt qualitativ ein großes Maß, eine große Intensität aus 'expresses qualitatively a great measure, a great intensity'	steigerungsspezifisch	*Druck* 'pressure'	intensity	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
stark	Macht besitzend, einflussreich 'Possessing power, influential'	—-	*Partei* 'party'	power	collocation
stark	(von Werten) ständig, gleichbleibend '(of values) constant, unchanging'	stabilitätsspezifisch*	*Währung* 'currency'	stability	collocation
starr	mit voller Intensität 'with full intensity'	steigerungsspezifisch*	*Beharren* 'insistence'	intensity	collocation
starr	unbeweglich, steif 'immobile, stiff'	bewegungsspezifisch	*Mimik* 'facial expression'	movement	free
starr	unnachgiebig, keiner Wandlung fähig 'unyielding, incapable of change'	—- (erziehungsspezifisch)	*Bürokratie* 'bureaucracy'	rigidity	collocation
steil	gewagt 'bold'	—-	*These* 'claim, proposition'	description	collocation
steil	von der Waagerechten stark abweichend, entweder ansteigend oder abfallend 'strongly deviating from the horizontal, either rising or falling'	dimensionsspezifisch	*Straße* 'street'	dimension	free
steil	beachtlich, beträchtlich 'considerable, substantial'	steigerungsspezifisch	*Hierarchie* 'hierarchy'	intensity	collocation
steil	sehr schnell 'very fast'	geschwindigkeitsspezifisch*	*Karriere* 'carreer'	speed	collocation
stürmisch	stark windig, mit Sturm 'very windy, with storms'	klimatisch	Nacht	climate	free
stürmisch	wild, heftig 'wild, violent'	—-	*Protest* 'protest'	description	collocation
stürmisch	leidenschaftlich, ungestüm, hitzig 'passionate, impetuous, hot-tempered'	Gefühl empfindend	*Affäre* 'affair'	emotion	collocation
stürmisch	mit voller Intensität 'with full intensity'	steigerungsspezifisch*	*Jubel* 'cheering'	intensity	collocation
stürmisch	wild, leidenschaftlich 'wild, passionate'	—-	*Kuss* 'kiss'	manner	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
stürmisch	sehr schnell vor sich gehend, sich sehr schnell vollziehend 'proceeding very quickly, taking place very quickly'	geschwindigkeitsspezifisch	*Expansion* 'expansion'	speed	collocation
süß	eine angenehme Empfindung hervorrufend 'evoking a pleasant feeling'	Gefühl erweckend	*Rache* 'revenge'	emotion	collocation
süß		—-	*Versuchung* 'temptation'	epithet	collocation
süß	niedlich, knuddelig, goldig 'cute, cuddly, sweet'	bewertungsspezifisch, charakterspezifisch	*Baby* 'baby'	evaluation	collocation
süß	im Geruch dem süßen Geschmack ähnlich, vergleichbar 'similar in smell to the sweet taste'	geruchsspezifisch	*Geruch* 'smell'	smell	collocation
süß	wie Zucker, Honig schmeckend 'tasting like sugar, honey'	geschmacksspezifisch	*Kuchen* 'cake'	taste	free
teuer	einen hohen Preis oder hohe Kosten aufweisend oder verursachend 'having or causing a high price or high costs'	kostenspezifisch	*Auto* 'car'	cost	free
tief	von oben (weit) nach unten reichend, sich weit nach unten erstreckend 'reaching from above (far) downwards, extending far downwards'	dimensionsspezifisch	*Brunnen* 'well'	dimension	free
tief		—	*Seele* 'soul'	epithet	collocation
tief	sehr stark, sehr groß, intensiv 'very strong, great, intense'	steigerungsspezifisch	*Trauer* 'grief'	intensity	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
tief	infolge niedriger Schwingungszahl dunkel tönend 'dark sounding due to low vibration frequency'	geräuschspezifisch	*Stimme* 'voice'	sound	collocation
tot	gestorben, nicht mehr am Leben 'deceased, no longer alive'	zustandsspezifisch (belebtheitsspezifisch)	*Fliege* 'fly'	state	collocation
tot	nicht weiterfunktionierend, nicht mehr in Betrieb 'no longer functioning, no longer in operation'	funktionalitätsspezifisch	*Gleis* 'track'	function	free
wild	mit wenig Selbstbeherrschung/ Selbstkontrolle 'with little self-control/ self-restraint'	charakterspezifisch (Gefühl empfindend)	*Geselle* 'fellow'	character	collocation
wild	heftig, ungestüm 'fierce, impetuous'	—-	*Party* 'party'	description	collocation
wild	zornig 'angry'	Gefühl empfindend	*Geschrei* 'screaming'	emotion	collocation
wild	sehr stark, sehr groß, intensiv 'very strong, great, intense'	steigerungsspezifisch*	*Durcheinander* 'mess'	intensity	collocation
wild	heftig, ungestüm 'fierce, impetuous'	—-	*Prügelei* 'fighting'	manner	collocation
wild	hektisch, zappelig, unbändig, ruhelos 'hectic, fidgety, unruly, restless'	bewegungsspezifisch	*Wasser* 'water'	movement	collocation
wild	nicht gezähmt, freilebend, frei in der Natur vorkommend, naturbelassen 'not tamed, free-living, freely occurring in nature, natural'	—- (tierspezifisch)	*Fluss* 'river'	naturalness	free
wild	nicht genehmigt, vom Gesetz abweichend 'not authorized, deviating from the law'	geltungsspezifisch	*Parken* 'parking'	validity	collocation

Continued on the next page

Table I.1: Continued from previous page

adjective	definition	art. concept GN	example NN	attribute	status
windig	unzuverlässig, zweifelhaft, nicht glaubhaft 'unreliable, doubtful, not credible'	gewissheitsspezifisch (privativ)	*Händler* 'trader'	certainty	collocation
windig	mit Wind, voller Luftbewegung 'with wind, full of air movement'	klimatisch	*Nachmittag* 'afternoon'	climate	free
zäh	schwerflüssig, zähflüssig 'heavy-bodied, viscous'	konsistenzspezifisch	*Brei* 'porridge'	consistency	free
zäh	ausdauernd 'enduring'	konstitutionsspezifisch	Bursche	constitution	collocation
zäh	langwierig 'tedious'	—-	*Diskussion* 'diskussion'	description	collocation
zäh	sehr stark, sehr groß, intensiv 'very strong, great, intense'	steigerungsspezifisch*	*Bemühung* 'effort'	intensity	collocation
zart	leicht reagierend, empfindsam 'easily reacting, sensitive'	charakterspezifisch (Gefühl empfindend)	*Seele* 'soul'	character	collocation
zart	schwächlich, anfällig 'weak, fragile'	konstitutionsspezifisch	*Jüngling* 'youth'	constitution	collocation
zart	bildlich: sanft auf die Sinne wirkend 'fig.: gently affecting the senses'	—-	*Lyrik* 'poetry'	description	collocation
zart	nicht sehr groß, gering 'not very large, low'	steigerungsspezifisch*	*Anzeichen* 'indication'	intensity	collocation
zart	ohne Kraftaufwand, sanft 'without force, gentle'	—-	*Berührung* 'touch'	manner	collocation
zart	fein, weich, dünn 'fine, soft, thin'	beschaffenheitsspezifisch	*Spitze* 'lace'	texture	free

TABLE I.1: Sense definitions of adjectives in the GerCoAt dataset, artifitial concept in GermaNet the sense belongs to, example noun, final attribute after annotation and adjudication, collocational status.