

# Annotating Quantified Phenomena in Complex Sentence Structures Using the Example of Generalising Statements in Literary Texts

**Tillmann Dönicke**  
Göttingen Centre for  
Digital Humanities  
University of Göttingen  
tillmann.doenicke@  
uni-goettingen.de

**Luisa Gödeke**  
Department of  
German Philology  
University of Göttingen  
luisa.goedeke@  
uni-goettingen.de

**Hanna Varachkina**  
Department of  
German Philology  
University of Göttingen  
hanna.varachkina@stud.  
uni-goettingen.de

## Abstract

We present a tagset for the annotation of quantification which we currently use to annotate certain quantified statements in fictional works of literature. Literary texts feature a rich variety in expressing quantification, including a broad range of lexemes to express quantifiers and complex sentence structures to express the restrictor and the nuclear scope of a quantification. Our tagset consists of seven tags and covers all types of quantification that occur in natural language, including vague quantification and generic quantification. In the second part of the paper, we introduce our German corpus with annotations of generalising statements, which form a proper subset of quantified statements.

## 1 Introduction

Quantification is a core element of human language because it allows us to make statements about groups or classes of entities, in contrast to statements about individually referenced entities.

One subtype of quantified statements are *generalising* or *generic* (for now summarised as *generalising*) statements that involve quantification over assumed members of a class rather than contextually given entities. These generalising statements are particularly interesting for NLP applications that operate on discourse-level, e.g. in knowledge extraction (e.g. [Bhaktavatsalam et al., 2020](#)) and argumentation mining (e.g. [Becker et al., 2016](#)). But also in computational literary studies, generalising statements can be viewed as indicators (i.e. features) for meta-level phenomena such as passages of (self-)reflection (cf. [Lahn and Meister, 2016](#), p. 184) or passages addressing real-world issues within a fictional text (e.g. [Vesper, 2014](#)).

Put very concisely, one traditionally differentiates between (determiner or generic) quantification

on NP-level and (adverbial or generic) quantification on clause-level ([Krifka and Gerstner, 1987](#); [Partee, 1990](#); [Krifka, 2016](#)), which is also reflected in certain annotation schemes (cf. [Friedrich et al., 2015](#)). However, for such higher-level applications, where the presence of quantification or, more specifically, generalisation serves as a feature, the syntactic or semantic structure of quantified statements plays only a subordinate role. Therefore, performing a syntactic and/or semantic analysis during the annotation would be laborious but not expedient—especially in domains where sentences tend to employ a complex syntactic structure, such as literary texts. Moreover, quantified statements are not always marked by an overt linguistic marker but can also be covertly quantified in the case of generic statements.

For generalisation specifically, previous work commonly differentiates between generalisation over (kinds of) individuals and generalisation over recurring events/situations ([Krifka et al., 1995](#); [Carlson, 2011](#); [Friedrich and Palmer, 2014](#)). Similar to the syntactic categorisation, this is not only an insufficient differentiation if one is interested in generalising statements as a whole; it further constitutes a limitation, since it is possible to quantify (and thus generalise) over other types of entities than the two just mentioned. We shall expand on the theoretical syntactic and semantic considerations in Section 2 and give an overview of related practical challenges in literary texts in Section 3.

Considering both the theoretical and practical aspects, we developed a tagset and annotation guidelines for generalising statements that are neither bound to syntactic nor to semantic properties and preserve only the information which is most important in our view: the type of quantification (universal/existential/vague etc.). This shallow annotation scheme allows a comparatively fast annotation of generalising statements, which is especially valu-

Quantifier	Restrictor	Nuclear scope
determiner,	subordinate clause,	main clause,
adverb,	if-clause,	assertion,
negation,	common NP,	predication,
generic	topic	focus

Table 1: Examples for tripartite-structure components (Partee, 1990, p. 10)

able in the literary domain. Since our annotation scheme can be used for quantified statements in general, we will present it as such in Section 4 and turn to generalising statements in Section 5. Afterwards, Section 6 presents details about our corpus and annotation process. Sections 7 and 8 discuss related and future work, respectively.

## 2 Quantification

In English (and in German), quantificational notions are typically triggered by determiners, e.g. *all*, *most* etc., or adverbs, e.g. *always*, *usually* etc. Following Lewis (1975), Kamp (1981) and Heim (1982), we assume quantified statements to consist of a tripartite logical form consisting of the quantifier  $Q$ , the restrictor and the nuclear scope:

$$(1) \quad Q[x : restr(x)][scope(x)]$$

We subsume determiners and adverbs to both function as quantifiers in this model:

- (2) a. Most horses have four legs.  
b.  $MOST[x : horse(x)][four\text{-legged}(x)]$
- (3) a. Usually, a horse has four legs.  
b.  $USUALLY[x : horse(x)][four\text{-legged}(x)]$

This approach enables us to include various clause-level forms of quantification into a unifying analysis. The forms can differ in syntactic realisation; Table 1 shows some examples for how quantifier, restrictor and nuclear scope can be realised in natural language, beyond determiners/adverbs and common noun phrases.

In addition to the syntactic diversity of quantification, it can range over all types of semantic entities. While quantification over individuals, as in (2) and (3), and events/situations, so-called *habituals* (e.g. Rimell, 2004) as in (4), may be most notable, it is also possible to quantify over e.g. times, as in (5), and locations, as in (6).

- (4) a. John usually drives to work

b.  $USUALLY[e : agent(e, j) \wedge to.work(e)][drive(e)]$

- (5) a. It snows every winter

b.  $\forall[t : winter(t)][GEN[e : e \subseteq t][snow(e)]]$

- (6) a. It snows in Austria

b.  $GEN[l : in.Austria(l)][GEN[e : location(e, l)][snow(e)]]$

## 3 Challenges in Literary Texts

We are interested in the distribution of such structures as “general statements” or “statements of universal validity” in German fictional texts. Therefore, we are aiming at annotating quantified statements in any form they may occur in. Our corpus consists of fictional texts written or published between 1650 and 1950. Hence, we are not only confronted with complex sentence structures, which are typical for literary texts, but also with older versions of German. We need an annotation concept that lets us capture quantified expressions in all their variety. Although our research does not primarily focus on the surface quantification, but on the generalising function that these structures fulfil, our work has a solid foundation in (formal) linguistics, as we will show in Section 4. Therefore, the transfer of theoretical knowledge about quantification into computational linguistics turns out to be a challenge—especially for analysing the literary domain. Particularly, we are facing three main challenges: First, the default formal analysis of quantification by defining quantifier, restrictor, and scope (as established in the previous section) can be highly complex in sentences of fictional writing. On top of that, German allows comparatively long and complex multi-clause sentences, especially in older language variants. This issue is illustrated in (7), where we already cut out several embedded if-clauses (German: *wenn* ‘if’) out of this one sentence. The English translation in (7’) does not take the various if-clauses on and splits them up into separate sentences:

