Adam Mickiewicz University in Poznań
Department of Scandinavian Studies

Generics in Norwegian – A Cognitive Analysis

Anna Kurek-Przybilski

A monster is not a monster if it doesn't scare you.

Jonathan Carroll

Contents

List of Tables

List of Figures

1   Introduction

Human language not only makes communication between people possible and effective, but also very diverse. We can talk about something in a concrete and literal way but we can also construct abstract and ambiguous utterances, as well as we can generalise about a given object or a group of people, saying for instance that *Cars have four wheels* or that *The French like wine*. Even though not all cars might have four wheels and there are certainly some French people who do not appreciate wine, such generalisations are not perceived incorrect or awkward when it comes to their interpretation. Sentences like these are an expression of GENERICITY – a phenomenon that allows the speakers of a given language to make generalisations. Both the ability to utter such statements and interpret them is a fascinating quality of the human language, and it has been observed in many languages that have been studied to date and it is a feature of any human language to express genericity.

The expressions of genericity are as diverse as the phenomenon itself. Depending on the language group, as well as the context in which generic sentences appear, they can include generic NPs and be written in either present or past tense. The latter is not the most common when it comes to genericity but sentences with past tense generics are nevertheless possible. What is more, certain predicates promote generic readings by referring to whole kinds, e.g. *to be numerous*. Those are called KIND-PREDICATES and will be discussed in more detail in chapter 2.

The available research on the matter focuses on a few main areas, such as theoretical studies, corpus linguistics, cognitive approach to generics, as well as a great number of experimental studies conducted in recent years. All of these have proved to be successful in different domains. For instance, purely theoretical studies such as Carlson (1977) and (1982), Carlson and Pelletier (1995) and Mari et al. (2013a) among others, have contributed greatly to the development of the applied terminology. Corpus studies and analyses of individual generic sentences, as for instance Oosterhof (2008), have made it possible to outline the NP types that take on generic meaning in certain languages, whereas experimental approaches to the phenomenon have proved that genericity might not only be a linguistic but also a psychological matter (see for instance Gelman and Tardif 1998; Leslie 2007; Leslie et al. 2011; Ionin et al. 2011).

All of the studies mentioned above and discussed further in this dissertation have

great value and are often novel in the domain. However, the majority of the available publications focus mainly on widely spoken languages such as English, Chinese, French or German, or whole language groups (for instance the account of genericity in Romance languages was given by Kabatek 2013). There are a number of studies devoted to languages with fewer speakers such as Karczewski's (2016) account on generics in Polish, Carlsson's (2012) analysis of Swedish, the comparative study of Farkas and de Swart (2009) where Hungarian generics are analysed in contrast to English and French or Molnár's (2014) account of Hungarian and German generics. However, a comprehensive analysis of Norwegian generics is lacking.

The goal of this project is to bridge the gap in research and to provide an empirical study of Norwegian generics. Since the phenomenon is rather diverse, the main research question concerns therefore the way generics can be rendered in Norwegian, considering especially the NP types used in such texts and sentences. What is more, the role of the context and speakers' cognition in expressing genericity will be considered. These aspects of genericity will be examined by using different strategies and by focusing on actual language use from a perspective of cognitive linguistics, especially the models proposed by Radden, Langacker and Leslie, which will be presented in the following chapter.

This chapter is divided into three main sections. The first section describes different types of references and the notions connected to the phenomenon. The second section lays out the methodological basis for the study design, namely the MIXED METHODS RESEARCH (MMR) and the statistical methods used in the empirical chapters of the dissertation. The structure of the dissertation is presented in the last section of this chapter.

1.1 Types of reference

Making a generalisation about a given object, a species or a group of people requires that a speaker *refers* to it. In the scholarly literature, genericity is therefore seen as a type of reference and it stands in the opposition to the specific reference. In this section different kinds of reference will be described, as well as the notions of reference and denotation themselves. However, the focus of this project is cognitive so logical and philosophical accounts of reference will be omitted.[1] The publications and theories mentioned in this section are mostly seminal works on the matter or general studies devoted to semantics.

Reference, as used in a linguistic context, asserts a feature or a set of features to a given object (Lyons 1977: 177-197). In other words, a proposition, when uttered, creates a link to an element of the world – it refers to it by picking out given features that define the described object. There are also different types of reference. For example, in the

---

[1]For a detailed discussion of reference, sense and denotation, the reader is referred to the works of Frege ([1892] 2010) and Lyons (1977) among others.

sentence (1) a speaker may mean that barking is in fact a feature characteristic for most (or maybe even all) dogs, based on one's proposition.[2]

(1)   Dogs bark.

On the other hand, a receiver of such information might still classify a non-barking dog as a dog, even if it happens that a given animal may not match the characteristics from the sentence (1). Referring to all dogs in this case is a way of asserting barking to the whole kind, not only a given group of animals. In this case the reference is GENERIC, namely it concerns a wide range of representatives of a given group.

As Lyons points out, reference is an utterance-dependent notion (Lyons 1977: 180), not an NP-dependent one. This means that in the sentence (1) it is not the noun form itself that refers to all dogs but rather the whole sentence that promotes such reading. Asserting the act of barking to the species, makes the receiver of the information connect it intuitively to all individuals from the group 'dogs'. We can also imagine that in a different context, one might as well refer to some particular group of dogs that tend to bark, without stating the fact that most or all dogs do that.[3] What is more, it seems crucial that the speaker chooses an appropriate referring expression in order to achieve a desired meaning of a given sentence. Such an expression might come in a form of a proper name, a definite noun-phrase or a pronoun (Lyons 1977: 180).

One can distinguish words and phrases that can be used to refer to someone or something, the so-called REFERRING EXPRESSIONS (Lyons 1977: 181). These are used in sentences such as (2), where both 'the cows' and 'cows' are referring expressions. The example in (2a) is an INDIVIDUATIVE REFERENCE, and the sentence in (2b) is a GENERIC REFERENCE.[4]

(2)   a.   The cows are over there.

       b.   Cows are four-legged animals.

The two notions designate types of references, where the speaker either talks about particular instances or the whole kind in question. Languages with grammaticalised articles, definite and/or indefinite, utilise the articles to modify the reference. Another classification, used also in cognitive grammar, includes INDEFINITE and DEFINITE reference (Radden and Dirven 2007: 87). In the case of indefinite reference, the speaker assumes that the hearer does not have access to the instance in question through their general knowledge or the context, whereas the definite reference implies such knowledge.

---

[2]Different types of generic generalisations are discussed in section 2.4.3.

[3]Lyons discusses also the issue of reference and truth (Lyons 1977: 181-185).

[4]The notion of NON-REFERRING EXPRESSION, which stands in contrast to the referring one, is perceived by some researchers as a generic reference (cf. 'A cholecystectomy is a serious procedure', Saeed 2009: 26). What is more, in some normative grammar books generic reference is sometimes associated with 'general reference', see e.g. Teleman et al. (1999) and Pettersson (1976) among others. However, the two should not be confused with each other as general reference is time-bound, which cannot be said about generic reference (Lyons 1977: 193-194).

Generic reference mentioned before in contrast to individuative reference, has a special status in that it is used when the speaker talks about the whole class or the whole kind and assumes that the discourse and/or the general knowledge of the hearer allow for a correct interpretation of such expression.

Generic reference is often context bound as its interpretation may vary when used in a different context. Such dependency implies also that both the speaker and the hearer interpret a given sentence in the same way. Whether or not an NP, a sentence or a text is considered generic, is often a matter of judgement and language intuition (see also sections 2.4 and 2.4.1).

When it comes to the opposition definite-indefinite reference, one differentiates between SPECIFIC and NON-SPECIFIC reference, which can be observed in the examples in (3) by Radden and Dirven (2007: 95).

(3)   a. I want to marry *an American.* He lives in Kalamazoo.   [indefinite, specific]

   b. I want to marry *an American.* He should be rich.   [indefinite, non-specific]

The sentence (3a) implies that from a group of Americans that the speaker has in mind (presumably the whole nation), there is one specific person that fits the description, namely he lives in Kalamazoo. In the example (3b), 'an American' is defined by the second sentence ('He should be rich'), implying that the speaker does not mean any particular American but anyone that would fulfil the given criteria.

Another notion that is closely connected to reference is DENOTATION. Certain linguists claim that the two notions basically designate the same process (cf. Saeed 2009), whereas others emphasise different meanings of reference and denotation. The contrast between the two was illustrated for instance by Kreidler:

> Reference is the relation between a language expression such as *this door, both doors, the dog, another dog* and whatever the expression pertains to in a particular situation of language use, including what a speaker may imagine. Denotation is the potential of a word like door or dog to enter into such language expressions. Reference is the way speakers and hearers use an expression successfully; denotation is the knowledge they have that makes their use of expressions successful. (Kreidler 1998: 43)

Lyons claims that in some cases denotation can be possible whereas reference cannot, as in the example below:

(4)   The present King of France is bald. (Lyons 1977: 182)

The sentence (4) shows how the abstract concept can exist independently of the reality. To most people who are familiar with the current political situation in France, such a sentence will seem grammatically correct but false. There are many more examples like this one where the DENOTATUM (the person or the object one refers to) exists only putatively (Allan 2014: 68). We can also imagine another scenario to be true: the speaker might

think that a given object or a person exists, whereas the hearer might suspect the opposite or interpret the utterance differently. Such a reference is therefore always context bound, both for the speaker and the hearer (Allan 2014: 68).

Context bound sentences do not need to be true in general but if the denotation is possible despite the lack of reference, communication is still considered successful. This is illustrated by Lyons in his example sentences (5a) and (5b):

(5)   a. Mr Smith is looking for the Dean.

        b. Mr Smith is looking for Professor Brown. (Lyons 1977: 192)

The speaker, in this case Mr Smith, might not know who the Dean is but the reference from the sentence (5a) is still possible for the hearer for whom the Dean refers to Professor Green for instance. In the sentence (5b) though, the reference might be possible only for the speaker who considers Professor Brown to be the Dean. In such case, the hearer, knowing that the Dean is in fact Professor Green and not Professor Brown, would think of one of the professors but not of the Dean (Lyons 1977: 192).

Referring to something and understanding such reference implies that one comprehends not only the particular context but, first and foremost, the meaning of the words used in a sentence. The question of meaning has been taken up by countless researchers and is the subject of study in several disciplines such as linguistics, philosophy, logic and psychology. The amount of definitions and interpretations of the notion is therefore abundant. Here we shall focus only on the linguistic accounts of meaning (the list of sources provided in this chapter is by no means exhaustive).

In the literature on the matter (see e.g. Lyons 1977 and Allan 2014, 2016) it is often mentioned that the act of referring to something or someone is closely connected to the sense or meaning of a given word or phrase. The matter has been widely discussed by Ogden and Richards in their classical work on meaning (*The Meaning of Meaning*, first published in 1923), and later re-defined and re-interpreted by others (cf. Lyons 1977 and Frege [1892] 2010).

Lyons defines the difference between sense and reference as *a distinction between reference and meaning* (Lyons 1977: 197), assigning the notion of sense exclusively to philosophy. This way, sense could be understood as the synonym of meaning, which has also been observed in recent works on general semantics (Allan 2016; Riemer 2016).

As has been already said, reference is the relation between the abstract concept and the object that this concept represents. By creating a link between the two, we refer to something or, simply, we *mean* something by uttering a given word or a sentence. In order to refer to a particular entity or a given person, place etc. we use proper names, as Frege points out:

> A proper name (word, sign, sign combination, expression) *expresses* its sense, *stands for* or *designates* its reference. By means of a sign we express its sense

and designate its reference. (Frege [1892] 2010: 40)

According to this statement, every expression has a given sense (or in Lyons' terms: meaning). As we have seen in the examples (5a) and (5b), an expression can have its meaning which would make a denotation possible but the reference might not be feasible. Another example illustrating this paradox is the classic sentence by Frege:

(6)   The Morning Star is the Evening Star.

Both *the Morning Star* and *the Evening Star* denote the same element, namely the planet Venus seen from the Earth. The expressions do not have the same sense but they do stand for the same reference (planet Venus). It is worth mentioning though that this is only the case for those who realise that the notions describe the same planet. For anyone who does not possess such knowledge, a part of this sentence will be false (Frege [1892] 2010: 41). Just like in the example (5) Mr Smith was using an expression with false reference (Professor Brown is not the Dean), similarly someone could claim that the Morning Star is not the Evening Star and therefore sentence (6) should be considered grammatically correct but false when it comes to its reference. This dependency can be summarised in the following way:

> The regular connection between a sign, its sense, and its reference is of such a kind that to the sign there corresponds a definite sense and to that in turn a definite reference, while to a given reference (an object) there does not belong only a single sign. The same sense has different expressions in different languages or even in the same language. (Frege [1892] 2010: 38)

What Frege suggests is that a given sign can only have one sense and one definite reference (*the Morning Star*=Venus) but a given object can be described with the use of more than one sign (Venus=*the Morning star, the Evening Star*).

Now that we have seen the link between sense and reference, let us focus on another opposition often mentioned in semantics, namely SENSE and DENOTATION.

> The status of the relationship between denotation and reference, on the one hand, and denotation and sense, on the other, is not, however, affected by our failure to draw these distinctions; and we could not do so without philosophical commitment, except at the cost of introducing a further set of technical terms. (Lyons 1977: 207)

The relationship between sense and denotation is slightly more complex than that of sense and reference. As we have already seen, reference occurs when a sign with a given sense makes a link to an object in the real world. Reference is therefore utterance-bound and context-bound. Denotation, on the other hand, *holds independently of particular occasions of utterance* (Lyons 1977: 208). A given lexeme, for instance 'dog', will therefore

denote a whole group of animals classified as dogs, whereas expressions such as 'my dog' and 'Katy's dog' refer to particular entities. Denotation can therefore be interpreted as an abstract category in human language.

Another major difference between oppositions sense-reference and sense-denotation is the so called TRUTH VALUE. The speakers of a given language haves the possibility to denote categories of objects that might not exist in reality or might no longer exist. One can therefore utter a fairly abstract sentence about unicorns or dinosaurs which will be understood by other speakers of the language. Such an utterance makes denotation possible, whereas reference would probably pose problems to most speakers of any language, as neither unicorns or dinosaurs exist.[5]

An interesting aspect of different reference types, also the generic one, is the use of tenses. Whereas individuative reference may be expressed with virtually any tense, there are certain restrictions when it comes to generic references. Lyons (1977: 194) points out that generic sentences, with certain exceptions, are not only timeless but also *tenseless* and *aspectless*.

Timelessness and *aspectlessness* mean that a given sentence holds irrespectively of the time, as is illustrated in (7). Lions being friendly beasts is not associated with any particular point in time, despite the fact that all sentences are written in present tense.

(7) a. The lion is a friendly beast.

     b. A lion is a friendly beast.

     c. Lions are friendly beasts.

The *tenselessness* of generic references manifests itself in the fact that the truths they express hold irrespectively of the tense they are uttered in. Tense used in such expressions is sometimes called the GENERIC TENSE (Dahl 1975: 99). According to Dahl, stating something in generic tense, be it in past, present or future, means that a state, a law or a characteristic feature holds at a certain time and is not valid for all time (Dahl 1975: 103). 'Certain time' implies that the fact that e.g. cows have for legs holds for the time being but also leaves the possibility to change the situation – maybe in the future most cows will have five legs due to genetic modifications. Potential changes in meaning which are implied by generic present, past and future were illustrated by Dahl (1975: 103) in the following example:

(8) When I was a boy, I wrote with my left hand, but now I write with my right hand, although I will probably write with my left hand again when I grow older.

As has been said, all sentences in (7) can be perceived as tenseless, as the present tense in generic sentences is neutral in relation to time (Lyons 1977: 194). However, an example such as (9), even though written in the past tense (generic past), can still be

_____

[5]See also Frege's ([1892] 2010: 42) take on the truth value.

interpreted as tenseless and timeless. The use of past tense is connected to the fact that the speaker knows the dinosaurs are extinct. Nevertheless, the part of the proposition that marks the tense does not bear the generic reading – that reading is rendered by the predicate and the NP (Lyons 1977: 194).

(9)   The dinosaur was a friendly beast.

However, the use of tense can influence the interpretation of a given sentence and this may be observed also in the case the of generic sentences. Even though they are indeed considered timeless and tenseless, certain exceptions can occur. For instance, the use of Present Continuous in English is possible in the so-called HABITUAL SENTENCES. Such sentences render regularities and habits, not necessarily generalisations as has been seen in the case of *classic* generic reference referring to a kind. Habitual sentences are considered generic by certain researchers, whereas others perceive only true kind-reference as generic.[6] In this dissertation, habituals are considered as a type of generics.

The sentence (10), written in the continuous tense, can have two interpretations – it might mean that Mary is drinking wine at the moment, as well as that wine drinking after dinner is her habit (habitual reading). The second interpretation is a generic one, whereas the first one provides a literal interpretation of the sentence. The use of continuous tense in a generic sentence is not typical but it is acceptable in given contexts and especially with episodic reading of habitual sentences (see e.g. Krifka et al. 1995: 37).

(10)   Mary is drinking wine after dinner.

Interpreting generic sentences depends also on TRUTH-CONDITIONS. Whether or not a sentence is true, it can be constructed and uttered. We can imagine someone saying for instance:

(11)   Unicorns are friendly beasts.

Most people, if not everybody, would consider the sentence above correct but not true. It is not the matter of unicorns being friendly or unfriendly but the lack of referent of the word *unicorns* that makes the sentence false. We have seen a similar dependency in example (4) about the present king of France being bald. Sense and denotation of such sentences make it possible for a speaker to create them but the lack of reference does pose problems from a semantic point of view.

The types of references discussed in this section and the notions connected to the phenomenon, show the complexity of the process of referring to something. Different types of references will be discussed in greater detail in the next chapter, where formal and cognitive theories on genericity are presented. Throughout the dissertation the notions of reference and denotation are used but the truth-conditions are not analysed. Since the

---

[6]See chapter 2 for further discussion on two senses of genericity.

material fort the study consists of generic texts and sentences, either created for the sake of the project (the pilot research) or retrieved from existing sources (the corpus study and the AJT survey), the issue of truth-conditions is not central in the analysis.

1.2   Design of the study

When choosing a study design for this project, the goal was to describe genericity from the point of view of cognitive linguistics but also in a way that would capture many aspects of the phenomenon. The design of the study was crucial in order to obtain credible and representative data. In order to achieve this, MIXED METHODS RESEARCH (MMR) strategy was utilised. There are numerous reasons for the popularity of MMR in recent years, such as the *triangulation* and *complementarity* of the data (Hesse-Biber 2010: 3-4). Triangulation allows to approach a research question from many different perspectives, which then are combined in the final analysis. Such an approach provides more credible results than single-method studies. Complementarity of such projects means that one analysis strengthens the other, for instance the quantitative (QUAN) description is supported by the qualitative (QUAL) one, or the other way around, depending on the research question and the studied field.

The use of mixed methods research not only allows to study a given phenomenon from different perspectives in one research project, but it also provides a broader understanding of the problem in question. In contrast to multimodal methods, which utilise either qualitative or quantitative tests (Hesse-Biber 2010: 3-4), the MMR relies both on qualitative and quantitative data analyses. Such methodology is known as the 'third methodological movement' or the 'third research paradigm' (Creswell and Clark 2018). The two other paradigms are qualitative and quantitative ones.

There are numerous study designs within MMR paradigm and the choice of a given strategy must be done in accordance with the research question. For the main core of this project,[7] CONCURRENT TRIANGULATION mixed methods design was chosen. The strategy combines two types of data analysis, namely QUAL and QUAN, which performed simultaneously are then combined in order to provide a more detailed account of the phenomenon (Riazi 2016: 47).

The MMR design of this project does not exclude or interfere with the cognitive framework applied for the interpretation of the findings. The qualitative part of this project consists in analysis of the material from the point of view of cognitive linguistics and the models described in chapter 2. The QUAN analysis is crucial in order to show how genericity can be expressed in Norwegian and what cognitive status is assigned to the phenomenon. The QUAN part on the other hand, presents a number of statistical tests that allow to structure the data and see its distribution, providing also numerical

---

[7]See section 1.3 for the description of the project.

data that supports the descriptive analyses.

1.3   Structure of the dissertation

The dissertation is divided into 7 chapters. Chapters 1 and 2 present a theoretical background of the dissertation, presenting the notions connected to genericity together with its two senses and the current state of research. The third chapter presents the existing studies on the phenomenon in Mainland Scandinavian languages. The following parts of the thesis concern the analysis of generics in Norwegian: chapter 4 is the description of the pilot study conducted in Norway, chapter 5 gives a description and an analysis of the tailor-made corpus of generic texts used in the project. In chapter 6, I provide a description and an analysis of another survey conducted among the native speakers of Norwegian and based on the Acceptability Judgement Task. Chapter 7 is a comparison of the data retrieved both from the corpus and the surveys. Chapter 7 contains also main conclusions and implications for further research.