- (7) Wenn Luciane, meine Tochter, die für die Welt geboren ist, sich dort für die Welt bildet, [...]; wenn sie durch Freiheit des Betragens, Anmut im Tanze, schickliche Bequemlichkeit des Gesprächs sich vor allen auszeichnet und durch ein angeborenes herrschendes Wesen sich zur Königin des kleinen Kreises macht, wenn

die Vorsteherin dieser Anstalt sie als kleine Gottheit ansieht, die nun erst unter ihren Händen recht gedeiht, die ihr Ehre machen, Zutrauen erwerben und einen Zufluß von andern jungen Personen verschaffen wird, wenn [...]: so ist dagegen, was sie schließlich von Ottilien erwähnt, nur immer Entschuldigung auf Entschuldigung [...]. (Goethe, WV)

- (7') Luciana, my daughter, born as she is for the world, is there training hourly for the world [...] She distinguishes herself above every one at the school with the freedom of her carriage, the grace of her movement, and the elegance of her address, and with the inborn royalty of nature makes herself the queen of the little circle there. The superior of the establishment regards her as a little divinity, who, under her hands, is shaping into excellence, and who will do her honor, gain her reputation, and bring her a large increase of pupils; [...] while her concluding sentences about Ottilie are nothing but excuse after excuse. (Goethe, EA, p. 23 f.)

If-clauses, as in (7) can be considered as restrictors (compare Table 1); and the then-clause (German: *so* ‘then’) can be considered as nuclear scope. Our first problem manifests itself here: The if-clauses form a list of coordinated restrictors for only one scope, and it remains unclear how many individual quantified statements there are or whether the individual restrictors are meant to be joined by logical conjunction or disjunction. Resolving such issues would require a laborious and—in our case—redundant analysis.

For ease of presentation, we shall only use English examples in the following, taken from official translations. The original examples are provided in the appendix (B).

Second, we have to deal with ambivalent syntactic structures, leading to scope ambiguity. If a sentence carries more than one quantifier, different readings arise due to the dominant quantification.

- (8) Help upon the spot is the thing you often most want in the country. (Goethe, EA, p. 49)

In (8) we find two generic expressions (*help upon the spot* and *the country*) combined with the adverbial *often*. Third, the absence of overt markers

Tag	Description
ALL	overt universal quantification
MEIST	overt majority quantification
EXIST	overt existential quantification
ZAHL	overt numerical quantification
DIV	overt vague quantification
BARE	none of the above + covert quantification
NEG	any of the above + negation

Table 2: Tagset

for quantification is a greater problem than an overpresence. The generic NPs (cf. Leslie and Lerner, 2016), e.g. *business* and *life* in (9), certainly have a generalising function in this context, but are not overtly quantified.

- (9) Business requires earnestness and method; life must have a freer handling. (Goethe, EA, p. 46)

In the following section, we will present our annotation tagset, which allows us to tackle these issues.

## 4 A Tagset for Quantified Statements

The complete tagset is summarised in Table 2. Because of the challenges associated with identifying restrictor and scope of a quantified statement, we do not annotate them separately. Instead, we label the whole span which contains quantifier, restrictor and scope. The tags in our tagset represent the (semantic) type of quantification. We take clauses as the smallest unit of annotation, meaning that one quantified statement may comprise one clause, as in (10), or several clauses, as in (11–12). Punctuation at annotation boundaries is omitted.

- (10) [Most horses have four legs]<sub>MOST</sub>.  
 (11) [A whale which is ill yields no blubber]<sub>BARE</sub>. (cf. Burton-Roberts, 1976)  
 (12) [He who gets up early gets tired quicker]<sub>BARE</sub>.

We use brackets to indicate annotation spans and subscripts to denote tags. The following subsections motivate the individual tags.

### 4.1 Precise Quantification

Natural language employs a clear-cut set of mathematically precise quantifiers, whose meanings can be defined using set relations (see Table 3). All

Name	$Q$	$Q[x : restr(x)][scope(x)]$ iff
universal	$\forall$	$ S_{restr} \cap S_{scope}  =  S_{restr} $
majority	MOST	$ S_{restr} \cap S_{scope}  >  S_{restr} \setminus S_{scope} $
existential	$\exists$	$ S_{restr} \cap S_{scope}  \neq 0$
counting	$\exists^{Rn}$	$ S_{restr} \cap S_{scope}  \geq Rn$
proportional	$Q_{prop}^{Rn/m}$	$ S_{restr} \cap S_{scope}  \geq Rn/m \cdot  S_{restr} $

Table 3: Truth conditions for precise quantifiers;  $S_P := \{x : P(x)\}$  is the extension of  $P$

of these quantifiers are expressed by a number of lexemes at the surface of a sentence.  $\forall$  is expressed by *all, every, always, everywhere* etc., and MOST usually appears as *most(ly)* or *main(ly)*. Statements with these quantifiers should be labelled with the tags ALL and MEIST (German for “MOST”), respectively:

- (13) There is lime, you remember, [which shows the strongest inclination for all sorts of acids—a distinct desire of combining with them]<sub>ALL</sub>. (Goethe, EA, p. 55 f.)
- (14) [Men think most of the immediate—the present]<sub>MEIST</sub>; (Goethe, EA, p. 16)

$\exists$  is associated with the indefinite article *a/an* in classical Fregean semantics (Zalta, 2020), as (15a) and (15b) exemplify. Fodor and Sag (1982) note, however, that a statement as in (15a) rather is ambiguous between the quantified interpretation in (15c) and the referential interpretation in (15d) (cf. von Heusinger, 2000)<sup>1</sup>. We follow this analysis and do not consider indefinite NPs to be markers for existential quantification. Instead, statements with a meaning as in (15c) are treated as genuine generic quantification (see Section 4.4): and statements with a meaning as in (15d) must not be labelled since they do not contain quantification.

- (15) a. A man walks  
b.  $\exists[x : \text{man}(x)][\text{walk}(x)]$   
c.  $\text{GEN}[x : \text{man}(x)][\text{walk}(x)]$   
d.  $\text{walk}(\varepsilon_i x \text{man}(x))$

We use the tag EXIST for explicit existential statements instead. In English, such statements can be formulated with the expression *there is/are* or the verb *exist*:

<sup>1</sup>The expression  $\varepsilon_i x P(x)$  returns an entity which satisfies the predicate  $P$ , based on the choice function  $i$  (Avigad and Zach, 2020).