2 Theoretical models of genericity

In every language that has been studied to date,
generics have the least marked surface forms;
it is natural to conjecture that this is a linguistics universal.

Leslie (2007: 382)

The existing theories on genericity can be divided into three main categories, based on their approaches and interpretations of generics. One can therefore differentiate between 1) Calsonian-theories introduced by Carlson and other researchers from the Generic Group, 2) neo-Carlsonian theories developed by some members of the Generic Group and other researchers, as well as 3) cognitive theories which seem to be on the rise in recent years.

The first and second approach can be perceived as the *core* or somehow *classic* research on the matter as those theories were developed first. A more unified methodology and terminology was introduced, which was in fact the main purpose of *The Generic Book* published in 1995. Cognitive research in the field has shed new light on understanding of the notion of genericity, as well as it has allowed for more interdisciplinary approach. In newest studies on genericity, the researchers turn often to psychology and theories applied in disciplines such as cognitive linguistics, as well as neurolinguistics.

This chapter is organised in the following manner: I will first focus on the notion of genericity itself and present the existing research on the topic. The methodologies utilised in these studies will also be briefly presented. Section 2.1 concerns the use of the notions 'genericity' and 'generics' in linguistics, section 2.2 handles two senses of the phenomenon, whereas sections 2.3 and 2.4 provide an overview of the existing research. Most of the publications discussed in this chapter English and other widely spoken languages, as those are the subject of most of the seminal works. The available scholarly literature on genericity in Mainland Scandinavian languages will be described in chapter 3.

2.1 Generics and genericity

Even though it has been a known fact that language users can speak about kinds as well as individuals and that the two uses of nominal and verbal expressions may to some extent

differ, for a long time a coherent terminology describing this fact was lacking. This has inspired the Generic Group, consisting of Gregory N. Carlson, Gennaro Chierchia, Manfried Krifka, Godehard Link, Francis Jeffry Pelletier and Alice ter Meulen[1], to publish *The Generic Book* in 1995. Their main goal was to *develop a terminology to be recommended for use by other researchers* (Carlson and Pelletier 1995: viii). The notions established by the Generic Group have been widely used in numerous works on the matter that appeared after 1995 but not all linguists have adopted the terminology proposed by Carlson and Pelletier. For the most part, the notions used in this dissertation are those proposed by the Generic Group, unless they are terms coined after the publication of 1995, such as the generic generalisation types proposed by Leslie et al. (2011).

As has been mentioned in the previous chapter, one differentiates between GENER-ICS and GENERICITY. Both notions are used in several works on the phenomenon (see for instance Carlson and Pelletier 1995; Mari et al. 2013a), often without a very clear distinction between the two. It might therefore seem that the notions (at least to a certain degree) designate the same phenomenon in linguistics but when it comes to philosophy of language this is not necessarily the case.

In English, one can distinguish at least three main words that appear in discourse on the matter, namely *genericity, generics* and the adjective *generic*. The distinction between the two first notions present in English, does not seem to be the case in Norwegian for instance. The two words connected to the phenomena in Norwegian are the noun *generisitet* and the adjective *generisk*. In French literature on the matter one comes across the nouns *la généricité* and *le générique*, as well as the adjective *générique*. German terms include the noun *der Generizität* and the adjective *generisch*. German and Norwegian show the same structure of the notions connected to genericity – there is no equivalent to the English *generics*, whereas in French one might suppose that *le générique* (used sometimes also in the plural form *des génériques*, see Dahl 1985: 57) is indeed such an equivalent.

Since the majority of the works on genericity are in English, the notions *genericity* and *generics* appear quite often. What is more, some researchers differentiate between the two. Nickel claims for instance that

> generics (linguistic phenomenon) exhibit genericity (not obviously a linguistic phenomenon), and though a theory of generics is closely connected to a theory of genericity, the two are distinct (Nickel 2017: 437).

Even though the terms describe two very closely related phenomena, they do have slightly different interpretations. These might not be visible on the lexical level in other languages

---

[1]The researchers from this group are also known for developing Carlson's division into generics and kind-reference. According to some researchers, such division is present in certain languages (for Norwegian see e.g. Halmøy 2016), whereas in other works such semantic division is absent. In this dissertation I treat kind-reference as generics.

but the semantic difference between the two is universal.

As Nickel postulates, the difference between the terms is purely conceptual. First of the terms, namely generics, is related to natural languages in a strictly linguistic sense, not in an abstract one. It simply designates the ability to create and utter generic sentences (Nickel 2017: 437). Generics could therefore be considered as *actual realisations* of genericity in the same way as utterances are actual realisations of abstract units – sentences. The idea of features that can only be assigned to kinds, is the one connected to the conceptual level of the notion. At that point, the speaker assesses a given feature to a given kind – a feature that cannot be connected with an individual – in order to utter it afterwards with the use of language. Such features, when expressed in a form of generic sentences, become then generics in a given language. For instance sentences such as the classic example in (12), contain the so-called KIND-RESTRICTED PREDICATES (called also KIND-PREDICATES; in this case 'to be extinct'):

(12)   Dodos are extinct.

Those are predicates that can only apply to whole kinds, not to individuals of a particular kind (Nickel 2017: 437). One cannot for instance say that *a lion is extinct* or that *a sparrow is widespread* (Lyons 1977: 196). Sentences containing kind-restricted predicates are considered to be classic examples of generics – they do not allow for any other interpretations than generic ones. They are also limiting the choice of NP types that can be used in such utterances, as shown in the example above.

Nickel also points out that even though some generics are indeed very clear when it comes to their interpretations, sentences of this kind are still highly context bound (Nickel 2017: 438). This applies particularly to sentences which do not contain kind-restricted predicates, for instance:

(13)   a. Ravens are black.

        b. Dogs bark.

        c. Cows have four legs.

This context-dependence does not allow for any statistical approach to the issue as the variability of the phenomenon is too great and is in addition connected to truth-conditions. Every person has some knowledge about the world and can therefore judge whether a given sentence is probable (true) or not. Generics depend greatly on truth-conditions and plausibility of a given feature occurring in real life. Examples (14a) and (14b) show this dependency:

(14)   a. Ravens are black.

        b. Ravens are white.

Most speakers, referring to their general knowledge about the world, would probably not consider the second sentence to be true and neither generic. The reason for this is that

generics that do not contain kind-restricted predicates imply that a generalisation or a feature from that sentence concerns most individuals from a given group. Most but of course not necessarily all – in some very particular cases one could utter (14b) as a true sentence, for instance when talking about albino ravens (Nickel 2017: 438). The problem with such an interpretation though is that the probability of an albino raven occurring in the real world is very low and most speakers would not allow for a generic sentence about such a limited group.[2] An albino raven would still be classified as a raven but uttering a generalisation about an exception to the rule does not seem to be acceptable by most speakers. If it was, other examples of this kind would have to be considered true, for instance:

(15) a. Cows have three legs.

  b. Frenchmen do not drink wine.

Both examples could be true in some particular contexts but they cannot be considered generic in the classic meaning of this term as mentioned above. In order to create a generalisation about a kind or a group of objects or people, one needs to refer to a quality that the majority of this kind/group would possess. What is more, this type of generics, namely generalisations without kind-restricted predicates, is not prone to any statistical analysis and cannot therefore be evaluated this way (Nickel 2017: 439).

As I will explain it in the section on two senses of genericity, there are two main types of sentences that can have generic meaning. One of them are kind-referring sentences as discussed above, whereas others are a lot more context-dependent and therefore connected to the conceptual level of the notion (HABITUAL SENTENCES, called also CHARACTERISING SENTENCES in the literature on the matter (Krifka et al. 1995: 3)). The first type might be therefore connected more to the notion *generics*, whereas the latter to a slightly broader notion of *genericity*.

> Since the meaning and use of generics is probably the most direct evidential connection we have to genericity, it's only natural to frame a theory of genericity as a theory of the truth-conditions of generics (Nickel 2017: 441).

The distinction of generics and genericity, as important as it is, does not seem to pose difficulties in the analysis of the phenomena in a language. Both terms are used interchangeably in many linguistic works. Some researchers opt for one of the notions, others for the other whereas some do not differentiate between them at all and use them interchangeably. Also the most comprehensive work on the matter, namely *The Generic Book*, contains both names in seemingly similar contexts (see e.g. Krifka et al. 1995: 1-124 and Dahl 1985: 412-415). Generics are rather connected to such notions as *marking of generics in language*. This can indicate that the authors do respect the distinction without

---

[2]See also the so-called *false generalisations* in section 2.4.3.

expressing it explicitly in the theoretical section of the book.

In this study the term *genericity* will be applied when it comes to the linguistic phenomenon, whereas *generics* will be applied to actual generic terms, sentences and texts – the *realisations* of genericity.

2.2 Two senses of genericity

In the literature on the matter, it is often mentioned that genericity has two senses. This terminology, similarly as with other notions connected to the subject and presented in this chapter, was first established by Krifka et al. (1995). In recent research, one turns rather to newer and often interdisciplinary approaches, where the division of generic sentences is slightly different. In order to understand the evolution of the methodology on genericity, I shall first focus on the *classic* approach, namely the division into KIND-REFERRING SENTENCES and HABITUAL SENTENCES (Krifka et al. 1995: 2-4).

Kind-referring sentences, as the name suggests, refer to kinds. The phenomenon can be seen in the examples below:

(16)   a.   Beavers build dams.

b.   Foxes are mammals.

Other predicates utilised often in kind-referring sentences are called 'kind predicates' (Krifka et al. 1995: 10; known also as 'kind-restricted predicates' as has been mentioned before). Those are constructions such as *be extinct*, *be rare*, *be numerous* etc. Kind-predicates, again as the name suggests, may be used only with NPs that designate kinds and are most often bare plurals. Indefinite and definite singular nouns are unacceptable and somewhat awkward as shown in the sentences in (17) below.

(17)   a.   Pandas are rare.

b.   Lions are numerous in Africa.

c.   ? A panda is extinct.

d.   ? The lion is rare.

The two forms above identified as unacceptable, can occur in generic contexts. Hawkins perceives such sentences as indefinite and definite generics (Hawkins 2015: 214) and gives the following examples:

(18)   a.   A lion is a noble beast.

b.   The lion is a noble beast.

c.   Lions are noble beasts.

d.   The lions are noble beasts.

The four sentences presented above differ from the examples in (17) in that they do not include kind-referring predicates. Therefore, it is fully acceptable to form generic sentences with indefinite and definite singular forms without them sounding awkward or unnatural. What is more, Hawkins' examples show that both singular and plural forms can be used in generic sentences – in singular and in plural. This view differs slightly from other theories in which plural definite form is completely excluded or, as we shall see in section 2.4.1, reserved for humans.

It is also worth mentioning that technically all NP types have the potential to take on a generic reading but not all forms use this potential, not in all contexts and, above all, not in all languages. As Behrens states it, in most languages all *generic* NP types[3] are ambiguous as they always have another possible reading (Behrens 2005: 276). What is more, *it is never the case that all different generic types permitted in a language would be intersubstitutable in all possible (generic) contexts* (Behrens 2005: 276). One can therefore divide reference types into categories, where generic reference also has some subtypes (see section 2.4 for further discussion).

The nature of definite and indefinite reference allows for both generic and specific uses of each of the NP types.

> The indefinite article still refers exclusively, the definite article still refers inclusively within pragmatically defined parameters. It is therefore no accident that one and the same morpheme can perform both functions. Again, locatability and grammaticality facts prove the fundamental similarity between generic and non-generic reference. (Hawkins 2015: 214)

Apart from the fact that the same morpheme can take on more that one function, the context remains the decisive factor in interpretation of a given sentence. This is particularly visible in the case of indefinite reference. If indefinite reference is used specifically, it is only the speaker who knows which entities of a group the sentence concerns, whereas in the case of generic indefinite reference it is both the speaker and the hearer who are able to identify the referents (Hawkins 2015: 215).

Whether the speaker uses definite or indefinite form, the interpretation of such reference has to be verified by the context. Reference can be specific when the speaker means a given entity or a group of entities, or generic when the entities picked out in the sentence represent a feature or features of a bigger group. Hawkins calls generic references *non-specific* references in *specific* contexts (Hawkins 2015: 215). Those *specific* contexts can be an effect either of the speaking situation when, for instance, the speaker is able to point at something and refer to it specifically, of the predicates used in the sentence (for example kind-predicates) or a wider context in which a given sentence appears. It is

---

[3]In none of the studied languages there seems to be a noun form reserved only for generic references (Behrens 2005: 277). Therefore stating that a noun form is generic means that it has such a reading in a given context, not by default.

therefore possible that a sentence analysed out of the context would have a specific reading, whereas in a wider context it might be interpreted as generic (Behrens 2005: 279).

A similar ambiguity in interpreting the sentences depending on their form, can be observed in habitual sentences as well. This is particularly visible in languages that have formal distinction when it comes to functions of particular noun phrases. Behrens refers to French, where the reading of habitual sentences depends highly on the context, as both in habitual and non-habitual sentences it is possible to use definite forms which are also applied in generic constructions (Behrens 2005: 285):

(19) a. Jeanne mange *les pommes.*
          Jeanne eats   the apples.
          Jeanne eats the apples.

      b. Jeanne mange *des pommes.*
          Jeanne eats   Ø-apples.
          Jeanne eats apples.

Habitual sentences, called also characterising sentences, are said to express the second sense of genericity (cf. Carlson and Pelletier 1995), namely regularities and characteristics of someone or something. Characterising sentences do not need to include generic noun phrases or even make generalisations about kinds to be considered generic[4]. As Behrens claims, *habitual sentences (. . .) resemble traditional generic sentences in that they express a typical characteristic of their subjects* (Behrens 2005: 288). Let us consider the following examples:

(20) a. Yesterday, we had a very interesting discussion *about the potato.* (The teacher
          told us that it was first cultivated in South America. . .)

      b. John smokes a cigar after dinner.

      c. I love *beavers.*

      d. *The beaver* has always fascinated me. (Behrens 2005: 289, original emphasis)

Sentence (20b) is considered to be a *classic* characterizing sentence – it states a characteristic feature about John, the subject of the sentence (cf. Behrens's definition above). Sentence (20c) could also be interpreted as habitual since it states a generalisation about the speaker. Some researchers do not consider such generalisations generic as they do not concern kinds but rather individuals (Behrens 2005: 288). On the contrary, sentences (20a) and (20d) contain generic noun phrases but they do not make generalisations about kinds either. Nevertheless, they are considered generic habitual sentences. This type of characterising sentences will be analysed as generic also in this project.

The interpretation of habituals as generic or non-generic depends on the notion of genericity that one applies. Behrens proposes therefore the division between *generic*

---

[4]Some researchers do not consider habitual sentences generic. There is no agreement among the linguists when it comes to the interpretation of such sentences but the analyses where both types of generics are considered are abundant. In this dissertation I will therefore treat habituals as generics.

sentences and *generic phrases* (Behrens 2005: 289) which will be discussed in greater detail in section 2.2.2.

2.2.1 Generic NPs

As has been stated before, genericity can be expressed both at the NP-level, as well as at the sentence level. What is more, generic NPs do not need to occur in generic sentences and the other way around – generic sentences do not need to contain generic NPs to be interpreted as generic (Behrens 2005: 288). In English, generic NPs can, technically, occur in all possible forms (definite and indefinite, singular and plural) but it will always be the context and the reading of a given sentence that will play a central role in its interpretation.

As we have mentioned before, certain NP types tend to have a generic reading in most cases (for instance BPs), whereas others will depend on the context in which a sentence is uttered or on the form of the subject of such a sentence (e.g. definite plural generic). What is more, in the case of habitual sentences, the NP type does not seem to play the central role, whereas in *classic* generic sentences (kind-referring) the NPs' form and its role seem to influence the reading of a sentence.

On this basis, one can differentiate between two main analysis methods. One way of analysing generic sentences is the *bottom-up* strategy, where linguists look rather at the forms of NPs, definiteness and their function in the sentence. The other strategy, the *top-down* approach, focuses on generic sentences and their structure, rather than on the functions of each element of such a sentence (Behrens 2005: 287).

In the analysis presented in further chapters of this dissertation, I will employ both strategies, as the analysed material is rather complex. In order to properly analyse both the surveys and the corpus texts, it is necessary to look at both levels of genericity – the NP-level and the sentence level, depending on the source materials.

Generic NPs can occur in different forms, definite or indefinite:

> One obvious difference between definite and indefinite noun-phrases, used generically, is that, with definite noun-phrases, both a collective and a distributive interpretation is possible, but with indefinite noun-phrases (in the singular) the collective interpretation is excluded. (Lyons 1977: 196)

According to Lyons, English native speakers are rather consistent when it comes to the interpretation of NPs. For instance, bare plurals will most of the time be considered generic, whereas definite plural nouns will, for most English speakers, designate a specific group of entities.

Languages can show different levels of grammaticalisation when it comes to generic NPs and their form. For instance, in English the most used NP types in generic sentences are bare plurals, singular definites and singular indefinites, with very few exceptions. The

exceptions allow for bigger liberty in the choice of NP types in such utterances and are, of course, context-dependent. Such a wide variety of choices might not be the case in every language though.

French, on the other hand, is an example of a language with fully grammaticalised articles where the singular bare noun is considered incorrect in generic contexts, even when it comes to mass nouns that tend to occur as bare nouns for example in English and Norwegian. What is more, the French indefinite singular article seems to be the only indefinite form accepted in a generic context (Galmiche 1985: 2).

The most often used forms of a generic NP are the definite ones, singular and plural, with a very limited use of indefinite forms:

(21) a. Le lynx a   une vue   perçante.
          The lynx has a   vision sharp.

          The lynx has a sharp eye.

      b. Les lynx   ont une vue   perçante.
          The lynxes have a   vision sharp.

          The lynxes have a sharp eye.

      c. Un lynx a une vue perçante.
          A lynx has a vision sharp.

          A lynx has a sharp eye. (Galmiche 1985: 2)

The rule seems to be consistent and applicable to most (if not all) generic readings. Generic interpretations are of course context dependent and prone to numerous readings, also non-generic ones. Despite that, there are some exceptions where indefinite form would indicate a generic reading without the possibility for different interpretations:

(22) Un lapin   albinos est en voie d'extinction (l'albinos   des   Pyrénées).
      A   rabbit albino is on way to extinction (the albino of the Pyrenees).

      The albino rabbit is about to extinct (the albino of the Pyrenees). (Galmiche
      1985: 9)

An albino rabbit in the sentence above is understood generically as it is a subspecies. Therefore, one can imagine saying that "a rabbit species" (*une espèce de lapin*), namely the albino of the Pyrenees, is about to extinct (Galmiche 1985: 9). It is worth mentioning that in English one would rather opt for a definite form when talking about subspecies (Mari et al. 2013b: 77) but both indefinite and definite forms are substitutable in such contexts (cf. Lyons (1977: 196)).

2.2.1.1   Cognitive status of familiar nouns and well-established kinds

A particular type of generic NPs are the so-called WELL-ESTABLISHED KINDS (WEKs). Another category that is closely connected to WEKs are FAMILIAR NOUNS. Even though

some familiar nouns may be WEKs, the two notions are sometimes used synonymously. However, their cognitive status is slightly different, as will be shown in this section.

The interpretation of WEKs and familiar nouns is very much dependent on one's language intuition, the context in which the noun appears, as well as the predicate in a given sentence (cf. Pettersson 1976, Carlsson 2012 among others). Thus, the notion of a well-established kind is not a well-established concept itself. Such NPs are regarded as familiar to the speakers of a given language, which can be both the result of one's general knowledge or the cultural context one lives in (different cultures might have different WEKs and familiar nouns). As Borthen (2007: 156) puts it:

> [t]he notion of being 'well-established' is a bit vague in the existing literature, but it seems reasonable to assume that *well-established* means *familiar* in the sense of GHZ (1993) [Gundel et al. 1993].

The publication mentioned by Borthen refers to the so-called Givenness Hierarchy (Borthen 2007: 144). In short, the classification allows to identify English NPs according to their grammatical form (my emphasis):

in focus > activated > **familiar** > uniquely identifiable > referential > type identifiable

Each of the statuses presented in the model above receive a certain cognitive interpretation, e.g. familiarity is described as the ability to recall a representation of an object from one's memory, based on the discourse (Borthen 2007: 144). Familiarity is perceived by Borthen on a par with being well-established. Both notions are connected to speakers' knowledge about the world and the ability to interpret the discourse.