(16) [Thirdly, there are those people who investigate the sea bed as if it were a meadow]<sub>EXIST</sub>. (Fontane, Stechlin, p. 288)

(17) [But they still do exist, they’ve got to exist or else they’ve got to exist *again*]<sub>EXIST</sub>. (Fontane, Stechlin, p. 130)

There are different theories on how to analyse such existential statements, differing in the question whether existence is a quantifier or a predicate and, in case of the latter, what kind of predicate it is (McNally, 1998; Moltmann, 2013). Although we do not have a preference for either analysis, we can observe that the verb *exist* must sometimes be analysed as a scope predicate rather than a quantifier. For example, it could be analysed either as existential quantifier or predicate of a covert quantifier in (18), whereas the quantifier analysis is not possible in (19) because of the overt quantifier MOST. The difference between (18b) and (18c) is very subtle and it is not always easy or even possible to identify the correct analysis—especially because a generic quantifier can also have an existential interpretation (Cohen, 2004). Therefore, and because we prefer to keep all occurrences of *exist* in one class, we label all occurrences with EXIST and treat cases like (19) as double quantification (see Section 4.3).

- (18) a. [Fairy-tale creatures exist]<sub>EXIST</sub>  
b.  $\exists[x][\text{fairy-tale.creature}(x)]$   
c.  $\text{GEN}[x : \text{fairy-tale.creature}(x)][\text{exist}(x)]$
- (19) a. [Most fairy-tale creatures exist]<sub>EXIST+MEIST</sub>  
b.  $\text{MOST}[x : \text{fairy-tale.creature}(x)][\text{exist}(x)]$

The last type of precise quantifiers are numerical quantifiers, which either express absolute counts ( $\exists^{Rn}$ ) or proportions ( $Q_{prop}^{Rn/m}$ ). Numerical quantifiers are composed of numerals, such as *one, two, half, third, dozen, hundred, percent, million*, corresponding to  $n$  (and  $m$ ). Numerals are optionally combined with an expression like *at least, exactly, up to* etc., corresponding to a mathematical relation  $R \in \{=, <, >, \leq, \geq, \dots\}$ :

- (20) a. Five men walk  
b.  $\exists^{=5}[x : \text{man}(x)][\text{walk}(x)]$
- (21) a. At least five men walk  
b.  $\exists^{\geq 5}[x : \text{man}(x)][\text{walk}(x)]$

- (22) a. Up to two thirds of men walk  
 b.  $Q_{prop}^{\leq 2/3}[x : \text{man}(x)][\text{walk}(x)]$

Numerical quantification should be labelled with the tag ZAHL (German for “NUMBER”):<sup>2</sup>

- (25) The county had always gone Conservative and it was a matter of honor to go Conservative again, as Luther had said, “[Even if the world were full of a thousand devils]<sub>ZAHL</sub>.” (Fontane, Stechlin, p. 140)

The reader might argue that “almost all” as in (26) also has a mathematical definition (that is “all but finitely many”) and should thus receive a separate tag. In natural language, however, *almost* modifies the truth value of a statement rather than its quantification (Kilbourn-Ceron, 2014). In fact, *almost* can appear in combinations with other quantifiers (see (27)) and without any quantifier (see (28)) as well, hence we do not include a separate tag for “almost all” in our tagset. We also treat similar modifiers like *hardly*, *nearly*, *more or less* etc. as not affecting the type of quantification.

- (26) [And so for starters he’s got to conquer everything, almost all the towns roundabout and all the castles for sure]<sub>ALL</sub>. (Fontane, Stechlin, p. 82)  
 (27) [Almost five men walk]<sub>ZAHL</sub>  
 (28) “You must let me make what will seem a wide sweep; we shall be on our subject almost immediately.” (Goethe, EA, p. 53)

## 4.2 Vague Quantification

In addition to the precise quantifiers, one can find a broad range of vague quantifiers in natural language, whose truth conditions cannot be defined precisely. Some lexemes are *few*, *some*<sup>3</sup>, *many*,

<sup>2</sup>A potentially conflicting case is  $Q_{prop}^{>1/2}$ , which is mathematically equivalent to MOST. Following Hackl (2009), who provides evidence for a cognitive difference between *more than half* and *most*, we label these expressions as follows:

- (23) a. [Most of the men walk]<sub>MEIST</sub>  
 b. MOST $[x : \text{man}(x)][\text{walk}(x)]$   
 (24) a. [More than half of the men walk]<sub>ZAHL</sub>  
 b.  $Q_{prop}^{>1/2}[x : \text{man}(x)][\text{walk}(x)]$

<sup>3</sup>The authors of this paper intensely discussed whether the German *manch*, which has a similar meaning to that of *some*, should be DIV or EXIST, because some scholars analyse *manch/some* as existential quantifier (e.g. Löbner, 2005; Chierchia and McConnell-Ginet, 2000, p. 310). We decided to label *manch* with DIV since it usually implies an indefinite but

*rarely*, *occasionally*, *commonly*, *often*; and multi-word expressions like *as a rule* or *in general* can also express vague quantification. The vagueness makes it difficult to determine how many semantically different quantifiers there are—if not every lexeme represents its own quantifier. For example, is *often* the same as *frequently*? We therefore group all vaguely quantified statements under the tag DIV (for “diverse”):

- (29) “[Our excellent superior commonly permits me to read the letters in which she communicates her observations upon her pupils to their parents and friends]<sub>DIV</sub>. (Goethe, EA, p. 43)  
 (30) “It concerns our friend the Captain,” answered Edward; “you know the unfortunate position [in which he, like many others, is placed]<sub>DIV</sub>. (Goethe, EA, p. 13)

## 4.3 Multiple Quantification

As mentioned in Section 3, several quantifiers can occur within a statement; or several quantified statements can be nested as in (31). Annotations from overlapping statements do not affect each other, hence this is no multi-label case in the sense of having multiple tags for one statement. It can become a multi-label case if one merges tags on token or clause level, though, e.g. for measuring annotator agreement or evaluating a quantification tagger.

Statements that contain more than one overt quantifier, on the other hand, should receive all corresponding tags, as in (32). We treat the assigned tags as a set, meaning that every of the five tags for overt quantification can be assigned only once to a statement, even if there are e.g. several quantifiers qualifying for the ALL tag, as in (33).