In English, familiar nouns are definite singulars, as in the examples in (23) (Borthen 2007: 148):

(23) a. <u>The coke bottle</u> has a narrow neck.

b. <u>The blue whale</u> is the largest animal on earth.

c. Bell invented <u>the telephone</u>.

'The coke bottle' and 'the blue whale' are examples of noun phrases that function as kinds, whereas 'the phone' can be understood both as a concept of the device, as well as a specific reference depending on the context. In Norwegian such nouns are definite (singular or plural) but they can also occur as bare nouns.

2.2.2   Generic sentences and generic terms

Now that the two senses of genericity and the status of generic NPs have been established, let us take a closer look at generic sentences (generic phrases) and generic terms. Both

notions appear in literature on the matter but they differ slightly in meaning.

The opposition between GENERIC SENTENCES and GENERIC PHRASES was coined by Behrens in her work on typological aspects of genericity. The two notions are applied in a slightly different manner so that it opens for a broader and more precise analysis of generics, without perceiving characterising sentences as non-generic (Behrens 2005: 288-289).

The argument often used against analysing habituals as generics is that the characteristics expressed in such sentences often concern individuals, rather than whole kinds. This way, characterising sentences cannot be analysed as generic as they do not state anything about kinds but only about some, or even one member of a given group. This is called by Behrens for a *narrow* broad sense of genericity (Behrens 2005: 289). In order to account also for those sentences, she then proposes a *broad* approach in which the notion *generic phrase* is used. Generic phrases are not bound by any syntactic restrictions (Behrens 2005: 289) and could therefore be used to all instances that are not considered *classic* generic phrases, namely kind-referring phrases.

Generic phrases, according to Behrens, can, but do not have to, occur with kind predicates such as *be extinct* and *be rare*. They can also occur with predicates that presumably are not generic but in certain contexts do acquire generic reading, as in the examples below:

(24)   a.   Polar bears *are mammals.*   [kind predicate]

       b.   Polar bears *are dangerous.*   [non-generic predicate]

The fact of being dangerous can be used both in specific and generic reference. In specific reference it indicates that somebody or something is dangerous at the moment of speaking about it [non-generic predicate]. Used with a kind-referring expression, the predicate, still being non-generic, expresses a generalisation over most or all members of the kind, as in the example (24b). On the contrary, being a mammal is not a feature that can be acquired for a certain period of time or that can apply only to some entities of a given kind.

Generic sentences and phrases discussed by Behrens (2005) are not new notions in research on genericity. Since the very beginning of such research, the linguists have brought about a number of notions with quite a similar meaning. Carlson (1982) for instance refers to GENERIC TERMS and GENERIC SENTENCES (Carlson 1982: 145). Generic sentences are perceived by Carlson similarly as by Behrens, namely those are *classic* kind references. Carlson (1982: 145) does seem to differentiate between kind-referring sentences and habitual sentences, proposing various notions used in philosophy and linguistics to describe such sentences. Nevertheless, the examples proposed by Carlson are rather generic sentences in the classic meaning of the term, whereas habituals are perceived as a separate category.

What is more, the notion *generic terms* coined by Carlson, seems to be synonymic with generic NPs and kind-denoting NPs. As examples of generic terms Carlson gives *dogs, mountains, unfriendly policemen* (Carlson 1982: 146). He then discusses whether it is plausible or not to say that generic terms denote kinds or, as he quotes after Twardowski, *genera* (Carlson 1982: 146).

In the literature on the matter and in the examples reviewed in this chapter one finds that generic terms cannot be perceived as notions exclusively denoting kinds. Such a statement is rather radical. Carlson and other researchers have pointed out that it would imply that for instance the term *dogs* holds for all dogs in the world, not allowing any exceptions (Carlson 1982: 146). This, as we have seen in (13c), is a false approach. Moreover, perceiving generic terms as solely denoting kinds would mean that there are categories devoted to generics only. In none of the studied languages such a category has been found (cf. Behrens 2005).

Carlson (1982: 148) points out that even though kind predicates can only be attributed to kinds and not members of a given kind, does not mean that generic terms denote. However, predicates do not divide in only two categories (kind denoting and *regular*). They do express rather complex semantic relations, such as the ones in the examples below:

(25)   a. Dogs are mammals. (*all*)

       b. Dogs eat meat. (*most*)

       c. Dogs give milk to their young. (weaker than *most*) (Carlson 1982: 148)

Examples such as those in (25) will be discussed further in the following sections of this chapter. What is worth mentioning though is that throughout this dissertation the notions of *generic terms, kind-denoting NPs* and *generic NPs* will be treated as synonyms. As has been stated before, after the publishing of *The Generic Book* a number of studies have been conducted. However, many of the researchers still utilise different notions to refer to the same concept. While quoting or making references to these studies I will follow the original terminology (as above in the description of Carlson's discussion on generic terms and sentences) but the preferred notions used in this research are kind-denoting NPs and generic NPs.

Generic terms are also said to function similarly as proper names, especially in the *so-called* constructions (Carlson 1982: 150). The examples in (27) sound awkward as they do not contain proper names, whereas the sentences in (26) are correct (Carlson 1982: 150). The same concerns generic terms - in certain sentences non-generic terms would sound awkward, just like the examples below without proper names do.

(26)   a. Giorgione is so-called because of his size.

       b. Cardinals are so-called because of their color.

       c. Machine guns are so-called because they fire automatically.

(27) a. The man over there is so-called because he is over there.

   b. Most machine guns are so-called because they fire automatically.

   c. All devilfish are so-called because of their satanic appearance.

In his analysis Carlson discusses also kind-denoting predicates, mentioning that they do not need to occur in sentences with natural kinds or nominal kinds (Carlson 1982: 153). On the contrary, they are restricted to kinds and cannot denote members of kinds. Further discussion on Carlson's approach, a somehow *classic* approach, will be presented in section sec:gen-formal.

The notions *generic sentences, generic phrases* and *generic terms* are closely related and their application in generic constructions are also similar. Generic sentences, as stated both by Behrens and Carlson, are sentences that express generalisations – either about kinds (*classic* generics) or about individuals (habitual sentences, also called characterising sentences). Generic terms are a notion utilised by Carlson and it concerns noun phrases that denote kinds, such as *dogs, cats*. Generic phrases, on the contrary, is a term coined and used by Behrens. It concerns the phrases used in habitual sentences, which are also considered generic.

2.2.3 Generic texts

In analyses on genericity one usually takes into account two linguistic levels, namely NP-level and sentence-level discussed above. Behrens (2005) proposes a new level of analysis, namely generic texts.

Generic texts present generalised knowledge about kind or particular stereotype situation (Behrens 2005: 289). In generic texts one finds a larger number of generic noun phrases than in non-generic texts but

> [t]his does not imply that a generic text contains only generic sentences of the classical type or that every mention of a linguistics expression allowing reference to the topic of a generic text is in actual fact to be interpreted as a generic NP (Behrens 2005: 290).

The fact that generics are present in a given text, e.g. in a novel, does not mean that the whole text is therefore generic. Since generic texts state generalisations over kinds or refer to stereotypical situations, it is rather non-fiction texts that contain numerous generic sentences. Nevertheless, this is not the rule. Behrens (2005) herself has analysed a fiction text, namely the text of *Le Petit Prince* in French, English, German, Hungarian and Greek. This proves well that even literary texts can be considered generic, even though this genre is not usually recognised as generic by its nature.

A number of Swedish generic texts were analysed in Carlsson (2012). The texts utilised for the analysis on genericity in Swedish were encyclopaedia texts about natural

kinds. The analysis of this dissertation also focuses on non-fiction generic texts which will be discussed in greater detail in chapter 4.

In her study on genericity Behrens has proposed two approaches, namely heuristic and interpretative ambiguity tests. The first approach is the one often used in the analysis of individual sentences, as well as in corpus-based studies. The researchers base their judgements on their own linguistic intuition (Behrens 2005). This approach is quite common and can be found in most works on genericity, for instance in Carlson's and Chierchia's analyses, as well as in *The Generic Book* and the like.

Interpretative approach is a strategy utilised in survey-based studies (Behrens 2005: 291), such as this one. It requires that the researchers consult native speakers of a given language and take into account their judgement of generics. In my dissertation I apply both: heuristic in the case of corpus and interpretative in the case of the surveys. Interpretative ambiguity test is often based on actual utterances, not necessarily sentences created for the sake of the analysis. In this way the respondents' judgements can be analysed as those occurring in a spoken language.

The approach proposed by Behrens is novel in two ways. First of all, the third linguistics level, apart from the NP-level and sentence level, was not considered in previous analyses. Generic texts defined by Behrens allow the linguists study genericity in a wider context than a noun phrase or a sentence. The second innovative aspect is the approach to the analysis itself. One can either analyse the sentences basing on their own language intuition, ask native speakers of a given language for judgement or, as will be applied in this dissertation, combine the two strategies.

2.2.4   Generic anaphora

Genericity can be expressed not only in statements that contain generic NPs or kind-restricted predicates. The phenomenon can occur also in anaphoric expressions, which can also be generic. Already Carlson (1977: 425) mentions generic anaphora in examples such as the sentences in (28):

(28)   a.   Kelly is seeking a unicorn, and Millie is seeking *it*, too.

b.   Kelly is seeking a unicorn, and Millie is seeking *one*, too.

c.   Queenie is seeking *unicorns*, and Phil is seeking *them*, too.

The examples above illustrate that both indefinite singular form and bare plural form can serve as antecedents in anaphoric expressions, also those interpreted generically. In his further discussion Carlson analyses mainly the use of BPs in this function, showing that not all expressions of this kind are interpreted in the same manner. The opposition is particularly visible when it comes to transparent and opaque reading.

The notion of the REFERENTIAL OPACITY was coined by Quine (2013) in his seminal work entitled *Word and Object*. Referential opacity is introduced in the chapter on

'Vagaries of Reference' and, in short, it is presented as a way of interpreting expressions. The reading of a given expression can be either opaque or transparent. If the reading is opaque it implies that the expression is ambiguous to some extent and can be interpreted in more than one way, depending on the speakers intentions and knowledge (Quine 2013: 131).

Opacity is expressed on the semantic level with what a speaker believes/means. An example of that was presented in the sentences in 5. What a speaker wishes to convey in his or her statements may be consistent with their knowledge and/or believes but does not need to be true – Professor Brown might be the Dean and then the two sentences would mean the same thing. On the other hand, Mr Smith looking for Professor Brown might not know that he is not the dean. Such sentences are therefore opaque as they can suggest more than one possible referent.

Referential opacity is present also in generic anaphora, where the pronouns may refer to their generic counterparts but they could also convey transparent readings. The nature of the antecedent determines what kind of reading a statement has. For instance, an antecedent that is a mass noun such as 'furniture' in the 29 favours opaque reading, whereas the antecedent from the 28a does not (Carlson 1977: 425-426).

(29)   Cedrick is seeking furniture, and Hiram is seeking *it*, too.

Even though the anaphoric expression in 29 might be seen as referring to the same piece of furniture that Cedrick is seeking, this is not the case (Carlson 1977: 426). The example in 28c on the other hand refers rather to the activity of searching for unicorns than looking for a particular group of those animals. The BP 'unicorns' is a generic antecedent and the pronoun 'them' has therefore generic reference as well.

What is more, according to Carlson (1977: 432) *a generic may serve felicitously as antecedent for an existential.* The examples of generic anaphora in such contexts are numerous and some of them are provided by Carlson (1977: 433):

(30)   a. My mother hates *raccoons* because *they* stole her sweat corn last summer.

   b. *Raccoons* have stolen my mother's sweet corn every year, so she hates *them* a lot.

(31)   a. My brother thinks *snakes* are nasty creatures, but that hasn't stopped me from having *them* as pets my whole life.

   b. I've had *snakes* as pets my whole life, but my brother still thinks *they*'re nasty creatures.

(32)   a. Martha told me that *beans* don't grow as well in this climate, but *they* grew well for me last year.

   b. *Beans* grew quite well for me last season in spite of Martha's warning that *they* can't grow in this climate.

(33) a. I didn't believe that *goats* liked *tin cans* until I actually saw *them* eating *them* last week.

  b. Before I actually *saw* goats eating tin cans last week, I didn't believe *they* liked *them*. (Carlson 1977: 433)

What all the sentences above have in common is the fact that in each of them the generic noun phrase serves as the antecedent for the anaphoric expressions. The examples contain indefinite plural forms which, according to Carlson and many other researchers in the field, are most frequently used generic nouns. Further discussion on the NPs' forms in generic contexts will be presented in the following sections on formal approaches to genericity.

2.3   Formal and modal approaches to genericity

The notions provided in previous sections and the classification of generic sentences with two main types (kind-referring and habitual sentences) are parts of the formal paradigm of genericity proposed and developed by the researchers from the Generic Group and their successors. The approach proposed by the researchers is very complex and an exhaustive description of these theories lies outside the scope of this work. Therefore I will characterise briefly the main assumptions of the formal and modal theories, since this paradigm has served as the foundation for all modern accounts of genericity – even those rejecting the logical analysis.

The Generic Book published in 1995 was the first compilation of the existing research on genericity. The work standardised and unified the terminology that had been in use from the 70s. What is more, the work of (Carlson and Pelletier 1995) was a milestone in the research on the matter as it also developed new theories and approaches while taking into account older assumptions. This section consists of two main parts in which I will discuss the so-called CARLSONIAN APPROACHES to genericity (started by Carlson in his numerous works on the matter and continued later by others) and the NEO-CARLSONIAN THEORIES (introduced by Chierchia and the like) that arose in the years after *The Generic Book* was first published.

Both Carlsonian and neo-Carlsonian approaches are based on the Logical Form (LF) and are considered to be of a rather formal character. The theories are also said to be modal, as they often propose models for analysis of generic utterances which are perceived to be of a quantificational character (Lazaridou-Chatzigoga et al. 2015: 470).

Since there are numerous models and operators connected to the semantics of genericity, I shall only focus on the most acknowledged ones and those that are considered somehow *classic* in the literature on the matter. The main focus of this thesis is the cognitive approach to genericity in language and I will not apply the LF-analysis in my research. Yet, the achievements of Carlson, Chierchia and the like are crucial in understanding the phenomenon.

2.3.1 Carlsonian theories

Many of the theories on genericity are based on the work of Carlson, starting already from the year 1977 when the article entitled *A unified analysis of the English bare plural* was published. Later, Carlson took up the issues of temporal expressions and genericity (Carlson 1979), generic terms and sentences (Carlson 1982), as well as the logical form of generic utterances (Carlson 1983). All these papers and theories have contributed to the major study on genericity, namely the publication of *the Generic Book*.

Carlson's take on genericity relies greatly on the phonologically null generic operators such as *Gn* (Carlson 1977, 1982), which has later developed into the widely used and acknowledged GEN-operator (see e.g. Carlson and Pelletier 1995; Mari et al. 2013b). In short, Gn-operator, as well as GEN, can be assigned to a given predicate in order to obtain a generic reading, as shown in the examples below (Krifka et al. 1995: 22):

(34)   a. John smokes.   **Gn(smoke)(John)**

           b. Italians smoke.   **Gn(smoke)(Italians)**

           c. Italians know French.   **Gn(know.French)(Italians)**

In his seminal work of 1977, Carlson analyses mainly the generics of English bare plurals but a short description of other NP types in such contexts is also given. For instance, a hypothesis is made that indefinite singular form with the article *a/an* might, in certain contexts, function as a counterpart of bare plurals. In order to support for this, Carlson gives examples of generic sentences where both noun forms can be used and interpreted in a similar manner (Carlson 1977: 415):

(35)   a. *A mammal* bears live young.

           b. *Mammals* bear live young.

In the following parts of his work, he then explains that such a hypothesis might be neglected and that BPs *cannot be plural of 'a'* (Carlson 1977: 429). The assumption was later discussed by numerous linguists and agreed upon, for instance by Mari et al. (2013b: 25):

> (...) kind-referring definite singulars in languages like English are **not** trivial variants of kind-referring bare plurals. (my emphasis)

In the analysis of 1977, THE OPACITY PHENOMENA were also mentioned. The notion of referential opacity, as mentioned before, was first introduced by Quine and discussed by numerous linguists afterwards (see for instance Lyons 1977: 192). Carlson analyses *opaque* sentences the meaning of which can be ambiguous as in the example below:

(36)   Minnie wishes to talk with a young psychiatrist. (Carlson 1977: 417)

The opacity of the sentence (36) (THE OPAQUE READING) makes it impossible to interpret

it in one way. A hearer might then presume that Minnie wishes to speak with *any* psychiatrist, as long as it is a young psychiatrist. When the so-called TRANSPARENT READING of the sentence (36) is taken into account though, Minnie might in fact have in mind a particular psychiatrist she wishes to speak to (Carlson 1977: 417).[5]

The sentence in (36) can, according to Carlson, have at least two different semantic structures:

(37)   a. (Epl x) (young psych. (x) & M. wishes M. talk with x)

        b. M. wishes (Epl. (x) (young psych. (x) & M. talk with x)

where (37a) represents the transparent reading and (37b) shows the opaque interpretation (Carlson 1977: 417). What is interesting, is that the use of indefinite article *a* in (36) does not in fact exclude a narrow-scope meaning where Minnie has in mind one particular psychiatrist (Carlson 1977: 417).

One of the most important claims made by Carlson is the one that there is a semantic difference between individuals and the so-called stages of individuals. This hypothesis was formulated as follows:

> The stages aren't simply the things that *are*; they are more akin to things that *happen*. That is, stages are conceived as being much more closely related to events than to objects. (Carlson 1977: 448)

'Stages' are closely connected and perhaps even synonymous to 'properties' (Carlson 1977: 448-449), as they both characterise an individual but are not a part of it. These can be observed in the examples in (38), where different stages (properties) of Jack are being listed.

(38)   a. Jack is intelligent.

        b. Jack is tall.

        c. Jack likes wine.

As can be seen, the fact of being intelligent, tall or liking wine does not mean that without those features Jack would not be Jack. Those are therefore only his stages which can last for a period of time, be applicable at any time or even be replaced by other stages.

Such an approach to analysing generic sentences might suggest that genericity can in fact occur both at the NP-level (individual), as well as at the sentence-level (stages of individual). Since stages relate to events, one might suppose that these will appear in habitual sentences, rather than in kind-referring sentences where generic NPs denote kinds seen as individuals in Carlson's analysis.

In his later work, Carlson has further modified his theory and focused also on the atemporal clauses with *when*, such as in the examples below:

---

[5]The case of opaque and transparent readings was already presented in the examples about Professor Brown and the Dean in (5).

(39)   a. Giraffes are intelligent when they are tall.

        b. Cats are intelligent when they have blue eyes. (Carlson 1979: 52)

Even though the sentences in 39 may sound awkward to speakers of English, the use of *when* does not exclude their generic reading. Carlson (1979) discusses the existence of what he calls *a generic verb phrase operator* which seem to exclude any temporal references from sentences such as (39). A number of restrictions are given in order to support the atemporal when-hypothesis and analysis. The restrictions concern the structure of the main clause, use of NPs and the generic interpretation (Carlson 1979: 51-53).

An interesting claim made by Carlson in his work of 1979 is that he sees kinds *not as being sets of objects, as is commonly supposed, but rather as being individuals themselves* (Carlson 1979: 54). What is more, another category of entities was proposed, namely STAGES, which, according to Carlson, are subtypes of entities and *time-space slices of individuals* (Carlson 1979: 54). Such an approach was novel in many ways, mostly because most researchers perceived, and still do perceive, kinds as collections of individuals.

For each of the claims made, Carlson proposes semantic structures that illustrate the theory. Apart from the LF, a number of hypotheses are made and confirmed. One of them being a claim that most predicates apply to stages and to kinds themselves (Carlson 1979: 56). The main focus in the paper falls on the generic operators which influence the reading of a given sentence. If the operator is present, a generic reading is preferred, whereas the lack of the generic operator GN in a sentence suggests a specific or *an event-type* reading.

The theory of Gn-operator accounts for ambiguities of sentences which can arise when the generic reading is unavailable in a given context. In this sense, a generic reading seems to be more straightforward than a non-generic one. This dependency is best seen in the semantic structures proposed by Carlson, where the presence or absence of the Gn-operator lets one interpret a given sentence in a generic or non-generic way. Interestingly enough, cognitive linguists also seem to be searching for the difference between these two readings, only using different tools. Instead of marking of genericity at the semantic level, cognitive approaches focus on the speakers' language intuitions (see section 2.4).

In yet another of his works, the article of 1982, Carlson delves further into the matter of generic sentences. This time it is not only the semantic structure of an NP or a DP that is taken into account, but a much wider scope in which generics are being analysed as a sentence-level phenomenon. Two main issues of this study are: (1) the interpretation of generic terms and (2) truth-conditions of generic sentences (Carlson 1982: 145). Such an approach gives a more ontological character to the study, which is also one of the main features of most formal theories of generics.