- (31) His entrance into the regiment more or less coincided with the beginning of the reign of Friedrich Wilhelm IV, [and whenever he mentioned that fact, he took pleasure in poking a bit of fun at himself by stressing that “[all great events have their accompanying secondary phenomena]<sub>ALL</sub>.”]<sub>ALL</sub>

substantial (vague) number of entities (Dudenredaktion, n.d.). (Furthermore, the determiner does not quite fit in the EXIST class from a morphological perspective.) We suggest that one should classify statements with *some* as DIV by the same argument. Note that *manch* always causes a quantificational interpretation (with singular and with plural NPs), whereas *some* can also cause a referential interpretation (like the indefinite article in (15d)) when combined with a singular NP (Winter, 1997). In the latter case, no tag would be assigned.

(Fontane, Stechlin, p. 3)

- (32) [Whoever eats meat sometimes is a murderer forever]<sub>ALL+DIV</sub>  
 (33) [Every Pope loves all his subjects equally]<sub>ALL</sub>

The individual quantifiers within a statement do not always employ an unambiguous hierarchy, or the hierarchy becomes apparent after labourious semantic analysis only. Therefore, we do not opt for an ordering of the tags in the case of multiple quantification.

There are more morphosyntactically complex quantifiers in natural language than we could discuss in the previous subsections (see Keenan and Paperno (2012) for an extensive overview). Complex quantifiers should be decomposed whenever no single tag is applicable, which can also result in a multi-label annotation:

- (34) [All but two men walk]<sub>ALL+ZAHL</sub>

#### 4.4 Generic Quantification

In opposition to the other quantifiers discussed so far, the generic quantifier GEN is covert, i.e. it is not marked by a specific lexical item. Instead, there is a broad range of surface forms that can mark genericity. The statements in (35), for example, (cf. Carlson, 2011) all make a general claim about lions.

- (35) a. The lion is ferocious  
 b. Lions are ferocious  
 c. A lion is ferocious  
 d. GEN[x : lion(x)][ferocious(x)]

While (35) shows generic statements about entities, (36) shows generic statements about events. Note that we only show one possible analysis in (36c), although several interpretations are possible due to scope ambiguities.

- (36) a. John eats meat  
 b. John used to eat meat<sup>4</sup>  
 c. GEN[e : eat(e) ∧ agent(e, j)][  
 GEN[y : meat(y)][patient(e, y)]]

Consecutively, (37) is a generic statement over both entities and events:

- (37) a. [Lions eat meat]<sub>BARE</sub>

<sup>4</sup>Ignoring tense. The German *pflügen zu* ‘use to’ can also be used in present tense; unfortunately, there seems to be no equivalent present-tense construction in English.

- b. GEN[x : lion(x)][  
 GEN[e : eat(e) ∧ agent(e, x)][  
 GEN[y : meat(y)][patient(e, y)]]]

We are aware that some semanticists would replace some of the generic quantifications in (37b) by existential quantifications or even non-quantificational expressions. This illustrates, however, how difficult it is to find covert generic quantifiers compared to overt quantifiers as in (38).

- (38) a. [Most lions always eat some meat]<sub>ALL+DIV+MEIST</sub>  
 b. MOST[x : lion(x)][  
 ∀[e : eat(e) ∧ agent(e, x)][  
 SOME[y : meat(y)][patient(e, y)]]]

With increasing complexity of sentence structures—as in fictional texts—, it is simply impossible to determine all covert quantifiers unambiguously in the annotation process. However, if no overt quantifier appears in a statement and the statement still has a quantificational meaning then there must be a covert quantifier somewhere. Hence quantified statements without any overt quantifier should be labelled with the tag BARE:<sup>5</sup>

- (39) [The country people have knowledge enough]<sub>BARE</sub>, [but their way of imparting it is confused]<sub>BARE</sub>, [and not always honest]<sub>NEG</sub>. [The students from the towns and universities are sufficiently clever and orderly, but they are deficient in personal experience]<sub>BARE</sub>.

#### 4.5 Negation

Negation can occur in different syntactic positions and cause problematic cases during the annotation. If the quantifier or the scope in a universally or existentially quantified statement is negated, its meaning could be expressed as both a universal or existential quantification, following the negation rules for quantifiers:

- (40)  $\neg\forall[x : restr(x)][scope(x)]$   
 $\equiv \exists[x : restr(x)][\neg scope(x)]$

<sup>5</sup>It might be confusing why the passage about the country people in (39) is fragmented whereas the passage about the students is not. According to our annotation guidelines, two or more subsequent quantified statements should be joined if they receive the same tag and the restrictor or the scope stay the same. This condition is not fulfilled for the former passage where the restrictor firstly shifts from *country people* to *their way of impairing it* and the tag secondly changes from BARE to NEG.

$$(41) \quad \neg\exists[x : restr(x)][scope(x)] \\ \equiv \forall[x : restr(x)][\neg scope(x)]$$

The case becomes even more complicated with ambiguous negation lexemes. The determiner *no* could be analysed as  $\neg\exists$  or  $\neg$  GEN, hence the statement in (42a) could be analysed as both (42b) and (42c).

$$(42) \quad \begin{aligned} \text{a. } & [\text{No lion sleeps}]_{\text{NEG}} \\ \text{b. } & \neg\exists[x : lion(x)][sleep(x)] \\ & \equiv \forall[x : lion(x)][\neg sleep(x)] \\ \text{c. } & \neg \text{GEN}[x : lion(x)][sleep(x)] \\ & \stackrel{?}{\equiv} \text{GEN}[x : lion(x)][\neg sleep(x)] \end{aligned}$$

This means that one could find arguments to label (42a) with any of EXIST, ALL or BARE. Resolving (ambiguous) negation would require to know whether it applies to the quantifier, restrictor or scope of a statement, and a set of detailed definitions for how one should annotate cases as in (40–42). Again, such a detailed analysis does not fit our aim of developing a simple annotation procedure. The simplest solution to this issue is to assign a special tag NEG for quantified statements with any negation in it. NEG then replaces all other tags that one could assign:

$$(43) \quad \text{“[And there are many cases [...] in which we are obliged, and in which it is the real kindness, rather to write nothing than not to write]}_{\text{NEG}}\text{.” (Goethe, EA, p. 20)}$$

## 5 Generalising Interpretations

In the previous section, we presented a tagset for quantification. However, our research does not focus on quantified statements in general but only on generalising statements, which we consider to be a subset of quantified statements. The main purpose of our research is to find generalising statements in fictional works of literature to investigate their narratological function. Our working definition for generalisation results from previous work on the re-interpretation of universal quantifiers: Löbner (2005) already notes two interpretations for *every* (originally the German counterpart *jede*) as in (44), namely 1) a concrete quantification over contextually determined instances, and 2) a generic quantification over assumed instances. Similarly, Leslie et al. (2011) found that adults frequently judge universal statements as in (45) true, despite knowing that there are counterexamples. Leslie

		Trueness erroneously accepted	
		no	yes
Instances assumed	no	(i) All students in the semantics class take notes	(ii) All students in the semantics class are broke
	yes	(iii) All triangles have three sides	(iv) All ducks lay eggs

Table 4: Quantified statements with varying characteristics

et al. (2011) conclude that *all* can be interpreted as a generic quantifier instead of a universal quantifier, and calls this the “generic overgeneralisation effect”.

- (44) Every child is entitled to a place in school  
(45) All ducks lay eggs

According to these works, an “overgeneralised” (universal) statement seems to be characterised by two properties:

1. The quantification involves assumed instances, i.e. not all restricted instances are contextually determined.
2. The statement is accepted as true (in the context of utterance) although there is not enough evidence for its trueness (because of unknown instances), or there is evidence for its falseness (because of known counterexamples).

Quantified statements that fulfil both properties are clearly generalising whereas statements fulfilling none of them clearly are not. Statements that fulfil only one of the properties are harder to classify, which is why we want to briefly discuss them in the following.

For examples (i) and (ii) in Table 4, imagine a classroom situation in which all students take notes. Then there is no doubt that (i) is true. Furthermore, some (but not all) of the students look the worse for wear; hence there is not enough evidence to claim (ii) as being true. Thus, one could call (ii) a “generalisation” from a subset of the students to all of them, taking the ordinary meaning of “generalisation” into account. However, the additional (cognitive) process of generalising to any assumed instances is missing, which is why it is no generalising statement in our sense.

The restrictor in (iii) and (iv) includes all triangles or ducks, respectively. Hence the quantified instances include assumed/unknown instances. Still, even for triangles which one does not know, one can infer that they have three sides (because otherwise they would not be triangles by definition). Therefore, we do not consider such (*quasi-*) *definitional* statements (Leslie et al., 2011) to be generalising.<sup>6</sup>

Having only looked at universal quantification so far, we now claim that generalisation can occur with every natural-language quantifier, and frequently does so in literary texts. Many of the examples shown in previous sections are in fact generalising; additional examples for each tag are given in the appendix (A). Unlike Leslie et al. (2011) and Löbner (2005), we do not claim that the quantifier in a generalising statement is re-interpreted as generic quantifier. Based on our observations (cf. Sec. 6), we suppose that generic quantification is a frequent but not the only type of quantification used to express generalisation.

## 6 Corpus and Annotations

We currently construct a diachronic corpus of German fictional literature from 1650 to 1950. As of now, we annotated generalising interpretations in ten texts—excluding some additional texts from a pilot annotation for developing our annotation guidelines. CATMA<sup>7</sup> appeared to be most suitable for annotating fictional texts and became our tool of choice. In order to create a versatile dataset and save resources, we annotate only the beginning of every text (usually the first about 200 sentences).

Our annotation procedure is as follows: Each text is first annotated by two out of six student assistants. In a second step, two researchers glance over the text again, focusing on the statements that were annotated by at least one annotator, discuss the annotations and create an expert annotation as gold standard. Arguably, this procedure is prone to false negatives, i.e. statements that none of the annotators identified are likely to be missed while creating the gold standard.

We measure inter-annotator agreement on token-level (excluding punctuation) using  $\kappa$  (Fleiss et al., 2003), treating the occurring tag combinations as

<sup>6</sup>*Generic* is not the same as *generalising* by our definition: Since generic quantification can also be used to express definitional statements (e.g. *triangles have three sides*), not all generic statements are generalising statements.

<sup>7</sup><https://catma.de/>

GI+Q		GI		Q	
$\kappa$	$\sigma$	$\kappa$	$\sigma$	$\kappa$	$\sigma$
.67	.20	.68	.22	.85	.14

Table 5: Mean inter-annotator agreement ( $\kappa$ ) over all texts and corresponding standard deviations ( $\sigma$ ); see text for column meanings

Generalising statements						
ALL	MEIST	EXIST	DIV	BARE	NEG	Total
151	7	17	76	332	145	728

Table 6: Number of generalising statements in all texts

classes (i.e. none, ALL, ..., ALL+DIV, ...). Since the annotators vary between texts, we first compute the agreement separately for each text. The average agreement is shown in Table 5. On average, there is substantial agreement of 0.67 for annotating generalising statements (GI+Q). The relatively high standard deviation of 0.20 indicates that there is a great variance between texts and/or annotators. Additionally, we compute the agreement when just distinguishing between “no tag” and “any tag”, i.e. “not generalising” vs. “generalising” (GI). The values are almost identical to those for GI+Q, indicating that the overall agreement is mainly influenced by the agreement on what is a generalising interpretation. To estimate the applicability of our tagset, we also compute the agreement for only those tokens that received a tag by all annotators, i.e. those tokens where the annotators agreed that they are part of a generalising statement (Q). Here, we see an average agreement of 0.85, which is significantly higher than that for GI, indicating that annotating quantification is comparatively straightforward.

Annotation results can be found in Table 6. Within 2,791 annotated sentences (61,979 tokens), 728 generalising statements occur (in the gold standard), which have an average length of 17 tokens. We can see that most generalising statements use plain generic quantification (BARE); followed by universal quantification (ALL) and vague quantification (DIV). Generalising statements with existential quantification (EXIST) or majority quantification (MEIST) are far less common. Note, however, that the counts do not directly reflect a “generalisation potential” of the individual quantification types since some quantification types occur with a higher frequency than others in the first place. We removed the tag Z AHL from our annotation



guidelines because generalising interpretations for numerical quantifiers hardly occurred during the pilot annotation (and ever since then).

## 7 Related Work

While the data in the existing corpora that are annotated with phenomena related to generalisation usually originates from the domains traditional for NLP, such as news and internet communication, we investigate generalisation in fiction, a more complex domain. To the best of our knowledge, there are no corpora on generalisation or genericity for German and our work reduces this gap.

Most existing resources focus on noun phrases, often in the context of coreference resolution. A detailed survey on generics in the coreference resolution research can be found in [Nedoluzhko \(2013\)](#).