The truth value of a given generic sentence is certainly important when interpreting it, but it does not exclude generic meaning. This means that if a generic claim is made, such as that in (1) that assigns a feature of barking to all or most dogs, the claim can

still be considered generic even if it does not hold for each and every dog in the world. Carlson himself points out that one can easily find an exception to the rule that, for instance, all dogs bark, and yet most people would still be willing to ascribe the feature *dog* to a non-barking individual of the kind *canis lupus familiaris* (Carlson 1982: 147).

Another point that Carlson makes is that some predicates cannot occur in a sentence with individual objects but that they can only refer to whole kinds. These were mentioned before as kind-referring predicates – predicates that can only describe the whole species. In contrast to those, another type of predicates is discussed, namely the one that refers to less than 50% of a kind, namely the minority, and is yet used generically. An example of such a predicate can be *to give milk* as in the example below:

(40)   Dogs give milk to their young. (Carlson 1982: 148)

Even though it is only female dogs at a certain age and in certain situations, the sentence 40 would be interpreted as generic by most speakers. A similar issue was discussed by Leslie et al. (2011) in their study on Generic Overgeneralisation effect (GOG, see section 2.4.3), where a sentence *Ducks lay eggs* was given as an example of GOG.

Another issue discussed by Carlson are KIND-DENOTING TERMS and ENTITY-DENOTING TERMS. All predicates that can apply to entity-denoting terms, will also apply to kind-denoting terms as each entity is a part of a kind (Carlson 1982: 158).

The ambiguity concerning use of different tempora in generic sentences is presented in the examples below:

(41)   a. Dogs barked.
       b. Dogs bark.
       c. Dogs will bark. (Carlson 1982: 165)

Each of the sentences above has a generic reading and it is only the situation in which they could be uttered that accounts for their meaning. For instance, in (41a) the use of past tense might suggest either an episodic reading or a generic reading if the species was extinct. Sentence (41b) is the *classic* generic example, whereas (41c) could have been uttered before the kind *dog* existed and in the case that somebody was able to predict the future and state that dogs will be able to bark (highly improbable truth-conditions but a possible generic reading).

All sentences in (41) are habitual sentences – the type that by some linguists are considered generic, and these can be, according to Carlson, analysed in a similar matter as kind-referring sentences:

> (...) the analysis of generics should, if possible, be assimilated to the analysis of habitual sentences (to the point where we treat generics and habituals as subclasses of a more general category we will call *gnomic* sentences.) (Carlson 1982: 165)

Even though Carlson's claims from his early works seem to be well-grounded, the theory has evolved and Carlson himself introduced minor changes and new aspects. In the article from 1983 he focused on the LF of natural kinds and the so-called NOMINAL KINDS. The first type was described as either subsets or supersets, where all entities share exactly same features. Nominal kinds on the other hand, are seen as group of entities from which the majority share a given feature (Carlson 1983: 300). Predicates that refer to natural kinds must therefore be true for all individuals of that group, for instance a predicate *to be a mammal* will be true for all cats, whereas a predicate *to have four legs* might refer to a majority of that kind but not necessarily to each and every individual.

Theories including generic operators were further developed and modified by other researchers from the Generic Group. The work of Chierchia, which will be discussed in greater detail in the next section, has contributed to the development of the neo-Carlsonian theories. This proves that Carlson's original claim, even though slightly modified and proved insufficient for a comprehensive analysis, has remained an important element of study on generics. As Mari et al. (2013b: 8) sum it up:

> Carlson's analysis is very elegant: it presents a unified analysis of English bare plurals and predicts the correct existential and generic readings.

Carlson's approach to genericity can indeed be seen as *elegant* but recent research has shown that the formal analysis might not be sufficient when it comes to interpretation of generic sentences in different contexts. Since formal and modal approaches focus mainly on sentences constructed for purposes of the analyses, they do not take into consideration a wider context of analysed utterances, nor do they account for different possible readings that depend highly on the context (Lazaridou-Chatzigoga et al. 2015: 471). This does not mean however that modern approaches to genericity can neglect the achievements of formal linguists. It is in fact quite the opposite:

> [...] the literature on the processing and acquisition of genericity has often ignored or misrepresented the relevant linguistic analyses and stands to benefit from the wealth of insights and the systematicity found in the theoretical linguistics literature. (Lazaridou-Chatzigoga et al. 2015: 471).

The theories presented in the section above are much more complex and this short description was certainly not meant to exhaust the subject. For a more detailed discussion of the Carlsonian approaches to genericity see for instance Carlson and Pelletier (1995); Carlson (2010); Pelletier (2010); Mari et al. (2013a); Lazaridou-Chatzigoga et al. (2015) among others.

2.3.2   Neo-Carlsonian theories

The importance of Carlson's theories and approaches has motivated many researchers to develop and better adjust the claims so that new analyses and models could be introduced. One of the linguists who has contributed by far the most to the development of the so-called neo-Carlsonian theories is Chierchia. Since the publishing of *The Generic Book*, Chierchia continued his research, which in 1998 led to the publishing of his seminal work on genericity, namely *Reference to kinds across languages* (1998).

One of the central points in Chierchia's analysis is its cross-linguistic perspective which was missing in Carlson's theoretical studies on genericity. Chierchia's interest falls on the role and functions of generic NPs in languages such as Chinese, English, Italian, as well as other Germanic and Slavic languages. Apart from this multilingual perspective, Chierchia follows Carlson in some ways and disagrees with him in others. Mari et al. (2013b: 9) have meticulously analysed the differences between Carlsonian and neo-Carlsonian approaches, summarising them as follows:

1. First, Chierchia doesn't use the notion of stage of an individual, which played a crucial role in Carlson's analysis, since it gives rise to a distinction between two types of predicates, i.e. inidividual-level predicates on the one hand and stage-level predicates on the other.
2. Secondly, Chierchia analyses the semantic contribution of plurality in English bare plurals. Indeed, contra Carlson, who compares English bare plurals with proper names (they are analysed as constants at the logical form and are bare, i.e. built without any determiner), Chierchia proposes a compositional analysis of English bare plurals, in which the semantic import of the plural morpheme is analysed. Bare plurals result from the composition of a plural morpheme with a singular predicate to form a plural predicate, which is nominalised.
3. Finally, Chierchia proposes a formal and compositional analysis of kind-referring DPs. He addresses the ontological issues related to the structure of the domain of reference of discourse entities, and he make explicit the relations between singular individuals, plural individuals, kinds and properties. He intriduces new operators, the up and down operators, which allow for an account of the relations between individuals and properties.

As presented above, there are quite a few differences in Chierchia's approach to genericity and Carlson's one. I shall briefly analyse neo-Carlsonian theories started by Chierchia in 1998, focusing on the main differences between the two contrasting views, as they were pointed out by Mari et al. (2013b).

The main element of Chierchia's study is the so-called NOMINAL MAPPING PARAMETER. Put simply, the theory of Nominal Mapping Parameter concerns the features of an NP that allow it to denote certain objects or groups of objects. When a given NP denotes an individual, *clusterings* of these denoted objects are created in the semantic structure (Chierchia 1998: 358). As Chierchia points out, depending on the analysed language, a given NP can create different clusterings. A particular mapping parameter might be considered characteristic for one language group, and a different one for another language:

> Given certain assumptions on the nature of the domain of quantification, minimal changes on what NPs can denote (which we will call the 'Nominal Mapping Parameter') lead one to expect certain clusterings of properties of bare nominal arguments. Such clusterings appear to be exactly what differentiates important language families. (Chierchia 1998: 358)

According to Chierchia, the Nominal Mapping Parameter and the neo-Carlsonian approach could solve the issue of BPs' ambiguity in English. The ambiguity is that English BPs can both denote kinds (cf. sentence (1) *Dogs bark*), as well as can they can favour a specific reading of a given utterance (such as *Dogs are in the garden*). The generic interpretation of BPs is the one suggested by Carlson. The second reading, a specific one, is also possible but depends highly on the context.

Chierchia's work appeared exactly 20 years after Carlson's seminal paper on genericity. The differences that emerge from Chierchia's view compared to the original theory of Carlson are therefore apparent, especially when it comes to the understanding and analysis of English kind-referring NPs. Chierchia points out main differences between the neo-Carlsonian approaches and suggests that most objections to the Carlsonian model can be rejected when it comes to neo-Carlsonian approach thanks to the new analysis model.

First of all, Chierchia allows for the generic use of definite article in his analysis, calling is SINGULAR GENERIC 'THE' (Chierchia 1998: 379). He agrees that singular definite form of a noun might well be interpreted generically, for instance in sentences as (42a) and (42b):

(42) a. The tiger is rare.

  b. The tiger roars. (Chierchia 1998: 379)

This assumption is proven by the fact that both sentences are correct and considered natural for the English speakers. As Chierchia himself points out, a corresponding mechanism was also observed by (Krifka et al. 1995) where the main focus of generic definite nouns concerns the so-called well-established kinds (WEKs, see the subsection 2.2.1.1).

The use of definite form is justified by Chierchia and depicted in his analysis model with the use of generic operators. He proposes his own analysis models and operators,

which are not retrieved from Carlson's analysis but base greatly on his main assumption, namely that generic sentences always have a phonologically null operators in their semantic structures. Since Chierchia's analysis methods base on the Universal Grammar (UG) and LF, we shall not delve further into that matter here.

On the contrary, the use of a plural definite form would not be possible in these sentences, as that would imply an episodic reading which normally does not occur with kind-predicates such as *be rare*:

(43)   a. *The tigers are rare.

        b. *The tigers roar.

As shown in the examples above, English, unlike the Romance languages, *has the definite article, but disallows generic or kind-oriented uses of it* (Chierchia 1998: 393). As we shall see in the following sections and chapters, Norwegian seems to accept definite plural forms in generic sentences, even though such uses of definite forms are rare and reserved for certain types of NPs (see section 3.2 and chapters 5 and 6).

As mentioned before, Chierchia's paper can be seen as ground-breaking, mostly because of its cross-linguistic approach which seemed to confirm the correctness of the proposed theory. Contrasting approaches to genericity in Germanic (English) and Romance (Italian) languages are presented, as well as a brief reference to Chinese and Russian. Such a wide perspective lets the reader see a broader scope of Chierchia's study, as well as the evolution of the Carlsonian approach in favour of the new hypotheses and theories.

In conclusions to his discussion, Chierchia points out the double nature of nouns, namely that they can appear as predicates, whereas as kind-denoting nouns they take on an argumental role (Chierchia 1998: 399).

Having discussed briefly the two formal approaches to genericity, let us now focus on the newer theories, namely the so-called 'simple view' on genericity and cognitive approach to genericity. The simplified theory of generics is based on the logical analysis of generic sentences and will be briefly described as a continuation and modification of Carlsonian and Neo-Carlsonian approaches. The latter one will form the base for the analysis presented in the following chapters of this dissertation. My main focus will be on applications of these theories, as well as the studies which have proved the cognitive approach to be more suitable for survey-based research and therefore for the purpose of this work.

The *simple view* on generics   In recent years, the standard view on generics, namely the approach proposed by Carlson and other researchers in the Generic Group, has been challenged and partially replaced by the so-called *simple view*. The theory proposed initially by Liebesman (2011) and further discussed by other researchers (see e.g. Leslie 2015 and Collins 2018), implies that there is no such thing as *Gen*-operator and that all

generic sentences have the same logical structure as atomic sentences. The theory seems to have certain flaws and does not explain all instances of generics but its controversial character might undermine over 40 years of research on the phenomenon.

The simple view, as opposed to the standard view represented in Carlsonian and neo-Carlsonian approaches, rejects the existence of all generic operators such as G', Gen, Gn and Gen. Instead, Liebesman opts for a simplified logical analysis of generic sentences with bare plurals.[6] The examples in (44) below show generic sentences with the material seen as generic ('wooden') and a reference to kind ('boots'):

(44) a. That is wooden. (Uttered while demonstrating a table.)

b. Boots are made for walking.

According to Liebesman, neither of the sentences in (44) contains Gen in their logical structure[7] and can therefore be analysed in the same way as a simple atomic sentence would be.

The absence of the Gen operator that would bound the variables in the sentences, although controversial, is not a new idea. In his seminal work on generics, Carlson (1977) proposed an approach similar to Liebesman's 'simple view' but rejected it due to numerous exceptions and sentences that could not be analysed with the use of this method.

In his paper, Liebesman discusses the possibilities that come with the simple logical analysis of generics. By addressing the issues that standard theories cannot account for, Liebesman claims his approach to be the ultimate theory of genericity, which, even though simple and concise, still needs to be adjusted in certain points. If the simple view proves sufficient to explain and semantically represent certain problems, not yet solved by the standard view, over 40 years of research might prove futile. The researchers have not yet managed to completely defend Gen-based theories, neither have they managed to undermine the simple view. The debate that has started in recent years is therefore still ongoing (cf. Leslie 2015, Sterken 2016 and Collins 2018).

2.4   Cognitive approach to genericity

The cognitive approach to genericity has made it possible to study the phenomenon in many different languages, very often in a form of cross-linguistic studies and with the focus on the cognitive status of generics (e.g. the theory of 'generics as default' coined by Leslie). When it comes to formal and modal theories presented in the previous sections, most of them concerned widely-studied languages and might therefore not be fully compatible

---

[6]Carlson's analysis (Carlson 1977) focuses on bare plurals and their (in)ability to refer to kinds. The simple view proposed by Liebesman also concerns the functions of bare plurals in English.

[7]For the logical sentence analysis the reader is referred to the original paper. The subject is out of the scope of this project and the main assumptions of the theory are presented in order to contrast the Carlsonian theories with modern views.

with other language groups. Cognitive linguists have proved that a more psychological approach allows for a much broader perspective in research which resulted in publishing of various studies not only on English but also, for instance, Mandarin Chinese (see e.g. Gelman and Tardif 1998) or Polish (Karczewski 2016) among many other languages.

In the following sections I shall focus on the notion of generic references as it is understood by the cognitive linguists, as well as the theories applied in research on genericity in English and other languages which will be utilised in the following parts of this dissertation. Since the study presented in this dissertation is based on both surveys and a corpus analysis, a wide scope of approaches was needed in order to structure the data collection and the analysis itself. The main core of the theoretical background consists therefore of approaches proposed by, among others, Radden and Dirven (2007), Radden (2009), Leslie (2007) and Leslie et al. (2011), as well as Ionin et al.'s approach to Acceptability Judgement (Ionin et al. 2011). The latter will be discussed further in chapters 4 and 6.

2.4.1 Understanding the generic reference

Radden and Dirven (2007) propose a following division of different types of reference in a language, where the generic reference allows for indefinite and definite forms. Figure 2.1 depicts these relations.
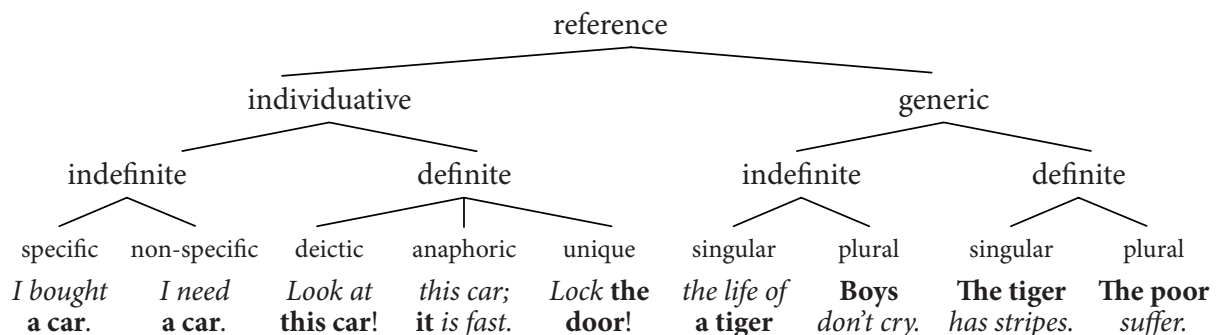
Figure 2.1: 'The main types of reference' (Table 5.5) in Radden and Dirven (2007: 111).

As it can be seen above, generic reference in English allows for the use of definite and indefinite forms but this does not apply to all Germanic languages. Norwegian for instance, seems to follow that pattern to some extent – singular and plural forms of generic nouns are allowed, as well as BNs which are not depicted on the scheme above. Further discussion on the matter can be found in section 3.2.

One of the most crucial claims of the cognitive grammar when it comes to generic references, is that CATEGORIES differ from CLASSES (Radden and Dirven 2007: 106). Categories are seen as abstract representations of kinds, whereas classes are actual, physical entities.

What is more, both a class and a category have their members. Members of a category are all subkinds, as for instance *Bengal tiger* and *Siberian tiger* belong to a category *tiger* (Radden and Dirven 2007: 106). When it comes to members of a class those are all individual entities of a given group, for example each and every dog on this planet can be considered a member of the class *dogs.*

Radden and Dirven compare generic reference to specific reference (called also individuative reference), where a speaker has to pick out an entity from the world of categories and make a mental connection to it. In the case of generic reference, that entity is the whole kind (Radden and Dirven 2007: 106). The process can be illustrated in the figure 2.2 below.

type of thing, category: 'tiger'

instance:
generic referent: class of tigers

speaker

hearer

grounding: referring expression: *the tiger*
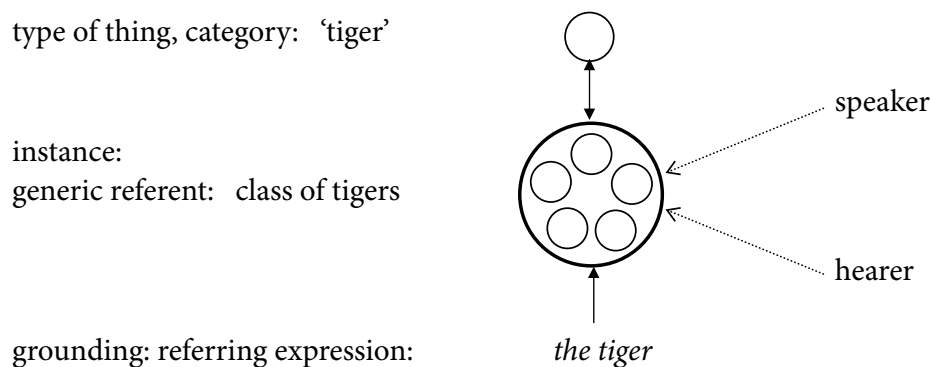
Figure 2.2: 'An act of generic reference: *the tiger (has stripes)*' (Figure 5.5) in Radden and Dirven (2007: 106).

One of the main characteristics of a category in a generic sense, is that all or most of its members should share the majority of the features typical for that category. This would be, as Radden and Dirven call it, an ideal situation. Nevertheless, it is very rarely the case. What we will see in the section on GOG, generic references are often subjects to overgeneralisations which means that a feature that is characteristic only for half or less of the population tends to be assigned to a whole species, as presented in the examples below:

(45)   a. Mammals give milk to their young.

        b. Ducks lay eggs.

Radden and Dirven point out that generic reference is, from a grammatical point of view, not very different from individuative reference since all languages that have been studied so far seem to be using the same referring expressions when talking about kinds, as they do when referring to specific objects (Radden and Dirven 2007: 106). This claim was confirmed by all linguists in their works on genericity (see for instance the phonologically null operators in formal approaches to the issue).

According to Radden and Dirven (2007: 107) and many other linguists, English allows for four different NP types that can appear in generic contexts and these are:

(46) a. *A tiger* hunts by night.   [indefinite singular generic]

    b. ? *An Italian* is fond of children.

(47) a. *Tigers* hunt by night.   [indefinite plural generic]

    b. *Italians* are fond of children.

(48) a. *The tiger* hunts by night.   [definite singular generic]

    b. ? *The Italian* is fond of children.

(49) a. ? *The tigers* hunt by night.   [definite plural generic]

    b. *The Italians* are fond of children.

The examples presented above do not include BN as a possible NP type for a generic reference and most of the linguists agree on that (see e.g. Carlson and Pelletier 1995; Mari et al. 2013a; Chierchia 1998).[8] As will be demonstrated the use of BNs in a generic context with count nouns is acceptable, at least for some nouns. Further discussion on the matter will be presented in section 3.2, as well as chapters 5 and 6.

In further parts of the analysis, Radden and Dirven give a short description of each of the NP types used in generic contexts. They stress the fact that not all of the forms convey the same meaning and that their individuative counterparts also bear different semantic functions.

Indefinite singular generic

Radden and Dirven describe the indefinite singular generic as a counterpart to individuative reference with the use of this form. They point out that it is *an element for its class*, meaning it is an entity that bears the features of the whole class but can be referred to with the use of the singular noun form.