[Friedrich et al. \(2015\)](#) provide a survey of genericity-annotated corpora for English. They note that ACE corpora ([Mitchell et al., 2003](#); [Walker et al., 2006](#)) are most widely used, e.g. by [Reiter and Frank \(2010\)](#), to identify generic noun phrases using a supervised approach.

[Friedrich et al. \(2015\)](#) were the first to suggest an annotation scheme for generic statements where both clauses and their subject NPs are annotated with the labels “generic” and “non-generic”. However, they only consider kind-referring generics and exclude e.g. habitual statements. In a subsequent work, [Friedrich et al. \(2016\)](#) investigate both habituals and kind-referring generics as two separate situation entity types, alongside with states and events, in a sentence classification task.

Many of the existing works use a limited set of labels: generic/non-generic or generic/specific. [Herbelot and Copestake \(2010, 2011\)](#) use a tagset similar to ours: ONE, SOME, MOST, ALL, QUANT. However, their concept is quite different: The authors assume that covertly quantified NPs are not generic but underspecified and label those NPs according to how many members of a class they refer to.

[Bhakthavatsalam et al. \(2020\)](#) create a large knowledge base of generic statements. The metadata in this knowledge base includes the term (restrictor) and the quantifier. They also include statements with an overt quantifier but “generic interpretation”.

Contrasting the previously mentioned shallow annotation schemes, [Bunt et al. \(2018\)](#) / [Bunt \(2019\)](#) propose an annotation scheme for quantifi-

cation that consists of several layers for syntactic and semantic representations but does not incorporate solutions for generics/habituals, yet.

[Donatelli et al. \(2019\)](#) suggest to expand the existing Abstract Meaning Representation (AMR) framework for the semantic annotation of sentences ([Banarescu et al., 2013](#)) by marking aspect and tense. As for aspect, they include such categories as “habitual” that characterise a regular recurrence of an event or state, and “stable” that characterises states and includes generalisations over kinds.

In comparison to the discussed approaches, we do not limit ourselves to NPs or clauses but annotate entire statements. Our tagset provides tags for overt quantification as well as covert generic quantification. Regarding the annotation of generalising statements, we jointly consider generalisations about entities, events and other types, that have been predominantly studied separately in the past.

## 8 Conclusion and Future Work

In this paper, we presented an annotation scheme for quantified phenomena using the example of generalisation. Our tagset matches every quantifier occurring in natural language to a particular tag, based on semantic criteria. We propose a shallow annotation that combines quantifier, restrictor and nuclear scope into a common annotation span, where the smallest unit of annotation is a clause. This approach is suitable for large-scale annotations which aim to investigate the distribution of quantified phenomena in a corpus, or to mark quantified statements to serve as feature input in follow-up applications. As a first step in this direction, we introduce our corpus of fictional texts that are annotated with generalising interpretations (among other phenomena).<sup>8</sup> Moreover, we are working on an automatic tagger for generalising statements.

## Acknowledgements

Our special thanks go to Anke Holler and the other members of the MONA project group, who advised us on the development of this work. We further thank the anonymous reviewers for their valuable comments. This work is funded by Volkswagen Foundation (Dönicke, Gödeke), and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 424264086 (Varachkina).

<sup>8</sup>Our corpus and annotation guidelines are accessible at <https://gitlab.gwdg.de/mona/korpus-public>.

## References

- Jeremy Avigad and Richard Zach. 2020. [The Epsilon Calculus](#). In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, fall 2020 edition. Metaphysics Research Lab, Stanford University.
- Laura Banarescu, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffith, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, and Nathan Schneider. 2013. Abstract meaning representation for sembanking. In *Proceedings of the 7th linguistic annotation workshop and interoperability with discourse*, pages 178–186.
- Maria Becker, Alexis Palmer, and Anette Frank. 2016. [Argumentative texts and clause types](#). In *Proceedings of the Third Workshop on Argument Mining (ArgMining2016)*, pages 21–30, Berlin, Germany. Association for Computational Linguistics.
- Sumithra Bhakthavatsalam, Chloe Anastasiades, and Peter Clark. 2020. [GenericsKB: A knowledge base of generic statements](#). *arXiv preprint arXiv:2005.00660*.
- Harry Bunt. 2019. [A semantic annotation scheme for quantification](#). In *Proceedings of the 13th International Conference on Computational Semantics - Long Papers*, pages 31–42, Gothenburg, Sweden. Association for Computational Linguistics.
- Harry Bunt, James Pustejovsky, and Kiyong Lee. 2018. [Towards an ISO standard for the annotation of quantification](#). In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).
- Noel Burton-Roberts. 1976. [On the generic indefinite article](#). *Language*, 52(2):427–448.
- Gregory Carlson. 2011. Genericity. In Claudia Maienborn, Klaus von Heusinger, and Paul Portner, editors, *Semantics: An International Handbook of Natural Language Meaning*, chapter 47, pages 1153–1185. De Gruyter Mouton.
- Gennaro Chierchia and Sally McConnell-Ginet. 2000. *Meaning and grammar: An introduction to semantics*, 2nd edition. MIT press.
- Ariel Cohen. 2004. [Existential generics](#). *Linguistics and Philosophy*, 27(2):137–168.
- Lucia Donatelli, Nathan Schneider, William Croft, and Michael Regan. 2019. [Tense and aspect semantics for sentential AMR](#). *Proceedings of the Society for Computation in Linguistics*, 2(1):346–348.
- Dudenredaktion. n.d. „manch“ auf Duden online. URL: <https://www.duden.de/node/151861/revision/151897>.
- Joseph L. Fleiss, Bruce Levin, and Myunghee Cho Paik. 2003. *The measurement of interrater agreement*, 3rd edition, chapter 18. John Wiley & Sons.
- Janet Dean Fodor and Ivan A Sag. 1982. [Referential and quantificational indefinites](#). *Linguistics and philosophy*, 5(3):355–398.
- Theodor Fontane. 1995. *The Stechlin*. Translated by William L. Zwiebel. Germ Series. Camden House. URL: <https://books.google.de/books?id=7yh0ZJjNgxC>.
- Theodor Fontane. 2012. [Der Stechlin](#). In *TextGrid Repository*. Digitale Bibliothek. URL: <https://hdl.handle.net/11858/00-1734-0000-0002-AECF-D>.
- Theodor Fontane. 2017. [Der Stechlin](#). In Berenike Herrmann and Gerhard Lauer, editors, *KOLIMO. A corpus of Literary Modernism for comparative analysis*. Originally published in Fontane (2012).
- Annemarie Friedrich and Alexis Palmer. 2014. [Situation entity annotation](#). In *Proceedings of LAW VIII - The 8th Linguistic Annotation Workshop*, pages 149–158, Dublin, Ireland. Association for Computational Linguistics and Dublin City University.
- Annemarie Friedrich, Alexis Palmer, Melissa Peate Sørensen, and Manfred Pinkal. 2015. [Annotating genericity: a survey, a scheme, and a corpus](#). In *Proceedings of The 9th Linguistic Annotation Workshop*, pages 21–30, Denver, Colorado, USA. Association for Computational Linguistics.
- Annemarie Friedrich, Alexis Palmer, and Manfred Pinkal. 2016. Situation entity types: automatic classification of clause-level aspect. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1757–1768.
- Johann Wolfgang Goethe. 2012. [Die Wahlverwandtschaften](#). In *TextGrid Repository*. Digitale Bibliothek. URL: <https://hdl.handle.net/11858/00-1734-0000-0006-6A93-D>.
- Johann Wolfgang von Goethe. 19--?. [Elective affinities : a novel](#). URL: <https://archive.org/details/electiveaffiniti00goetuoft/page/12/mode/2up>.
- Johann Wolfgang von Goethe. 2017. [Die Wahlverwandtschaften](#). In Berenike Herrmann and Gerhard Lauer, editors, *KOLIMO. A corpus of Literary Modernism for comparative analysis*. Originally published in Goethe (2012).
- Martin Hackl. 2009. [On the grammar and processing of proportional quantifiers: most versus more than half](#). *Natural language semantics*, 17(1):63–98.
- Irene Roswitha Heim. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts.
- Aurelie Herbelot and Ann Copestake. 2010. [Annotating underquantification](#). In *Proceedings of the*

- Fourth Linguistic Annotation Workshop, pages 73–81, Uppsala, Sweden. Association for Computational Linguistics.
- Aurelie Herbelot and Ann Copestake. 2011. [Formalising and specifying underquantification](#). In *Proceedings of the Ninth International Conference on Computational Semantics (IWCS 2011)*.
- Berenike Herrmann and Gerhard Lauer. 2017. KOLIMO. A corpus of Literary Modernism for comparative analysis. URL: <https://kolimo.uni-goettingen.de/about>.
- Klaus von Heusinger. 2000. [The reference of indefinites](#). In Klaus von Heusinger and Urs Egli, editors, *Reference and anaphoric relations*, pages 247–265. Springer.
- Hans Kamp. 1981. A theory of truth and semantic representation. In Jeroen A. G. Groenendijk, Theo M. V. Janssen, and Martin J. B. Stokhof, editors, *Formal Methods in the Study of Language (Part I)*, pages 277–322. Mathematisch Centrum, Amsterdam.
- Edward Keenan and Denis Paperno. 2012. *Handbook of quantifiers in natural language*, volume 90. Springer.
- Oriana Kilbourn-Ceron. 2014. [Almost: scope and covert exhaustification](#). In *West Coast Conference on Formal Linguistics (WCCFL)*, volume 32, pages 121–130.
- Manfred Krifka. 2016. Quantification and information structure. In *The Oxford handbook of information structure*.
- Manfred Krifka and Claudia Gerstner. 1987. An outline of genericity. Seminar für natürlich-sprachliche Systeme der Universität Tübingen.
- Manfred Krifka, Francis Jeffry Pelletier, Gregory Carlson, Alice ter Meulen, Gennaro Chierchia, and Godehard Link. 1995. Genericity: An introduction. In Greg N. Carlson and Francis Jeffry Pelletier, editors, *The Generic Book*, pages 1–124. University of Chicago Press.
- Silke Lahn and Jan Christoph Meister. 2016. *Einführung in die Erzähltextanalyse*, 3 edition. J.B. Metzler, Stuttgart.
- Sarah-Jane Leslie, Sangeet Khemlani, and Sam Glucksberg. 2011. [Do all ducks lay eggs? the generic overgeneralization effect](#). *Journal of Memory and Language*, 65(1):15–31.
- Sarah-Jane Leslie and Adam Lerner. 2016. [Generic Generalizations](#). In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, winter 2016 edition. Metaphysics Research Lab, Stanford University.
- David Lewis. 1975. Adverbs of quantification. In Edward L. Keenan, editor, *Formal Semantics of Natural Language*. Cambridge University Press.
- Sebastian Löbner. 2005. *Quantoren im GWDS*, pages 171–192. Max Niemeyer Verlag.
- Louise McNally. 1998. [Existential sentences without existential quantification](#). *Linguistics and Philosophy*, 21(4):353–392.
- Alexis Mitchell, Stephanie Strassel, Mark Przybocki, JK Davis, George Doddington, Ralph Grishman, Adam Meyers, Ada Brunstein, Lisa Ferro, and Beth Sundheim. 2003. ACE-2 Version 1.0. *Linguistic Data Consortium, Philadelphia*. URL: <https://catalog.ldc.upenn.edu/LDC2003T11>.
- Friederike Moltmann. 2013. [The semantics of existence](#). *Linguistics and Philosophy*, 36(1):31–63.
- Anna Nedoluzhko. 2013. [Generic noun phrases and annotation of coreference and bridging relations in the Prague dependency treebank](#). In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pages 103–111, Sofia, Bulgaria. Association for Computational Linguistics.
- Barbara H. Partee. 1990. Domains of quantifications and semantic typology. Mid-America Linguistics Conference.
- Nils Reiter and Anette Frank. 2010. Identifying generic noun phrases. In *Proceedings of the 48th annual meeting of the association for computational linguistics*, pages 40–49.
- Laura Rimell. 2004. Habitual sentences and generic quantification. In *Proceedings of WCCFL*, volume 23, pages 663–676.
- Achim Vesper. 2014. *Literatur und Aussagen über Allgemeines*, pages 181–196. De Gruyter (A).
- Christopher Walker, Stephanie Strassel, Julie Medero, and Kazuaki Maeda. 2006. Ace 2005 multilingual training corpus. *Linguistic Data Consortium, Philadelphia*, 57:45.
- Yoad Winter. 1997. [Choice functions and the scopal semantics of indefinites](#). *Linguistics and Philosophy*, 20(4):399–467.
- Edward N. Zalta. 2020. [Gottlob Frege](#). In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, fall 2020 edition. Metaphysics Research Lab, Stanford University.