The similarity is that in both individuative and generic reference, a speaker has to pick out an entity that he has in mind in order to refer to it (Radden and Dirven 2007: 108). In the case of generic reference, that entity is the whole kind, for instance all birds or all Italians, as shown in the examples below:

(50) a. A bird has a beak, wings and feathers.

    b. An Italian is a native, citizen or inhabitant of Italy.

    c. ? An Italian is fond of children.

Examples (50a) and (50b) are perfectly correct, whereas (50c) sounds rather awkward. The reason for this is that the first two examples capture the *essence* of the whole group. Assigning this essence to one entity from a given group does not exclude the generic reading (Radden and Dirven 2007: 108). The third sentence, on the contrary, cannot be seen as correct since being fond of children is not considered an *essence* of all

---

[8]This does not concern mass nouns (Hawkins 2015: 215-216)

Italian people. However, the interpretation whether or not something is considered an essence of a person/an object depends very much on the context, as well as one's world view and even general knowledge. The example (50c) does not suggest that being fond of children is a characteristic feature of a given nation, as such features are often stereotypes connected to a given culture (e.g. being fond of pizza, gesturing while speaking etc.). What is more, the essence might not be interpreted in exactly the same way by all speakers of a given language.

As has been mentioned before, some predicates can only be used with generic NPs and these are for instance *be extinct, be rare, be numerous*, among others. It is therefore important to note here that English indefinite singular nouns cannot occur with kind-predicates, as these can only refer to absolutely all members of a category, not some or even the majority (Radden and Dirven 2007: 108). That is also why the sentences (51) cannot be considered correct or even acceptable in the English language:

(51) a. * A sparrow is extinct.

b. * A dog is rare.

c. * A tiger is numerous.

Kind-predicates can occur with a singular form of a noun but it needs to be the definite form and even then the use of such constructions is quite limited. Radden and Dirven (2007: 108) give an example of a subspecies such as *the Balinese tiger* in the sentence (52):

(52) The Balinese tiger has been extinct for 50 years.

Indefinite plural generic

In generic reference, the indefinite plural NP type represents *indeterminate elements for their class* (Radden and Dirven 2007: 108), namely some entities from a category, not necessarily all or the majority of them, which represent the whole category:

(53) The large majority of *Italians* are Roman Catholics and for centuries, this has affected their art. *Italians* are proud of their artistic heritage. (Radden and Dirven 2007: 108)

The example above shows two different interpretations of BP, namely the first sentence states it clearly that the fact of being Roman Catholic applies to the vast majority of the population. The hearer will therefore assume that it is not the whole nation that can be identified this way. In contrast, the second sentence implies that being proud is somehow the essence of Italians. As Radden and Dirven summerise it:

> People tend to generalise on the basis of relatively few experiences, and the indefinite plural provides the adequate referring expression to do so: it con-

veys generalisations based on vague, impressionistic judgements and allows for exceptions. (Radden and Dirven 2007: 108-109)

Indefinite plural, called also for BP (bare plural), is said to be the most common NP type used in generic sentences (see for instance Carlson 1977; Krifka et al. 1995; Chierchia 1998; Mari et al. 2013a). As has been mentioned before, BPs are most likely to be interpreted generically among all available NP types in English but even despite this fact they are not free of ambiguities.

Definite singular generic

The definite singular form is seen by Radden and Dirven (2007: 109) as a *prototypical element for the class as such* in generic contexts. It can, for instance, refer to subkinds as we have seen in the section above, but the definite singular is also seen as being somehow characteristic and representative for the whole category that a speaker refers to.

Also in this case, individuative and generic references share some features, namely they refer to particular entities. In individuative reference that entity is one element, known both by the speaker and the hearer, and in generic reference that one particular entity is the whole kind (Radden and Dirven 2007: 109). Similar observations were made by other scholars as well, (see for instance Carlson 1977; Chierchia 1998).

Definite plural generic

The definite plural generics stand for *many elements for their class* (Radden and Dirven 2007: 109). Compared to inviduative reference, which concerns all elements from a given set, definite plural generic reference makes a generalisation about, presumably, the majority of entities from a given group but not necessarily the whole category (Radden and Dirven 2007: 110).

(54)   a. *The Italians* are generally not inhibited when interacting with the opposite sex.

      b. Football is the main national sport and *the Italians* are well known for their passion for this sport. Italy has won the World Cup four times.

Sentences (54a) and (54b) show how the definite plural form refer to the majority of individuals, making it an generalisation over the whole category of Italians. What is more, one must bear in mind that this type of generic reference seems to be reserved to humans only and using definite plural form in order to refer to animals, plants or objects might yield incorrect and/or awkwardly sounding utterances (Radden and Dirven 2007: 110).

Since *people typically generalise on the basis of many individuals that share a salient attribute* (Radden and Dirven 2007: 110), the use of plural definite form is possible - a generalisation does not need to apply to all members of the category to be understood generically by most speakers of a given language. Another particularity of this generic

NP type is that it clearly allows for exceptions, whereas in individuative reference the use of this NP type does not:

(55)   a. Open the boxes. = Open all of the boxes.

      b. The Italians are friendly. = Most/some/all Italians are friendly.

Sentence (55a) states clearly that all of the boxes a speaker refers to, should be opened. Individuative reference makes it explicit with the use of definite plural form which, normally, does not allow for any exceptions. Sentence (55b) however, might mean that either most, some or even all Italians are friendly and the hearer is not able to judge how many Italians the speaker actually has in mind. Most probably he or she has only met few Italians who turned out to be friendly people and made an assumption on this basis. Since the generalisation concerns humans and the socio-cultural differences between them, such an utterance is considered perfectly clear and unambiguous (Radden and Dirven 2007: 110).

Another aspect of the so-called generic human groups, is that they can be referred to with the use of adjectives, such as *the poor, the old* or *the young* (Radden and Dirven 2007: 110):

(56)   a. *The old* are still running the country.

      b. *The young* will take over soon.

Both sentences refer to all members of each of the categories. This process is called by Radden and Dirven (2007: 110) the metonymy PROPERTY OF A THING FOR THE THING which will be discussed in a more detailed manner in the following section. In short, we can say that referring to *the old* means in fact picking out many members of the category, stating something about them and creating in this way a valid, generic reference. The same concerns the sentence (56b) and any other utterance we can imagine where a generalisation over a human group would be made.

Interestingly enough, there seem to be a limited number of adjectives that can function that way. One cannot say, for instance, *the happy, the new* or *the thirsty* (Radden and Dirven 2007: 110).

Radden and Dirven (2007: 111) give an overview of all generic reference types and how they function. The figure 2.3 shows how four NP types in English refer to either one, many or most members of a given category.

The authors propose the following reading of the scheme:

- Indefinite singular generic: small circles stand for elements of a category with the same essential attribute; one of the circles is picked out to represent the whole category.
- Indefinite plural generics: three identical circles stand for elements from a category that share a characteristic feature; two other circles represent entities that do not

<div align="center">

i n d e f i n i t e   d e f i n i t e

</div>

singular	plural	singular	p l u r a l	
A tiger hunts by night.	*Americans are patriotic.*	*The tiger hunts by night.*	*The Italians are musical.*	*The young will take over.*
one single element	indeterminate elements	a prototypical element	salient elements	salient property
essential attribute	characteristic attribute	distinct attribute of a species	attribute of a human class	

Figure 2.3: 'Types of generic reference' (Table 5.4) in Radden and Dirven (2007: 111).

possesses that feature but still belong to the category.

- Definite singular generic: the big circle in the middle is an entity chosen to represent a whole category. It is a prototype for the category and posses a distinctive feature of a whole kind.
- Definite plural generic: in one of the readings a majority, or simply many, humans share a given feature that represents a category or a given property describes the whole category and it not necessarily characteristic for the majority of its members.

When it comes to indefinite singular form, it is seen as an essential attribute that needs to be assigned to one particular entity, in order for it to be understood generically and to represent the whole category. With indefinite plural form on the other hand, a speaker picks out many entities from a given category and assigns to them a characteristic feature. The feature does not need to apply to all members to be understood as generic. This is one of the main characteristics of the indefinite plural form in English. The BP form is most common in generic contexts, which has already been pointed out by numerous scholars.

The use of definite forms is slightly more restricted though. The singular form is used either when a speaker talks about a distinctive attribute that can be assigned to a kind, or when the attribute concerns a subkind, such as the Balnese tiger mentioned in the example (52).

On the contrary, the plural definite form is reserved almost entirely to humans and it can be understood as a reference to *all* members a speaker might have in mind and not necessarily to all members of a given category, as such a generalisation would be difficult to make. When talking about humans one does not usually need to refer to all people

in the world but rather to different socio-cultural groups inside the category *human*, for instance *the Italians* mentioned in the examples before.

A similar categorisation of generic reference types was proposed by Pettersson (1976) in his work on definite and indefinite forms in Swedish. The model seems to align more with the Norwegian language, mainly due to the fact that bare nouns are presented as possible generic forms.
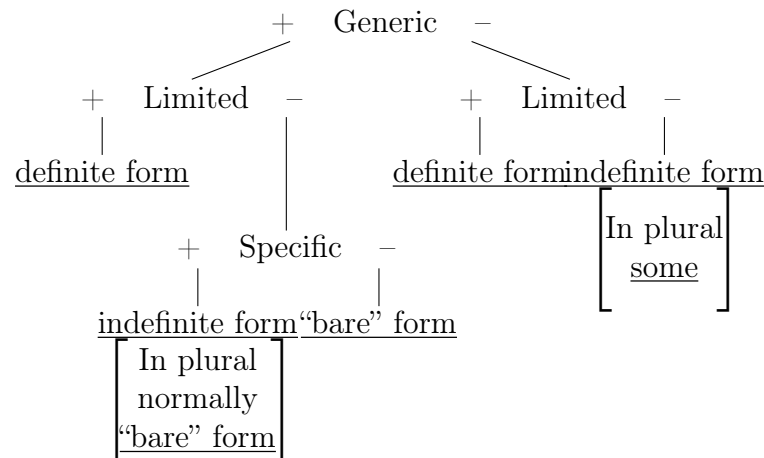
Figure 2.4: Reference types in Swedish according to Pettersson (1976: 121).

Pettersson divides generic nouns as either having the property 'limited' (*begränsad* +) or lacking it (*begränsad* -). Non-generic nouns are always specific (Pettersson 1976: 121) and those are associated with either indefinite form or a bare plural. Non-specific reference can be rendered with a bare noun, which was not depicted in the model for English above. English and Mainland Scandinavian languages do share a lot of features but what is correct and accepted in Scandinavian languages, does not need to occur also in English, as is the case with generic bare nouns.

The property depicted in the scheme as 'limited +' can be associated with definite nouns, whereas 'limited-' suggests indefiniteness. By using a definite noun, the speaker delimits the number of antecedents. This is not observed in the case of indefinite noun (Pettersson 1976: 122).

The "bare" form proposed by Pettersson, occurs in the case of non-specific reference, as can be seen on the graph. This element was missing in the models presented above, which all concerned English. In Swedish, and also Norwegian as will be presented in the following chapters, the bare noun is quite frequent in generic contexts and not only in the case of uncountable nous.

An interesting aspect can be also observed in the case of limited reference, both generic and non-generic – those are rendered with the use of definite forms. As has been seen before, definite plural generics in English are used exclusively to talk about people.

However, this is not the case in Norwegian and Swedish for that matter. The definite reference is still limited in its scope but its usage is wider in Swedish and Norwegian than in English.

2.4.2   Metonymy and generics

In recent years, various works on genericity seen from a cognitive point of view were published. One of the seminal works on the matter, where different types of generic references were given, is the one by Radden (2009). In his study on generics in English, Radden proposes a different analysis methods than those presented before in this chapter, namely *a metonymic and conceptual blending analysis* and makes three following claims:

1. Generic reference applies to types, where the type is invoked by way of an instance. Generic reference thus involves the metonymy INSTANCE FOR TYPE.
2. The characterizing type of generic reference allows for exceptions, i.e., it applies to a subtype of the type. Generic reference may thus also involve the metonymy TYPE FOR SUBTYPE.
3. Generic reference involves the conceptual blending of instance and type. (Radden 2009: 202)

In Radden's analysis, generic NPs substitute either some, the majority or all members of a given category. Similarly to his work of 2007, also here different types of generics are considered with the stress on their conceptual role in semantics. Generic reference, as presented by Radden, is understood in a broader sense, not only as a kind-reference. As the author says himself, *genericity is (…) much rather to be seen as forming a cline from full to marginal genericity* (Radden 2009: 199), indicating that both senses of the phenomenon, as discussed by Carlson and Pelletier (1995), are on the same spectrum.

The claim that *the type is invoked by way of an instance* implies that the singular instance influences the type and that both determine the generic meaning (Radden 2009: 204). This way the first claim can refer to kind-generics, where exceptions are not allowed – saying that a cat is a mammal implies that all cats ar mammals, even though the feature was asserted to an instance only.

When it comes to individuative and generic references pointed out by Radden, a metonymic reasoning is needed in order to differentiate between the two types. Metonymic reasoning means that both a speaker and a hearer are able to correctly identify whether an instance one speaks about is a TYPE or an INSTANCE. For example, *this jacket* in the sentence (57a) is an instance (one particular jacket), whereas *this jacket* in sentence (57b) refers to a type (all jackets of this particular style):

(57)   a. I really like *this jacket.*

      b. *This jacket* is our best-selling item. (Radden 2009: 201)

Usually it is the context in which an utterance is produced that determines the reading of an NP. Nevertheless, ambiguities may arise when for instance a non-linguistic aspect, such as gesturing, plays a role in a speaking situation, for example when a seller in a shop actually points at a given jacket and utters the sentence (57b). In such a case, an individuative reference is used in order to illustrate the generic meaning of the utterance (Radden 2009: 201). Most speakers will certainly understand that it is not in fact this particular jacket that was sold many times but a number of different items that looked exactly like the model that the seller is pointing at.

What happens in a situation such as the one described above, is a process of metonymy called INSTANCE FOR TYPE (Radden 2009: 201). This metonymic reasoning means that people are able to correctly distinguish generic generalisations from individuative references where an instance represents the whole category (type). Other studies conducted on this matter also seem to confirm this thesis (see for instance Gelman and Tardif 1998; Leslie 2007; Leslie et al. 2011).

In the metonymic analysis, a special attention is given to characterising sentences which clearly allow for exceptions. The well-known example about a lion implies that *usually* lions have such tails whereas in reality the sentence applies only to male lions (Radden 2009: 201):

(58)   A lion has a bushy tail.

The sentence is said to represent the metonymic relation where a TYPE (lion) stands for a SUBTYPE (male lion).

A similar relation can be observed when a category one refers to, stands for a subcategory. An example of (LINGUISTIC) SUBCATEGORY FOR CATEGORY metonymy can be the expression *housewife mother*, which in fact refers directly to the category *mother*, irrespectively of the fact whether she is a housewife or not (Radden 2009: 201-202).

Radden (2009: 224) proposes following functions to each of the generic NP types:

Each of the generic reference types proposed in Radden and Dirven (2007) and Radden (2009) have their own semantic roles. Indefinite singular, called 'representative generic', is used when an instance evokes a type (INSTANCE FOR TYPE, cf. Radden 2009: 202 and the discussion above), namely speaking about *a lion* means referring to the type *lion*. This metonymy works when a given instance possesses distinctive attributes that can also be assigned to the whole category (kind-reference).

The second type of generic reference, proportional generic, is in most cases an exclusive relation and can be observed in sentences such as (59):

(59)   Hedgehogs are shy creatures. (Radden 2009: 224)

	generic type	generic form	ex-/inclusiveness	generic meaning
(a)	representative generic	indefinite singular	exclusive	arbitrary instance representing its type
(b)	proportional generic	indefinite plural	exclusive/ inclusive	salient proportion of the type's reference mass
(c)	kind generic	definite singular	inclusive	prototypical subtype of a well-established type
(d)	delimited generic	definite plural	inclusive	delimited human set within a domain

Figure 2.5: 'Types of generic reference' (Table 2) in Radden (2009: 224).

The example shows an exclusive metonymy since there might be exceptions from the rule of hedgehogs being shy creatures. However, representative generic can be used in an inclusive manner when sentences with kind-predicates are formed:

(60)   Horses are mammals. (Radden 2009: 224)

Proportional generics are also interpreted as INSTANCE FOR TYPE metonymy. In the example (59), bare plural form refers to numerous members of the category but presumably not to each and every one of them, making it the same type of metonymy as in the case of representative generic. The same cannot be said about the example (60) which clearly evokes an inclusive relation - *being a mammal* is a kind-predicate and can therefore only refer to whole kinds, not only some members of a given species.

Radden perceives definite singular form as kind generic, in contrast to scholars representing formal theories who claim the bare plural to be model kind genericity (see for instance Carlson 1977; Carlson and Pelletier 1995; Chierchia 1998). What is similar in Radden's and the Generic Group's analyses is the claim that definite singular form is most often used with WEKs. *The kind is, however, constrained with respect to the level within its taxonomy and the Great Chain of Being* (Radden 2009: 224). This type of generic reference can involve either INSTANCE FOR TYPE or TYPE FOR SUBTYPE metonymy.

(61)   a. The tiger hunts by night.   INSTANCE FOR TYPE

       b. The albatross lays one egg.   TYPE FOR SUBTYPE

Delimited generic is an inclusive reference, reserved mainly to humans and most precisely to well-defined human groups (Radden 2009: 224), such as pasta-loving Italians in the example (62a) or *the rich* in (62b):

(62)   a. The Italians love pasta.

       b. The rich are controlling the country.

Delimited generics imply that the use of such sentences in the meaning of 'all', 'every' and so on is not possible. However, if such a reference was feasible, for instance

in a different language, one could perceive it as a regular reference, not a generic one. The issue of definite plural reference in generic context will also be discussed in empirical chapters of this dissertation (4, 5 and 6).

2.4.3 Types of generic generalisations

Among the many works on generics in English presented above, one finds multiple types of generic NPs, generic sentences and even generic texts. An overview proposed by Leslie et al. is an attempt at a comprehensive description of this diverse phenomenon, based on a great number of papers and studies conducted. Leslie et al. (2011: 19) propose the division of generic generalisations as depicted in the table 2.1 below:

Predication type	Example	Truth value of the generic	Description
Quasi-definitional	Triangles have three sides	True	Property must be universally true of all the members of the kind; no exceptions
Majority characteristic	Tigers have stripes	True	Property must be central, principled or essential (Gelman 2003; Medin and Ortony 1989) – namely, it must be directly related to the nature of the kind in question. It must also be prevalent though not universally had among members of the kind; while some exceptional members (e.g. albino tigers) fail to possess it, all the normal members of the kind must possess it
Minority characteristic	Lions have manes	True	Property must be central, principled or essential (Gelman 2003; Medin and Ortony 1989) – namely, it must be directly related to the nature of the kind in question. However, it must only be held by a minority of the kind. For our purposes we restricted these items to methods of gestation, methods of nourishing the very young, and characteristic physical traits had only by one gender

Predication type	Example	Truth value of the generic	Description
Majority	Cars have radios	True	Property must be prevalent among members of the kind, and must not be a principled connection (Prasada and Dillingham 2009).
Striking	Pit bulls maul children	True	Property must only be had by a small minority of the kind, and must signify something dangerous and to be avoided
False generalization	Canadians are right-handed	False	Property must be prevalent among members of the kind and there must be a sufficiently salient alternative property (e.g. being left-handed), so that the generic form of the predication sounds false or mistaken

Table 2.1: 'Various types of generic generalizations used in Experiments 1, 2, and 3' (Table 1) in Leslie et al. (2011: 19).

According to this classification, generic propositions are ordered by the meaning of predicates used in such sentences. The scale ranges from quasi-definitional generics, to which there seems to be no exceptions, to the so-called false generalisations where the value of generics is false.

The first category, namely quasi-definitional generics, asserts a property that is crucial and *universally true*. This means that quasi-definitional generics are always true and there are no exceptions that could be categorised in this way. An example given by Leslie et al. (2011) is the sentence:

(63)   Triangles have three sides.[9]

No person familiar with the basics of geometry would ever accept or utter a statement such as *Triangles have four sides*. A figure having four sides simply cannot be a triangle as having three sides is a distinctive feature of each and every triangle. Generic sentences of this type are quite rare since distinctive features are usually widely known to speakers of a given language. Quasi-definitional generics might occur most often in conversations with children or among them (see for instance experiments conducted by Gelman and Tardif 1998, as well as Gelman and Bloom 2007).

Sentences which present majority characteristic are a widely used generic sentence type. It asserts a feature that is true for the majority of group members or elements

---

[9]All examples in this section come from Leslie et al. (2011), unless specified differently.

from a given set. Such a feature can be, for instance, possessing stripes as in the example below:

(64)   Tigers have stripes.

The difference between quasi-definitional predications and majority characteristics is the fact that the latter allows for exceptions. We could imagine a tiger without stripes (for instance very rare pure white Bengal tiger), even though this would imply some sort of anomaly. nevertheless, classifying such an animal as a tiger probably would not pose problems to most speakers.

The third category presented in the table above concerns an interesting phenomenon, namely generic statements that are based on features possessed by the minority of a given group. It might seem that such sentences are not very common but in fact they occur quite often, especially when talking about natural kinds. Minority characteristics can concern for instance methods of nourishing the young or physical features:

(65)   Lions have manes.

The features presented in this category are so essential to the members of the group that possess them, that exceptions, even though numerous, are not taken into account by most speakers. The experiments conducted by Leslie et al. (2011) has shown that people are prone to overgeneralise, especially when it comes to this type of generics.[10] It is fully acceptable to say that

(66)   Ducks lay eggs.

even though it is only females in certain age and in certain period that do that. Nevertheless, the feature itself is so central to the kind that the exceptions from this rule do not make such generics false.

Majority predication, in contrast to majority characteristic, represent features that occur very often in a given group but are not principled. The example provided in the table 2.1 shows that most cars in fact do have radios but this feature does not define the car in any way – any car can be constructed without a radio and it will still fit to th category 'car' without being an anomaly, as it is the case with the albino tiger for instance.

The fifth category of predications used in generic sentences are striking features, assigned to the minority of the group. Such predications concern usually something dangerous and controversial:

(67)   Pit bulls maul children.

Mauling children is not a characteristic feature of this dog race, not is it central to the kind. Nevertheless, the fact of attacking someone, and especially children, is considered

---

[10]See section 2.4.3.1.

striking and therefore such characteristic might be interpreted as generic. What is more, the feature could be easily assigned to a different kind where the number of members possessing the feature would equally low:

(68)   Lions maul children. [own example]

The fact that any lion would be able to maul a child is obvious but since it does not happen as often as in the case of pit bulls, the sentence would be considered true but awkward. Another known example of such generics is the sentences:

(69)   Mosquitoes carry the West Niles virus. (Leslie 2007: 376)

Mosquitoes do not possess any distinctive feature that would make them carry the West Nile virus. Many different species are capable of doing that but the fact that some mosquitoes can be so dangerous is shocking and therefore the generalisation is assigned to the whole kind.

The last type of predicates discussed by Leslie et al. (2011) are false generalisations, as the predication made about Canadians in the example below:

(70)   Canadians are right-handed.

Even though the fact of being right-handed is prevalent and could probably be assigned to the majority of the Canadian population, the alternative (being left-handed) is too common to be ignored. in the case of false generics, the statistics do not matter but the property itself makes a given statement false. Technically, minority characteristics could be eliminated in the same way. The only difference is that in the case of natural kinds' characteristics, one does not focus for instance on ducks that do not lay eggs to describe the whole kind. There seem to be no alternatives to different methods of gestation – if a given kind bears their young in the womb, one would not consider egg-laying as an alternative and the other way around. One kind cannot have two methods of gestation, therefore generics concerning that feature will always be based on minority characteristic, not false generalisations.

Examples and categories presented by Leslie et al. (2011) show that predications do matter when it comes to reading of certain generics but it is in fact the cognitive abilities of each speaker that seems central in interpreting of such statements. Based on the experiments conducted in their research, Leslie et al. were able to test two hypotheses, namely generics-as-default hypothesis (GAD) and generic overgeneralisation effect (GOG).

2.4.3.1   Generics-as-default and Generic Overgeneralisation effect

Generics-as-default hypothesis presented by Leslie et al. (2011) assumes that quantified statements can be interpreted as generic ones. Such overinterpretation can occur when predication in a given quantified statement is considered to be a universal statement referring to the whole kind or a group (Leslie et al. 2011: 18). What is more, the GAD

hypothesis implies that the phenomenon concerns adults, not children. It is worth mentioning that many studies have focused on understanding of generics by children who, at a certain age, might confuse quantified sentences with generic ones. GAD hypothesis challenges this approaches by proving that also adults might be prone to overinterpretation of sentences that contain generalisations.

A number of studies conducted both on adults and children have shown that, in certain contexts and situations, speakers indeed are more likely to interpret quantified sentences as generic ones. Predications used in such statements play a central role in reading and interpretation of sentences (see the table 2.1 in previous section). In many cases generic sentences and texts contain predicates referring to the so-called *prototypical properties* (Leslie et al. 2011: 17). These can be for instance characteristics typical for a given species such as the colour (see example 14a about ravens being black), shape, size, gestation method and the like.

Among many predications that can be used to express generalisations over kinds, gender-specific properties seem to cause most misinterpretations (Leslie et al. 2011: 19). The fact of perceiving quantified sentences as generics has led the researchers to pose a following assumption: all generics might cause the GOG effect but it occurs most often in statements that contain characteristic properties (Leslie et al. 2011: 18).

The GOG effect is a theory based partially on GAD assumption, namely it implies that speakers are likely to make generic overgeneralisations even in the case of statements that are not strictly generic in their nature (for instance minority generics such as *Lions have manes* in (65)). In order to challenge this hypothesis and the assumptions connected to it, the researchers propose investigating the GOG effect in sentences that contain universal quantifiers such as 'most' and 'all'.

In the study conducted by Leslie et al. (2011) quasi-definitional sentences and false generic were used only for comparison purposes as, given the type of such statements, the GOG effect was not expected to occur in them. As the researchers claim, expecting the GOG effect to occur in quasi-definitional statements would be an error since such sentences do not allow for any exceptions. Therefore one cannot assume that the fact of assigning a feature to all members of a kind is an effect of an overgeneralisation (Leslie et al. 2011: 19).

The findings presented by Leslie et al. (2011) are as follows:

> [P]eople have a tendency to treat universally quantified statements as though they were generics. This tendency is what one would expect if generics involve more basic generalizations, while universally quantified statements involve more sophisticated, non-default ones. (Leslie et al. 2011: 29)

The experiments and results presented in the study show that the respondents were relying greatly on their cognitive reasoning when making assumptions about minority generics.

Even though they were aware of the fact that not all ducks lay eggs, they assumed such statements as true generics. The same process was observed when a group of respondents was primed with a knowledge about different duck species and the differences between. Nevertheless, that knowledge did not influence their reasoning and sentences such as 'all ducks lay eggs' were still considered true generics. This was due to the fact that the statements utilised in the experiments contained assumptions about prototypical properties, characteristic for a given kind.

The GOG effect, even though common in all of the experiments presented by Leslie et al. (2011), does not occur in all types of generic statements from the table 2.1: *the GOG effect was limited to minority characteristic and possibly majority characteristic items* (Leslie et al. 2011: 28).

The importance of the GAD hypothesis and the GOG effect can be seen especially when analysing numerous studies on understanding of generics in different groups of speakers. As Leslie et al. (2011) claim themselves, a number of theories and assumptions could be justified with the use of their findings, namely that the GOG effect does not need to be associated only with children. It may be true to assume to young speakers are especially prone to the GOG effect given their limited knowledge about the world and the developing understanding of certain language phenomena. Nevertheless, the experiments conducted in this study prove that children's tendency to overgeneralise in certain contexts may be due to a larger phenomena, rather than the lack of knowledge.

What is more, the mechanisms understood as GAD and GAG seem to function also in contexts completely new and unknown to the speakers. An example of such a study is the experiment conducted by Gelman and Bloom (2007). The experiment consisted in testing assumptions about 'dobles', creatures developed for the sake of the study. The respondents were presented with some facts about dobles and were supposed to judge a number of assumptions about the species afterwards.

The results provided by Gelman and Bloom (2007) show that not only were the speakers (both children and adults) able to correctly judge generic and non-generic assumptions about the novel kind, but they were also able to make new generic generalisations about it. When presented with a picture of a dobble without a certain feature, the respondents did not fail to distinguish exceptions (majority and minority generics) and still judge a statement as true generics (Gelman and Bloom 2007: 178-179). This phenomenon could in fact be explained with the GAD theory and the GOG effect.

2.5   Analysis model of the project

In the light of different research methods and theories presented above, it is clear that genericity is a complex phenomenon that can be analysed with the use of different methods and approaches. Since the nature of this study is two-fold (based on surveys and corpus

research), the use of cognitive approaches to the subject seems justified.

The theories that will be used for the analyses of the data are the cognitive models proposed by Radden and Dirven (2007) and Radden (2009). The matrix of generic references and their conceptual implications presented in Radden and Dirven (2007) will serve as a basis for the description of reference types in Norwegian. The conceptual blending analysis as proposed by Radden (2009) will be utilised to analyse the distribution of the NP types in the Norwegian material, as well as to study the cognitive status of generics in the language. Finally, a unified model for Norwegian generics based on the collected data will be proposed and contrasted with the original views based on the English data.

The model of generic generalisations proposed by Leslie et al. (2011) will be utilised for the interpretation of generic texts from the pilot study (chapter 4) and the sentences from the second survey (chapter 6). Since the main part of the project, namely the corpus analysis, concerned longer generic texts, the analysis models proposed by Radden and Dirven (2007) and Radden (2009) seem more suitable, as they allow to build a paradigm of generic NP types used in Norwegian.

As stated in chapter 1, the present study will utilise statistical tools for the quantitative analysis which will be presented in the final part of this chapter.

2.5.1 Statistical methods used in the project

The use of MMR methodology requires that the data be analysed both descriptively (QUAL) and statistically (QUAN). In this section, the statistical methods used throughout the dissertation will be described. The data from all three parts of the project, namely the pilot study, the corpus analysis and the AJT surveys, was analysed in the same way. Firstly, the descriptive analysis based on the cognitive framework is provided. Secondly, statistical tests are performed. Depending on the data's distribution, parametric or non-parametric analysis methods are used.

The results from all three parts of this project were analysed in the following way: with the use of simple statistics (distribution of the data, mean, median etc.), as well as more complex statistical methods such as the analysis of variance (ANOVA) and post hoc tests. The material gathered for this project consists of experimental data which does not always fit standard analysis models. Therefore non-parametric tests were applied where necessary.

The analysis of variance is used when more than one group (data set) is analysed. The analysis is performed in order to evaluate the differences between the groups (Mertens et al. 2017: 10). In this dissertation, the groups in the model are always five NP types, namely BN, two indefinite forms and two definite forms. The analysis of variance, similarly as other statistical models, relies on the null hypothesis. The null hypothesis is a reference point where all tested groups are equal (Mertens et al. 2017: 9). If the null hypothesis is

rejected after performing ANOVA, it means that there are statistical differences between the tested groups.

In each of the analyses the notions of p-value and F-statistic will be used. In short, p-value is used *to indicate statistical significance, not the effect size or importance of the hypothesized theoretical relationship* (Mertens et al. 2017: 152). Typically, p-value should be below 0.05 in order for the results to be significant. However, the value is used in many different analyses and tests in statistics, and it can differ depending on the conducted analysis, as well as its purpose.

The F-statistic is used in ANOVA to indicate potential differences between the tested groups. The F-statistic does not suggest where the differences are located, what their values are or whether certain groups differ more from the other ones (Mertens et al. 2017: 12). However, most of the computer programs report both the p-value and the F-statistic as those show whether the obtained results are significant or not. All of the statistical tests in this project were conducted and graphically depicted in R.

The analysis of variance is a very efficient tool since it allows to determine whether the groups in the analysed data differ from one another. However, as has been said, ANOVA does not show where the differences lie or whether one or more groups are different from the other ones. Performing only ANOVA is therefore insufficient and further tests are needed.

There are many ways of determining where the differences in the data are. The method used by many researchers and utilised also in this study are post hoc tests. They allow to determine which groups score differently, while still taking into consideration main assumptions of the analysis (Mertens et al. 2017: 12). The post hoc test that will be used throughout this dissertation in the case of ANOVA, is Tukey HSD.

The analysis of variance is a parametric statistical test which means that there are certain restrictions concerning the analysed data. Performing ANOVA requires therefore:

1. normal distribution of the data,
2. homogeneity of variance,
3. comparability of the groups.

If one or more of the requirements of ANOVA are not met (this can be verified e.g. with Levene's test and Shapiro-Wilks test), a non-parametric analysis should be performed (Mertens et al. 2017: 148). In this project, the non-parametric test of choice is Kruskal-Wallis test and an accompanying post hoc test is Dunn's test. The purpose of non-parametric tests is the same as in the case of ANOVA, namely they state whether there are differences between the tested groups, without indicating where potential differences are. Non-parametric tests are utilised when the distribution of the data is not normal or when the requirement of homogeneity of variance is not met. The results provided by

non-parametric tests are as valuable as the ones that can be obtained with the analysis of variance. However, due to lack of requirements (parameters) that need to be met, non-parametric tests can be utilised with different data sets and in different types of projects.

3 Research on genericity in Mainland Scandinavian languages[1]

Having discussed recent works on genericity and different approaches to the subject, let us now focus solely on genericity in Scandinavian Languages. This chapter is not meant to be an exhaustive description of the phenomenon but it serves as a theoretical foundation for the study described in the following parts of this dissertation. The main focus remains on the Mainland Scandinavian languages only, given the fact that the morphosyntactic structures of Danish, Norwegian and Swedish differ significantly from Faroese and Icelandic ones. Expressions of genericity are strongly dependent on the NP structure, as well as semantics of a given language. Therefore, the analysis of genericity in North Germanic languages such as Icelandic and Faroese is out of the scope of this dissertation. For a broader account on typological aspects of genericity see Behrens (2000). Dialectal differences in expressions of genericity in Mainland Scandinavian languages can be found in Dahl (2015) and Delsing (1993).

Studies on genericity have been developing in recent years, both when it comes to theoretical and empirical works. As has been mentioned in chapter 2, Scandinavian languages are not described in as much detail as for instance English or Romance languages. What is more, the existing data on genericity in those languages comes usually from general works on grammar or definiteness. Among such publications one needs to mention the works of Delsing (1993) and Julien (2005). Both of them mention generic references in Scandinavian languages but with regard to other grammatical phenomena.

Since, to my knowledge, there are no studies concerning generics in all Scandinavian languages (apart from certain experimental studies such as Skrzypek and Kurek 2018), I shall describe the existing material in three different sections, each of them concerning one language: Danish, Norwegian and Swedish respectively. The account of existing works on the matter in each of the languages concerns only main publications, namely seminal works and recent grammar books. Since the majority of the works discussed in this chapter concern a different topic than genericity itself, I have decided to focus on only few of them that present the phenomenon in most detailed way.

---

[1]Parts of this chapter have appeared in Norwegian in Kurek (2017).

3.1   Danish

Among the works on generics in Danish, the normative Danish grammar (*Grammatik over det Danske Sprog*, Hansen and Heltoft 2011) and Hansen's article (1994) on generic nouns are the most detailed ones. In Hansen (1994) we can find general rules concerning article use in generic contexts, whereas Hansen and Heltoft (2011) provide a detailed description of generic nominals, as well as a discussion on functions of each of the generic NP types.

In Danish, when it comes to countable nouns, one differentiates between four NP types that can take on a generic reading. The forms are indefinite and definite singulars as well as indefinite and definite plural forms. Bare noun as a grammatical form cannot occur in generic contexts with countable nouns (Hansen and Heltoft 2011: 500). The examples of all possible NP types are presented in (71).

(71)   a.   En tulipan kræver fugtig jordbund
             a   tulip   needs   humid soil

             Tulips need humid soil

        b.   Trompeten   er et blæseinstrument
             trumpet-DEF is a brass-instrument

             A trumpet is a brass instrument

        c.   Klæder   skaber folk
             clothes-Ø make   people

             Clothes make people

        d.   Japanerne   spiser med pinde
             Japanese-DEF eat   with chopsticks

             The Japanese eat with chopsticks (Hansen and Heltoft 2011: 137)

Even though each of the NP types presented above can occur in a generic sentence, it does not mean that the forms are interchangeable irrespectively of the context or a noun used in a given sentence. The authors provide a number of restrictions concerning the use of each of the NP types and their meaning in generic contexts.

Indefinite generics are discussed in detail in *Grammatik over det Danske Sprog* as indefinite forms are used in a number of different contexts, including generic ones. A matrix of indefinite forms is given in tables 3.1, 3.2 and 3.3,[2] after Hansen and Heltoft (2011: 477-478). Here we can observe that indefinite nouns have multiple readings and some of them include generics, e.g. uncountable nouns 'dust' and 'concrete' in table 3.2. As has been mentioned above, only uncountable nouns can occur as bare nouns in generic contexts in Danish.

Table 3.3 presents differences in meaning between non-generic and generic nouns, taking into account their markedness and the categories countable-uncountable. In the

---

[2]The tables were translated into English. For the original tables the reader is referred to Hansen and Heltoft (2011).

	Bare noun	Indefinite Individuliased	Indefinite Non-individuliased	Generic
Singular	pupil	a pupil	some pupils*	a pupil
	pig	a pig	some pigs	a pig
	tree	a tree	some trees	a tree
Plural	pupils	some pupils	**pupils**	**pupils**
	pigs	some pigs	**pigs**	**pigs**
	trees	some trees	**trees**	**trees**

Table 3.1: Overview of functions of the Danish bare noun, part I

	Bare noun	Indefinite Nominal function, incl. generic	Non-individualised
Uncountable	dust	dust	some dust
	concrete	concrete	some concrete

Table 3.2: Overview of functions of the Danish bare noun, part II

	Non-generic Individualised	Non-generic Non-individualised	Generic
Singular uncountable	dust	**some dust**	dust
	concrete	**some concrete**	concrete
Singular countable	a pupil	**some pupil**	a pupil
	a pig	**some pig**	a pig
	a tree	**some tree**	a tree
Plural	**some pupils**	pupils	pupils
	some pigs	pigs	pigs
	some trees	trees	trees

Table 3.3: Markedness of indefinite form.

following sections of the chapter, Hansen and Heltoft (2011) discus each of the NP types in generic contexts. An explanation of the notion of generic nominals is also provided:

> Ved et GENERISK NOMINAL forstås et nominal hvor substantivet i sammen-hængen betegner en hel art af genstande eller individer, ikke enkelte eksem-plarer. (Hansen and Heltoft 2011: 500)

> With a GENERIC NOMINAL one understands a nominal where the noun in the context refers to a whole kind of objects or individuals, not single specimens.

The definition of generics, even though simple and rather general, provides the readers with knowledge necessary to follow the discussion. At this point, the difference between *classic* generic sentences and habituals is not crucial as the general guidelines concerning article use in generic contexts focus on generic nominals. Habitual sentences on the other hand, do not need to contain generic nominals since predicates in such sentences might provide a generic interpretation.

The opposition definite/indefinite and one/many is not the same in the case of generic and non-generic reference in Danish, namely generic forms might sometimes be interchangeable (with a slight change in meaning in certain contexts), which cannot be said about the same forms in non-generic contexts (Hansen and Heltoft 2011: 501). As Hansen and Heltoft (2011: 501) point out, certain contexts are unavailable for some of the NP types, as in the examples in (72).

(72) a. Skrivemaskinen/*en skrivemaskine blev opfundet i slutningen af forrige århun-drede.

   The typewriter/* a typewriter was invented at the end of the previous century.

   b. Et kæledyr/*kæledyret kræver megen pasning.

   A pet/*the pet needs a lot of attention.

   c. Forkundskaber/*forkundskaberne er nyttige.

   Prerequisites/*the prerequisites are useful.

In the first example, the noun 'a typewriter' is unavailable in a generic context which can be due to the predicate used in the sentence, namely 'to be invented'. Certain predicates do not allow the use of indefinite singulars, as in English, and those are called kind-predicates (cf. Lyons 1977: 196 and Carlson and Pelletier 1995: 10).

The authors also draw attention to the similarity of Danish and French when it comes to the opposition between generic and non-generic plural references. In Danish, indefinite and bare noun generics cannot occur in *der* constructions. What is more, plural nouns in generic contexts do not cause stress reduction on the verb, contrary to the non-referential, non-generic forms (Hansen and Heltoft 2011: 501-502). A similar phenomenon

can be observed in French with the use of the so-called kind article, namely the definite article *les* as in (*on ne peut pas dresser les chats*, 'we cannot train cats') and the partitive article *des* in (*il dresse des chats*, 'he trains cats', Hansen and Heltoft 2011: 502).

The authors of *Grammatik over det Danske Sprog* raise a number of other aspects connected to genericity, such as the difference between generic and non-generic readings of the same NP types and generics in book titles. The first issue is illustrated with the examples in (73). The difference between a generic reference in (73a) and a specific reference in (73b) is provided not only by the context but also the subjects in each of the sentences.