## A Examples for Quantified Statements with Generalising Interpretations

- (46) [In all natural objects with which we are acquainted, we observe immediately that they have a certain relation to themselves]<sub>ALL</sub>. (Goethe, EA, p. 53)

- (47) [Heroism is the exception]<sub>BARE</sub> [and mostly the product of a separate situation]<sub>MEIST</sub>. (Fontane, Stechlin, p. 19)
- (48) [But there are also fruits which are not outward, which are of the true germinal sort, and which develop themselves sooner or later in a beautiful life]<sub>EXIST</sub>. (Goethe, EA, p. 43)
- (49) On New Year's Day he was to follow him, and spend the Carnival at his house in the city, where Luciana was promising herself infinite happiness from a repetition of her charmingly successful pictures, [as well as from a hundred other things]<sub>ZAHL</sub>; (Goethe, EA, p. 240)
- (50) In providing against accidents, [which, though common, yet only too often find us unprepared]<sub>DIV</sub>, they thought it especially necessary to have at hand whatever is required for the recovery of drowning men (Goethe, EA, p. 48)
- (51) “[We are strange creatures]<sub>BARE</sub>,” said Edward, smiling. “[If we can only put out of sight anything which troubles us, we fancy at once we have got rid of it]<sub>BARE</sub>.” (Goethe, EA, p. 25)
- (52) [A fellow from Friesack better not have a name like Raoul]<sub>NEG</sub>. (Fontane, Stechlin, p. 4)
- (14') [Die Männer denken mehr auf das Einzelne, auf das Gegenwärtige]<sub>BARE</sub> [...] (Goethe, WV)
- (16') [...] [da sind zum dritten die, die den Meeresgrund absuchen wie 'ne Wiese]<sub>EXIST</sub> (Fontane, Stechlin)
- (17') [Aber es gibt dergleichen noch, es muß dergleichen geben oder doch wieder geben]<sub>EXIST</sub>. (Fontane, Stechlin)
- (25') Die Grafschaft habe immer konservativ gewählt; es sei Ehrensache, wieder konservativ zu wählen. »[Und ob die Welt voll Teufel wär']<sub>BARE</sub>. (Fontane, Stechlin)
- (26') Da kommt hier so Anno Domini ein Burggraf ins Land, und das Land will ihn nicht, [und er muß sich alles erst erobern, die Städte beinah und die Schlösser gewiß]<sub>ALL</sub>. (Fontane, Stechlin)
- (28') Wenn es mir erlaubt ist, dem Scheine nach weit auszuholen, so sind wir bald am Platze. (Goethe, WV)
- (29') [Unsere vortreffliche Vorsteherin läßt mich gewöhnlich die Briefe lesen, in welchen sie Beobachtungen über ihre Zöglinge den Eltern und Vorgesetzten mitteilt]<sub>DIV</sub>. (Goethe, WV)
- (30') »Es betrifft unsern Freund, den Hauptmann,« antwortete Eduard. »Du kennst die traurige Lage, [in die er, wie so mancher andere, ohne sein Verschulden gesetzt ist]<sub>DIV</sub>. (Goethe, WV)

## B German Versions of Translated Examples

Numbers are identical to those of the corresponding translations in the main part of the paper. Note that the German version sometimes receives another annotation than its English translation because the quantification may differ.

- (8') [...] [und augenblickliche Hülfe ist doch immer das, was auf dem Lande am meisten vermißt wird]<sub>ALL</sub>. (Goethe, WV)
- (9') [...] trenne alles, was eigentlich Geschäft ist, vom Leben! (Goethe, WV)<sup>9</sup>
- (13') Gedenken wir nur des Kalks, [der zu allen Säuren eine große Neigung, eine entschiedene Vereinigungslust äußert]<sub>ALL</sub>! (Goethe, WV)
- (31') Dieser sein Eintritt ins Regiment fiel so ziemlich mit dem Regierungsantritt Friedrich Wilhelms IV. zusammen, [und wenn er dessen erwähnte, so hob er, sich selbst persiflierend, gerne hervor, »[daß alles Große seine Begleiterscheinungen habe]<sub>ALL</sub>«]<sub>BARE</sub>. (Fontane, Stechlin)
- (39') [Die Landleute haben die rechten Kenntnisse]<sub>BARE</sub>; [ihre Mitteilungen aber sind konfus]<sub>BARE</sub> [und nicht ehrlich]<sub>NEG</sub>. [Die Studierten aus der Stadt und von den Akademien sind wohl klar und ordentlich]<sub>BARE</sub>, [aber es fehlt an der unmittelbaren Einsicht in die Sache]<sub>BARE</sub>. (Goethe, WV)
- (43') [Und doch ist es in manchen Fällen [...] notwendig und freundlich]<sub>DIV</sub>,

<sup>9</sup>In English, this sentence is declarative, whereas in German it is imperative. Therefore, it is not annotated in German.

- [lieber nichts zu schreiben, als nicht zu schreiben]<sub>NEG.</sub> (Goethe, WV)
- (46') [An allen Naturwesen, die wir gewahr werden, bemerken wir zuerst, daß sie einen Bezug auf sich selbst haben]<sub>ALL.</sub> (Goethe, WV)
- (47') [Heldentum ist Ausnahmezustand]<sub>BARE</sub> [und meist Produkt einer Zwangslage]<sub>MEIST.</sub> (Fontane, Stechlin)
- (48') [...] [aber es gibt auch verschlossene Früchte]<sub>EXIST.</sub>, [die erst die rechten, kernhaften sind und die sich früher oder später zu einem schönen Leben entwickeln]<sub>BARE.</sub> (Goethe, WV)
- (49') Auf's Neujahr sollte ihm dieser folgen und das Karneval mit ihm in der Stadt zubringen, [wo Luciane sich von der Wiederholung der so schön eingerichteten Gemälde sowie von hundert andern Dingen die größte Glückseligkeit versprach]<sub>ZAHL</sub> (Goethe, WV)
- (50') [Da man auch die gewöhnlichen und dessen ungeachtet nur zu oft überraschenden Notfälle durchdachte, so wurde alles, was zur Rettung der Ertrunkenen nötig sein möchte, um so mehr angeschafft]<sub>DIV</sub> (Goethe, WV)
- (51') »[Wir sind wunderliche Menschen]<sub>BARE</sub>,« sagte Eduard lächelnd. »[Wenn wir nur etwas, das uns Sorge macht, aus unserer Gegenwart verbannen können, da glauben wir schon, nun sei es abgetan]<sub>BARE.</sub> (Goethe, WV)
- (52') [Wer aus Friesack is, darf nicht Raoul heißen]<sub>NEG.</sub> (Fontane, Stechlin)