(73)   a. man bør fælde en birk, men lade en bøg staå (generic)

             one should topple a birch but let a beech stand

    b. han fældede en birk, men lod en bøg staå (non-generic)

             he toppled a birch but let a beech stand (Hansen and Heltoft 2011: 503)

The subject *man* in (73a) is by itself generic which makes the whole sentence generic, irrespectively of the forms used. In sentence (73b) the subject *han* expresses a specific reference (to a certain person in question) which automatically makes the whole sentence non-generic. The very same context dependence can be observed when it comes to book and chapter titles, where certain forms are not allowed or are simply considered awkward:

(74)   a. Elefanten ('the elephant')

    b. Elefanter ('elephants')

    c. Elefanterne ('the elephants')

    d. *En elefant ('an elephant') (Hansen and Heltoft 2011: 503)

As the we read further in the text:

> [u]bestemt form singularis forekommer saaledes kun i ægte generiske sæt-ninger. Partielt generiske og ikke-generiske sætninger kræver af et generisk nominal at nominalet og dermed substantivet er pluralt eller har pluralisform. (Hansen and Heltoft 2011: 503)

> Indefinite singular form occurs only in truly generic sentences. Partially generic and non-generic sentences require from a generic nominal that the nominal and therefore the noun are plural or have a plural form.

Truly generic sentences are sentences with kind-predicates such as *be extinct* and those cannot refer to single members of a given kind but to whole kinds only. What is more, the definite singular form cannot be used with certain nouns and in certain generic contexts. As the authors claim, definite singular is restricted almost solely to kinds that have hyperonymes, such as *the saw* which belongs to a bigger group of tools (Hansen and

Heltoft 2011: 504). It is not crucial whether the speakers are aware of the hierarchy of species they talk about – certain nouns can occur as definite singulars just because they are perceived as belonging to a subcategory, as 'the eggplant' in the sentence 'the eggplant is edible' below.

(75) Ægplanten er spiselig

The eggplant is edible.

When it comes to definite generics, definite plural is also allowed in Danish, even though its use is limited to certain contexts. For instance, in most cases definite plural will be interpreted as anaphoric, deictic or inferentially definite (Hansen and Heltoft 2011: 506). However, when a speaker wants to refer to a characteristic feature that is shared by all members of a given kind, she can express it with definite generics:

(76) Hjerteanfaldene   er den hyppigste   dødsårsag i denne
heart.attacks-DEF are the most.frequent death.cause in this
befolkningsgruppe.
social.group

Heart attacks are the most common cause of death in this social group.

Hansen and Heltoft (2011: 506-507) differentiate between countable and uncountable nouns, as well as NEXUAL and INEXUAL nouns, providing a paradigm of generic NP types in Danish. The notions introduce the difference between the nouns that describe processes (nexual) and those that describe objects and materials (inexual), Hansen and Heltoft 2011: 172-173. The examples of each of the noun types are provided below.

(77)   a. En violin er et strygeinstrument.
a   violin is a string.instrument

A violin is a string instrument.

b. Violins er   strygeinstrumenter.
violins are string.instruments

Violin are string instruments.   (countable inexual)

(78)   a. Cement er opfundet af romerne
cement is invented by Romans-DEF

Cement was invented by the Romans.

b. Cementen   er opfundet af romerne
cement-DEF is invented by Romans-DEF

Cement was invented by the Romans.   (uncountable inexual)

(79) En stopprøve   er et politisk styringsredskab.
a screening.examination is a political tool

A screening examination is a political tool.   (countable nexual)

(80) Kærlighed gør   blind.
love   makes blind

Love makes one blind. (uncountable nexual)

As shown in the examples above, both categories can be divided further into countable and uncountable nouns. In each of the noun groups there are NP types that are not allowed in generic contexts, for instance we cannot say 'the screening examinations are political tools' as this would imply that each and every of such examinations plays such a role, which is not necessarily true.

In Hansen (1994) one finds a similar account of Danish generics, including the readings of each of the NP types. Similarly as in *Grammatik over det Danske Sprog*, in Hansen (1994) the author discusses also some particularities of genericity such as the acceptability of some indefinite singular generics and bare plural forms as a default generic form of any noun.

When it comes to the indefinite singular form, it can be used generically in certain contexts, namely when the sentence expresses a statement about all members of a given kind by referring to only one as a kind's model, as *en hval* ('a whale') in the sentences below. What is more, unlike plural indefinite, singular indefinite cannot obtain a generic meaning in a non-generic sentence.

(81) a. En hval   lever i vandet.
         a   whale lives in water
         A whale lives in water.

      b. En hval   lever ikke på landjorden.
         a   whale lives not on land
         A whale does not live on land. (Hansen 1994: 139)

Another noun form that does not occur in generic contexts very often is the definite plural form, even though certain contexts promote such readings of definite plural generics (as has been mentioned above), as in (82).

(82) a. Fuglene   har fjær.
         birds-DEF have feathers
         Birds have feathers.

      b. Bakterierne   gør stor nytte.
         bacteria-DEF-PL does big use
         The bacteria are of great use. (Hansen 1994: 143)

Indefinite plural form can be perceived as a default generic form of nearly any noun (Hansen 1994: 143), whereas this cannot be said about any other NP type that can be generic in certain contexts. There is a number of restrictions when it comes to use of singular forms, definite plural forms and bare nouns.

As Hansen (1994) points out, the fact that some forms are accepted as generic in certain contexts and incorrect/awkward in others, proves that genericity is rather a pragmatic problem, not a grammatical one. Similar conclusions were drawn by Carlsson (2012) in her study on genericity in Swedish texts (see section 3.3).

3.2 Norwegian

To my knowledge, there is no scholarly literature devoted solely to Norwegian generics. The subject is discussed for instance as a part of cross-linguistic studies as in Skrzypek and Kurek (2018), an aspect of definiteness (Norwegian normative grammar books) or it is mentioned in publications concerning the nominal system of Norwegian (e.g. Borthen 2003 and Halmøy 2016). One finds also minor works concerning the article system, as in Dyvik (1979), as well as publications that mention genericity as an aspect of other phenomena (e.g. the work of Lønning 1987 concerning mass terms and quantification where a few examples from Norwegian are given). An overview of the available Norwegian literature on the matter can be found in Kurek (2017).

Generics in Norwegian are rather diverse when it comes to the forms of generic NPs. Kulbrandstad (1998: 121) provides the following examples of Norwegian generics:

(83)   a. **Is** smelter ved 0 grader.
           ice melts   at   0 degrees
           Ice melts at 0 degrees.

      b. Du må   kunne stole på **en venn** /**venner**.
           you must can   count on a   friend /friends
           You have to be able to count on a friend/friends

      c. De   skriver stil   om   **katten**.
           they write   essay about cat-DEF
           They are writing an essay about a cat.

      d. Samler du på **frimerker**?
           collect you on post.stamps
           Do you collect poststamps?

      e. **Dinosaurene** døde ut for 60 milioner år   siden.
           dinosaurs-DEF died out for 60 millions years ago
           The dinosaurs died out 60 millions years ago.

The author comments on the examples in the following way:

> *I tilfeller hvor substantivet ikke refererer til noe spesifikt eksemplarer, men til hele arten, brukes dels bestemt, dels ubestemt form,* (Kulbrandstad 1998: 121).

> 'In cases where the noun does not refer to some specific entities but to the whole kind, sometimes definite form is used, sometimes the indefinite one.'

This rather vague explanation surely cannot be of much help for non-native speakers of Norwegian who would like to understand the differences in meaning in each of the examples in (83). However, certain guidelines do exist when it comes to article use, also in generic contexts.

There is a number of publications on the Norwegian grammar, some of them being meant for non-native speakers such as Strandskogen and Strandskogen (1995), MacDonald (1997 and 2009), Hagen (1998), Golden et al. (2008) among others. Apart from that, *Norsk referansegrammatikk* by Faarlund et al. (1997) is the most detailed work on the Norwegian grammar, meant both for native speakers of Norwegian, as well as for L2 learners.

In handbooks for foreigners, one finds rather scarce information on definiteness and the use of articles in Norwegian. Most of the books focus on the obvious, namely connecting indefinite articles with unknown information and the definite form with topics that have already been mentioned in the discourse. A slightly more elaborated comment on the use of Norwegian articles is proposed by Strandskogen and Strandskogen (1995: 52), where the authors say that *the definite article is used in a connection with a particular type or species of animal or object.*

Most of the Norwegian grammar books mention also different types of references. In some of the publications they are called 'special' and 'general' references (Strandskogen and Strandskogen 1995 and MacDonald 1997), in others 'specific' and 'general/generic' references (Faarlund et al. 1997, Kulbrandstad 1998, Golden et al. 2008, MacDonald 2009). The lack of consistency when it comes to the use of notions that denote basically the same phenomena is also known in English literature on the matter (cf. Lyons 1977 and Carlson and Pelletier 1995 among others).

Golden et al. (2008: 15) illustrates the difference between specific and general reference with the following examples:

(84)   a. **Fisk** er godt.
          fish   is good
          It's good to eat fish.   Generic

       b. **Fisken**   er god.
          fish-DEF is good
          The fish is good   Specific

       c. **Grønnsaker** er bra.
          vegetables   are good
          Vegetables are good.   Generic

       d. **Grønnsakene** er bra.
          vegetables-DEF are good
          The vegetables are good.   Specific

Bare noun *fisk* and the indefinite plural form *grønnsaker* obtain a general (generic) reading, whereas the definite forms are used as a specific reference. The notion of genericity and generic references is developed in more detail in Faarlund et al. (1997: 52), where generics are said to occur both in singular and plural, as well as definite and indefinite forms, which is illustrated in (85).

(85)   a.   **Ulven**   er et rovdyr.
          wolf-DEF is a hunting.animal

          The wolf is a hunting animal.

    b.   **En ulv** er et rovdyr.
          a   wolf is a hunting.animal

          A wolf is a hunting animal.

    c.   **Ulvene**   er rovdyr.
          wolves-DEF are hunting.animals

          The wolves are hunting animals.

    d.   **Ulver** er rovdyr.
          wolves are hunting.animals

          Wolves are hunting animals.

No clear difference between the NP types is given, except for the fact that they are all generic and that in this particular example the predicate 'to be a hunting animal' can be used with all four NP types. Furthermore, the authors provide an example (Faarlund et al. 1997: 292 after Barth 1980) where bare noun is used generically. The example is quite unfortunate in the sense that kinship terms are often used without articles in many different languages (Dahl and Koptjevskaja-Tamm 2001).

(86)   Hos baktamenene spiser aldri **mor**,   **far**   og **barn**   et eneste måltid
       with Bushmen   eat   never mother, father and children a single meal
       sammen.
       together
       In Bushmen's families, mother, father and children rarely eat together.

Norwegian bare nouns can be kind-referring or generic[3], both when it comes to countable and mass nouns. Borthen (2003) provides an analysis of Norwegian bare nouns, focusing solely on singulars,[4] also in generic contexts. In her PhD dissertation, she makes a number of claims concerning Norwegian bare singulars, stating among others that bare nouns are rather unlikely to occur in generic contexts, with a few exceptions (Borthen 2003: 30) which will be discussed later. The examples provided by Borthen (here the sentences in (87)) are based on her intuition as a native speaker of Norwegian.

(87)   a.   En katt har myk pels.
          a   cat has soft fur

          A cat has soft fur.

    b.   */??**Katt** har myk pels.
          cat   has soft fur

          A cat has soft fur

---

[3]The difference between the two references is a lot more pronounced than in English for instance.

[4]According to Borthen, Norwegian BNs are singular, even though other researchers do not agree with such classification of BNs (cf. Halmøy 2016).

   c. En bil er laget av metall.
      a   car is made of metal

      A car is made of metal.

   d. */??**Bil** er laget av metall.
      car   is made of metal

      A car is made of metal.

   e. Ola misliker jenter.
      Ola dislikes girls

      Ola dislikes girls.

   f. */??Ola misliker **jente**.
      Ola   dislikes girl

      Ola dislikes girls.

Such examples are called by Borthen sentences with *quasi-universal* generic readings (Borthen 2003: 30), where the subject or the object cannot occur as a singular bare noun and refer to the whole kind. However, the author does not claim that bare noun generics are completely unavailable in Norwegian. She provides a number of examples where the BN is fully acceptable and might even be considered the most suitable NP type Borthen (2003: 31).

(88) a. **Bil** er et kjøretøy.
       car is a vehicle

       A car is a vehicle.

    b. **Datamaskin** er et nyttig hjelpemiddel.
       computer   is a useful tool

       A computer is a useful tool.

The sentences in (88) are correct as they have a quasi-universal generic reading Borthen (2003: 32), meaning one could insert the determiner 'any' in each of the sentences and the meaning would remain the same. This cannot be said about the sentences without the quasi-universal reading in (89) that express regularities rather than general truths (cf. habitual sentences, Carlson and Pelletier 1995 and Mari et al. 2013b among others).

(89) a. Småbarn   spiser med **skje**.
       small.children eat   with spoon

       Small children eat with a spoon.

    b. Kari kjører (alltid) **bil** til jobben.
       Kari drives (always) car to work-DEF

       Kari always drives a car to work.

    c. Man bør   bruke **jakke** om vinteren.
       one should use   jacket in winter

       One should wear a jacket in winter.

The examples show that bare nouns can occur in generic contexts not only as subjects (even though the majority of them do) but also as objects and compliments. However, in most cases the (un)acceptability of the sentences is based on the speakers' language intuition and their own interpretations, as it also is the case in many of the theoretical studies on genericity. As Borthen puts it herself:

> Norwegian bare singulars can be generic. However, in generic statements of the type discussed in Carlson (1977), where the nominal in question gets a quasi-universal generic interpretation, Norwegian bare singulars are either out or highly exceptional, depending on how one interprets the examples in (19) [here 88]. (Borthen 2003: 32)

Some of the claims made by Borthen (2003) were later reevaluated by Rosén and Borthen (2017) in their empirical study based on the data from NorGramBank corpus. The authors refer to the original claims and examples but they also provide a new take on bare singulars in Norwegian. One of the aspects of the study are bare singulars in generic contexts, as presented in the examples in (90).

(90) a. **Hest** er et koselig dyr.
          horse is a nice   animal
          The horse is a nice animal.

      b. **Taxi** er dyrt.
          taxi   is expensive
          Taking a taxi is expensive.

Both sentences, (90a) and (90b), are *classic* generic sentences with bare nouns in subject position. While the example in (90a) can be translated into English with the use of definite singular, the subject in the example (90b) can only be rendered with a phrase 'taking a taxi' (Rosén and Borthen 2017: 221).

   Rosén and Borthen (2017: 224) discuss different construction types with bare singulars in Norwegian, introducing the so-called 'taxonomic' construction, as in (91). It means that bare singulars in (91a) and (91b) can be perceived as belonging to a bigger group of similar objects, other tools and other vehicles respectively (cf. hyperonyms mentioned by Hansen and Heltoft 2011).

(91) a. Det hjelpemiddelet som er mest brukt er **datamaskin**.
          the tool   that is most used is computer
          The type of tool that is the most used is the computer.

      b. **Buss** er et naturvennlig   kjøretøy.
          bus   is a nature.friendly vehicle
          A bus is a non-polluting vehicle.

A similar type of constructions where bare singulars can occur, are 'covert infinitival clauses' Rosén and Borthen (2017: 224), as in (92), where we can see that BNs can be

subjects (*sykkel*) as well as objects (*telt*).

(92)   a. **Sykkel** er kult.
           bike   is cool

           To ride a bike is cool.

       b. Jeg vil anbefale   **telt**.
           I   will recommend tent

           I will recommend (having/using) a tent.

The similarity of the examples in (91) and (92) is that both types of constructions can be interpreted generically, as they do not refer to any particular computers, busses, bikes or tents. It is rather the idea or a prototype of each of the objects that is expressed by a bare noun. In this sense, BNs become concepts of a computer, a bus and so on.

Borthen (2003) and Rosén and Borthen (2017) discuss only bare singulars, giving some examples of generic uses of this form. A slightly more detailed analysis was proposed by Halmøy (2016) who describes the nominal system of Norwegian. One of the many aspects of the Norwegian nouns analysed by Halmøy, is genericity which can be expressed with the use of all available NP types.

The work of Halmøy (2016) can be considered groundbreaking due to the fact that the author confronts the existing theories on generics and kind-reference with data from Norwegian. The data provided in the publication is based on the author's intuition and illustrates that what has been thought of as generics and kind-reference in English and related languages, seems to function differently in Norwegian.

Main claims made by Halmøy concern Norwegian bare nouns and indefinite plurals, and stand in stark opposition to Borthen's claims discussed earlier:

> [...] I will argue that the truly bare Norwegian Nouns **do not carry the features singular and indefinite** as is traditionally assumed, but that **they are marked for general number and are neutral with regard to definiteness**. I will furthermore propose that the Norwegian Indefinite Plural Noun is a true indefinite, not just neutral with regard to definiteness as is commonly suggested. [Halmøy 2016: 45, my emphasis]

These claims are revolutionary taking into account the scholarly literature on indefinites and generics (for an overview see chapter 2). Being marked for general number means that Norwegian BNs are countable but are in a way perceived as mass nouns or prototypes of a given kind (cf. *sykkel* and *telt* in (92)).

Furthermore, the author opposes the common claim that generics are directly triggered by kind-reference (a theory proposed by Chierchia 1998 among others) and provides a number of examples from Norwegian that seem to support this assumption. She also denies the possibility for the indefinite plurals in Norwegian to have three possible readings,

as is the case with English bare plurals (Halmøy 2016: 62):

> [T]he Norwegian Indefinite Plural is especially interesting in that it exhibits the rare property of distinguishing between generic and (true) kind predicates. While the English (bare) Plural has been famous since Carlson (1980) for receiving three interpretations as weak indefinites, generics and kinds, the Norwegian Indefinite Plural may only receive the two former readings.

This is illustrated with the following English (93) and Norwegian (94-96) examples (Halmøy 2016: 66-68), where the kind reading of Norwegian bare plurals is either unavailable or questionable:

(93)   a. Elks are not on the verge of extinction.   (Kind)

      b. Elks are magnificent animals.   (Generic)

      c. I saw elks and bears around the campsite.   (Weak indefinite)

(94)   a. Det   er **elger** i hagen.
        there are elks   in garden-DEF

        There are elks in the garden.   (Weak indefinite)

      b. Kjøpte du **lyspærer**?
        bought you light.bulbs

        Did you buy light bulbs?   (Weak indefinite)

(95)   a. **Elger** er flotte dyr   /pattedyr /har fire bein.
        elks   are pretty animals /mammals /have four legs

        Elks are pretty animals/mammals/have four legs.   (Generic)

      b. **Lyspærer** avgir mye varme.
        light.bulbs give a.lot warmth

        Light bulbs produce a lot of heat.   (Generic)

(96)   a. #**Elger** står i fare   for å bli   utryddet.
        elks   stand in danger to be extinct

        # Elks are threatened with extinction.   (Kind)

      b. #**Lyspærer** ble   oppfunnet av Edison.
        light.bulbs   were invented   by Edison

        # Light bulbs were invented by Edison.   (Kind)

      c. ?**Poteter** kom til Norge   først på 1600-tallet.
        potatoes came to Norway first in 16th.century

        Potatoes were first introduced in Norway in the 17th century.   (Kind)

The difference between (95) and (96) is that the former contains predicates that hold for all individuals of a given kind (all elks are mammals, all lightbulbs give light etc.), whereas the latter shows typical kind-predicates. We cannot say that each individual elk is threatened with extinction or that each and every lightbulb was invented by Edison

(Halmøy 2016: 68). This approach clearly contradicts the popular view that kind reference is directly connected to generics.

Halmøy (2016: 92) discusses also mass nouns in generic contexts, which again differ from English uncountable nouns when it comes to their form. In Norwegian (sentences in (97)), the preferred form of generic mass nouns is the definite form, whereas in English (sentences in (98)) mass nouns occur rather as BNs.

(97) a. (...)siden før mennesket oppfant **stålet** /* **stål**
    since before man-DEF invented steel-DEF /steel

    (...) since before man invented steel.

  b. **Risen** kom til Norge på midten av 1600-tallet.
    rice-DEF came to Norway on middle of 16th.century

    Rice came to Norway in the middle of the 17th century.

  c. **Vann** er livsviktig for alt velende på jorda.
    water is vital for all living on earth

    Water is vital for all life on earth.

  d. **Gull** er ikke bare sjeldent og meget vakkert...
    gold is not only rare and very beautiful

    Gold is not only rare and very beautiful.

(98) a. Man invented (*the) steel.

  b. (*The) water is becoming scarce.

  c. (*The) gold is rare.

As is shown in (99c) and (97d), also bare form of mass nouns is correct when it comes to generic contexts. The difference between the definite forms in the first two sentences and the bare forms in the latter two is that the former can access subkinds (types of steel and rice), whereas the latter cannot. A similar mechanism can be observed when it comes to the use of definite plurals in generic contexts (Halmøy 2016: 78).

(99) a. **Dinosaurene** /?#**dinosauren** er utryddet.
    dinosaurs-DEF /dinosaur-DEF is extinct

    The dinosaurs are/the dinosaur is extinct.

  b. **Hvalene** /?#**hvalen** er ytrydningstruet.
    whales-DEF /whale-DEF is extinction.threatened

    The whales are/the whale is threatened by extinction.

  c. **Pattedyrene** /?#**pattedyret** er ytrydningstruet.
    mammals-DEF /mammal-DEF is extinction.threatened

    The mammals are/the mammal is threatened by extinction.

(100) a. **Dinosaurene**, både tyrannosaurus rex, velociraptor etc, er utryddet alle
    dinosaurs-DEF both tyrannosaurus rex velociraptor etc are extinct all

    som en.
    as one

          The dinosaurs, tyrannosaurus rex, velociraptor etc, are all extinct.

      b. ***Dinosauren**, både tyrannosaurus rex, velociraptor etc, er utryddet.
         dinosaur-DEF both tyrannosaurus rex velociraptor etc is extinct

         *The dinosaur, tyrannosaurus rex, velociraptor etc, is extinct.

As the author points out, definite plural generics can denote super-kinds as in (99) but also sub-kinds (100), which cannot be said about indefinite plurals or definite singulars. Definite singular form can access kinds but in such cases *the reference is mass-like and the members of the class are conceived as homogenous* (Halmøy 2016: 79):

(101) a. **Elgen**   er drøvtygger.
          elk-DEF is ruminant

          The elk is a ruminant.

      b. **Elgen**   er sjelden på våre kanter av landet.
          elk-DEF is rare   on our parts of country

          The elk is rare in our parts of the country.

Norwegian, similarly to Danish and Swedish, expresses generics in multiple ways, each of them having certain restrictions when it comes to predicates and context. Normative grammar of Norwegian do not provide these, but the available literature (in particular Borthen 2003, Halmøy 2016 and Rosén and Borthen 2017) gives an overview of all generic NP types and contexts in which they can occur. An empirical study is needed in order to provide language data that is not based on the language intuition of one person.

3.3 Swedish

In *Svenska Akademiens Grammatik* (SAG) one finds the following notions connected to generics: *artbetydelse, sortbetydelse, typ, generell referanse*, among others. The first notion can be translated as kind-reference, the second as a subkind-reference, whereas the third notion, *typ*, is a class/kind (Teleman et al. 1999: 155, 234). The notion of general reference is a term widely utilised in scholarly literature on reference and generics.

    One noun can refer either to one member of a class, a whole kind or a subkind, e.g. singular nouns usually refer to prototypical members of a given kind (Teleman et al. 1999: 22), as in the examples below:

(102) a. **Diskmaskinen** är sönder.
          dishwasher-DEF is broken

          The dishwasher is broken.

      b. I större hushåll   sparar **diskmaskinen** mycket arbete.
          in bigger households saves dishwasher-DEF much   work

          In bigger households the dishwasher saves a lot of work.

In Swedish, similarly as in Danish and Norwegian, NP-level generics can be expressed with the use of four/five[5] different NP types. When it comes to countable nouns, the following NP types can be utilised in generic contexts (Teleman et al. 1999: 108):

(103)   En katt /Katten   /Katter /Katterna har   vassa klor.
        a   cat /cat-DEF /cats   /cats-DEF have sharp claws

        A cat/The cat/Cats/The cats has/have sharp claws.

Each of the sentences above can be paraphrased as 'all cats have sharp claws'. Generic reference can therefore be expressed both with definite and indefinite nouns, either in singular or plural (cf. Norwegian and Danish). However, certain restrictions apply, for instance:

- Indefinite singular generics imply that the generalisation concerns each an every member of the kind.
- Indefinite singular cannot be used as a kind-reference.
- Definite singular form can access sub-kinds.
- Bare plural can almost in all cases express a generic meaning but not necessarily a kind-reference.
- Certain fixed expressions and predicates favour the use of definite generics (e.g. *Priserna stiger igen*, 'The prices are rising again'; Teleman et al. 1999: 108-110).

The use of different NP types in different contexts is not caused only by the context or predicates used in a given sentence. The type of reference and its interpretation is also one of the key factors that allows the speakers differentiate between different meanings. One of the distinctions concerns the use of indefinite singulars in generics, as shown in table 3.4.

Entity *Individ*	**Subkind** *Sort*
He has bought himself a car. *Han har köpt sig en bil.*	Saab is a car that he likes. *Saab är en bil som han gillar.*
He came with the same car as me. *Han kom i samma bil som jag.*	He has the same car as me. *Han har samma bil som jag.*
Which car did he came with? *Vilken bil kom hon i?*	Which car does he have? *Vilken bil har han?*
Go and bring that car. *Gå och hämta den där bilen.*	That car over there is better than other brands. *Den där bilen är bättre än andra märken.*

Table 3.4: Individual reference vs. subkind-reference (Teleman et al. 1999: 23).

---

[5]The use of bare noun with countable nouns is somehow limited compared to Norwegian but can occur in certain expressions and/or with certain predicates.

In Swedish, plural forms can take two different plural suffixes, of which *-er* is used with kind-references (*sortbetydelse*, Teleman et al. 1999: 23), as in *många frö-er* (many seeds) being an equivalent to *många sorters frö* (many types of seeds). The expression *många frön* means many individual seeds, not seed types.

In his article on definite and indefinite forms, Pettersson (1976: 121) divides generic references into 'limited' (*begränsad +*) and 'non-limited' (*begränsad -*) ones. Limited generics are expressed with the use of definite form, whereas non-limited ones with indefinite form or bare nouns. The author explains that every non-generic noun always expresses a specific reference but specific nouns can also be used in generic contexts. An example of that is a generic sentence in (104) with definite singular noun.

(104)   Gärdsmygen   är en flyttfågel.
        Eurasian.wren-DEF is a   migratory.bird
        The Eurasian wren is a migratory bird.

The provided example shows the use of indefinite singular with a subkind. The sentence and the NP type used conform to the claim that in most Germanic languages subkinds are accessed by definite forms, especially singular ones (cf. 'the Berber lion' examples by Lyons 1977). The Eurasian wren is interpreted by Pettersson as a prototypical bird that represents the whole kind (Pettersson 1976: 124), whereas a subkind-reference is understood rather as referring to the whole group of entities (in this case all Eurasian wrens) through a definite singular reference.

In the text, Pettersson (1976) describes all references to kinds as generics, not differentiating between a generic reference *per se* and a kind-reference, as has been discussed e.g. in SAG, Halmøy (2016) and other scholarly literature on the topic (see chapter 2).

What is novel in the study is the notion of 'established kinds' (*etablerade klasser*, Pettersson 1976: 128) such as professions and nouns describing one's activities (cf. well-established kinds, Carlson and Pelletier 1995). These occur without articles when they have generic readings. Using a noun from this group with an article would usually suggest a specific reference and a different meaning.

(105)   a. Han är en bödel.
            he   is an executioner
            He is an executioner.

        b. Han är bödel.
            he   is executioner
            He is an executioner.

The difference between (105a) and (105b) is that the first sentence implies that the person belongs to the natural kind 'executioner', whereas the latter means that the kind in question is an established class (Pettersson 1976: 129). The difference may seem only stylistic but it does change the reading of the sentences. As Petterson mentions, the

choice of bare noun of indefinite form is highly subjective and can often depend on the context in the moment of speaking. What is more, the sentence in (105a) implies that being an executioner is not necessarily a profession of that person but rather that she behaves in a way that a prototypical executioner would behave in (being cruel, ruthless etc.).

The use of generic bare nouns in Swedish is limited to certain cases, one of them being sentences with *ha* (to have), as in (106) by Pettersson (1976: 130).

(106)   a. Anja har bil, men Kennet har bara cykel.
           Anja has car but Kennet has just bicycle

           Anja has a car and Kennet has only a bicycle.

        b. Jag har   körkort.
           I   have driver's.license

           I have a driver's license.

'Car', 'bicycle' and 'driver's license' are understood here as prototypes or notions, not a particular car, bicycle or driver's licence. The fact that these are used as bare nouns may suggest that also in Swedish certain common nouns can be perceived as well-established kinds.

Fixed expressions and well-established kinds are similar to uncountable generic nouns when it comes to the form they take on. Both of these noun types usually occur as bare nouns of definite nouns, each of the two having a slightly different meaning. The bare form suggests a more general and almost prototypical reference (cf. Pettersson's take on that matter), whereas the definite form suggests that the speaker focuses on what the noun denotes (Teleman et al. 1999: 111). The examples in (107a) illustrate this difference.

(107)   a. Guldet   /?Guld har sjunket i pris.
           gold-DEF /?gold has sank   in price

           Gold has become cheaper.

        b. Guld /?Guldet   fräts   inte av syra.
           gold /gold-DEF is.taken not by acid

           Gold doesn't dissolve in acid.

The definite form in (107a) can denote e.g. the price of gold in our time, in a given period etc., whereas the indefinite form in (107b) implies that being resistant to acid is a characteristic feature of the material (a truly generic reading, Teleman et al. 1999: 111).

Apart from the theoretical work, an empirical analysis of the matter is also available. The study of Swedish generics conducted by Carlsson (2012) is probably the biggest empirical research on genericity in Mainland Scandinavian Languages. It focuses both on NP-level and sentence-level genericity and is based on a corpus consisting of generic texts and a film transcription, both of which are descriptions of animal species. The choice

of this particular genre might have been dictated by the need to find unambiguous texts that would (in most cases) have only one possible reading, namely the generic one (see section 2.2.3).

Genericitet i text by Carlsson (2012) consists of two main studies: 1) a text analysis of 36 texts about animal species and 2) an analysis of a 47-minutes long film (fully transcribed) about the species *Macaca fascicularis*. In the first study, the author focuses on the distribution of different NP types in generic contexts (Carlsson 2012: 57). The second study is a continuation of the first analysis and its main goal is to examine context dependency of generic NPs (Carlsson 2012: 125).

What Carlsson (2012) claims already in the introduction to her work is that genericity can no longer be treated as *an all-or-nothing phenomenon* because the context is often crucial in deciding whether a sentence or an NP can be interpreted as generic or not. As her first study shows, the accessibility of non-generic expressions is much greater than the one of generics. It is therefore crucial to study the phenomenon in texts which give a wider context, instead of studying individual sentences.

Both the advantage and the disadvantage of Carlsson's study lie in the material. She does not rely on her language intuition in order to provide the data for the analysis, as was seen in the studies of e.g. Pettersson (1976), Hansen (1994), Borthen (2003) and Halmøy (2016). Instead, she opts for written texts and a transcription of the spoken language. The two data sets concern biological texts. On the one hand, focusing only on descriptions of animal species narrows the context and possible readings of NPs. On the other hand, the consistency in choosing the material makes it possible to analyse in detail one particular aspect of genericity, which is context dependency in this case.

The existing literature on generics in Mainland Scandinavian languages consists of normative grammar books and a few detailed works on the phenomenon, only one of them being an empirical study. The majority of the publications presented in this chapter focus on sentence analyses, not a broader context of generics such as analysing generic texts. Apart from Carlsson (2012), all researchers opt for language data created for the sake of their studies.

Even though the Norwegian scholarly literature on generics seems abundant at first sight (Borthen 2003, Halmøy 2016, Rosén and Borthen 2017 and to some extent Faarlund et al. 1997), there is none empirical study that would put the existing theories to test. An attempt to do so was the revised version of the original study of Borthen conducted by Rosén and Borthen (2017) and based on a Norwegian corpus. However, the main focus of the study were bare singulars, not generics in particular. The purpose of this work is therefore to provide an empirical study that puts the proposed theories to test and analyses generics in a broader context.

4 Pilot research

4.1 Method

In previous chapters an overview of the existing scholarly literature on genericity was provided, focusing mainly on the theories and approaches to genericity. In this section, an overview of different methodologies will be given, as well as the methodology chosen for this pilot study will be presented.

The pilot research for this project was conducted in 2017 and consisted of a survey published online (see the details in the next section), where the participants had to fill out texts with gaps. Before conducting the study, a number of different strategies and approaches were analysed in order to choose an appropriate one that would fit the resources available at the time. The goal of the pilot study was to collect a sufficient amount of data and to be able to see as many generic contexts as possible.

Many of the studies concerning genericity in English and other widely spoken languages, are corpus analyses. There are Norwegian corpora available online, however tagging them for genericity requires a lot of work and is very time consuming, since the texts not always can be downloaded. The corpus approach was therefore used in the main core of the study and not for this pilot research.

Another option of testing whether a certain aspect of language poses problems for native speakers is conducting a survey. There are different types of surveys, depending on the tested phenomenon. A survey based on the Acceptability Judgement Task seemed like a good choice for numerous reasons. First of all, this method allows to check how native speakers of a given language perceive the tested problem. What is more, the participants are given quite a lot of freedom when it comes to choosing among the available answers as they can provide one or many answers to one question. Finally, the respondents can grade each of the answers, showing in this way that certain phenomena may be interpreted in many ways.

When it comes to genericity, speakers' interpretation and language intuition are crucial, as it is a semantic, rather than morphosyntactic phenomenon. Among the studies based on surveys, the works of Oosterhof (2008) and Ionin et al. (2011) were considered as possible models for this pilot research. Oosterhof tested the semantics of generics in

Dutch and related languages, including also local varieties of Dutch.[1] His study is based on a corpus analysis combined with surveys conducted among native speakers of analysed languages. The idea of combining these two methods will be presented in chapters 5 and 6.

In his study, Oosterhof (2008: 110) conducted a survey that consisted of 64 characterising sentences and 18 kind-sentences. The sentences were then judged by 29 native speakers of local and regional varieties of Dutch and Frisian. The respondents could grade the sentences from 1 (completely unacceptable) to 5 (completely acceptable) (Oosterhof 2008: 114). An important aspect of this survey is that the test items were adjusted to the dialects spoken by the respondents, namely certain kind-predicates were omitted as they do not occur in those dialects or are expressed in a descriptive way (Oosterhof 2008: 112-113). The method proposed by Oosterhof is certainly efficient, especially when combined with a corpus study. However, the scope of the study was too wide for a pilot research.

The second possible model for a pilot study was the analysis conducted by Ionin et al. (2011). The analysis is based on Acceptability Judgement Task and concerns three languages: English, Spanish and Brazilian Portuguese. The respondents were given 20 texts with gaps. Each of the texts concerned one topic (one NP) and had one gap to fill. The respondents could fill the gaps with all possible NP types and in addition grade the answers on a scale from 1 to 4, where 1 meant an unacceptable sentence and 4 an acceptable one (Ionin et al. 2011: 973).

The method proposed by Ionin et al. allows to check respondents' intuitions and the way genericity functions in a given language. It seemed therefore like a good way of testing genericity in Norwegian. However, the method was slightly modified, namely only the first part of the study was used, namely filling out the texts with gaps. What is more, the rating of the answers was replaced with a place for respondents' comments. This was done in order to obtain a more detailed feedback and to be able to analyse the respondents' language intuitions. However, this did not function as planned and was corrected in the second survey based on the corpus material (see chapter 6). The structure of the survey and respondents' comments will be discussed in greater detail in the following section.

4.1.1 Survey structure and used tools

The survey[2] was conducted in 2017 among native speakers of Norwegian (the respondents are described in section 4.2 of this chapter). The survey was designed in Google Forms tool. It is a free software that allows to create simple surveys and share them with unlimited number of people. Due to the simplicity and accessibility of the tool, the

---

[1] There are many local varieties of Norwegian also but these are out of the scope of this study. Further discussion on this topic can be found in chapter 7.

[2] Since the project contains two surveys, the pilot study is referred to as 'Survey 1' in graphs and tables, whereas the AJT survey from chapter 6 as 'Survey 2'.

survey could be designed and published in a short amount of time and without any storage restrictions.

As has been stated before, the form of the survey was inspired by the study of Ionin et al., with certain modifications. The first part of the study remained the same, namely it included short generic texts, each one of them with a gap. The texts were written in Norwegian and proofread by a native speaker of Norwegian. The diversity of the texts' topics not only allowed for a wide range of contexts but it also made it difficult for the respondents to guess the subject of the survey. The part of the study of Ionin et al. that was modified concerned the rating of the answers. Instead of the rating scale, a comment field was inserted to each of the survey questions. The comments from the respondents were supposed to provide additional information on the chosen answers and their interpretation.

The survey did not include any filler items. Placing additional items in the survey to disguise the participants would only make the time necessary to answer the test longer and doing so could also result in fewer participants. What is more, the goal was to include as many different test items as possible, without making the survey too long. The number of participants was crucial for the pilot study so the survey's length and structure had to be adjusted according to the criteria mentioned above. However, filler items were included in the second and main survey of this project (see chapter 6).

Since the goal of this pilot study was to determine which NP types are used generically and in what contexts, the respondents could choose more than one answer in each of the survey questions. This way it was possible to analyse different generic contexts and the grammatical forms used in them.

All of the texts were written in *bokmål*, one of the two written standards of Norwegian. The majority of the Norwegian population utilises *bokmål* as their primary written language. Furthermore, choosing only one written standard allowed to unify the survey and the collected results. A similar study could also be conducted on *nynorsk* in order to test for any differences between the two standards. Further discussion on Norwegian, its variants and local varieties can be found in chapter 7.

The 30 texts in the survey concerned different nouns and NPs (one NP per text) in order to give a broader context. The nouns used in the survey were as follows[3]:

1. styrketrening – strength training[4]
2. eple – apple
3. appelsin – orange
4. venn – friend
5. politiker – politician

---

[3]The forms are given as BNs. See the Appendix A for survey texts.

[4]'Strength training' was used as a countable noun in the survey. The case was discussed with a native speaker of Norwegian.

6. hund – dog
7. religion – religion
8. smarttelefon – smartphone
9. lastebil – truck
10. sykkel –bicycle
11. melk – milk
12. musikk – music
13. barn – child
14. øl – beer
15. brødmat – *breadstuff
16. lydbok – audiobook
17. ryggsekk – backpack
18. stearinlys – candle
19. dunjakke – down jacket
20. stråhatt – straw hat
21. magefølelse – gut feeling
22. grønn lampe – green lamp
23. hvit kjole – white dress
24. stor trykkskjerm – big touchscreen
25. vanlig ølflaske – regular beer bottle
26. hvit skjorte – white shirt
27. keramisk komfyr – ceramic stove
28. grått hår – grey hair
29. rødt hus – red house
30. hvitt flagg – white flag.

Some of the nouns could be interpreted as well-established kinds, for instance 'white dress' or 'dog'. However, testing for WEKs was not one of the goals for this study and shall be discussed later (see chapters 5 and 6).

In the survey, there were 4 uncountable nouns, namely: 'milk', 'music', 'breadstuff' and 'grey hair'. Uncountable nouns, both in generic and specific contexts, occur either as bare nouns or as definite nouns. One could, theoretically, use an indefinite form such as 'a milk' but that would imply that one means one kind of milk, rather than an uncountable substance. The main focus was therefore on countable nouns that can take on either of the 5 NP types in generic contexts.

In order for the survey to show valuable results, a set of control questions was posted in the first part of the online form. These concerned the following aspects:

- age
- sex
- education
- origin
- languages spoken.

The first part of the survey was obligatory and it was not possible to proceed to generic texts without filling out the introductory part. When it comes to languages spoken by the respondents, both native and second/third languages were taken into account. The respondents had to answer in which language they count and dream, which allowed to verify whether Norwegian was their native language. The respondents who chose any other language in those two questions were not taken into account.

The link to the survey was published on a forum concerning the Norwegian language. This made it possible to reach people of different age, education and professions. Even though the forum was meant for people primarily interested in the language, the participants were not only linguists or teachers. The survey was available on the forum during two weeks.

4.2   Respondents

The total of 630 people answered the survey questions, of which 599 were identified as native speakers of Norwegian. As has been said before, the participants who chose a language different than Norwegian in questions concerning their 1st language and the language they dream/count in, were not taken into account. The table 4.1 and the graph 4.1 present the age and sex of the respondents.

Age	Women	Men	Total
16–19	3	1	4
20–25	30	12	42
26–30	50	10	60
31–40	89	21	110
41–50	108	35	143
50+	163	77	240
Total	443	156	599

Table 4.1: Survey 1 – age of the respondents
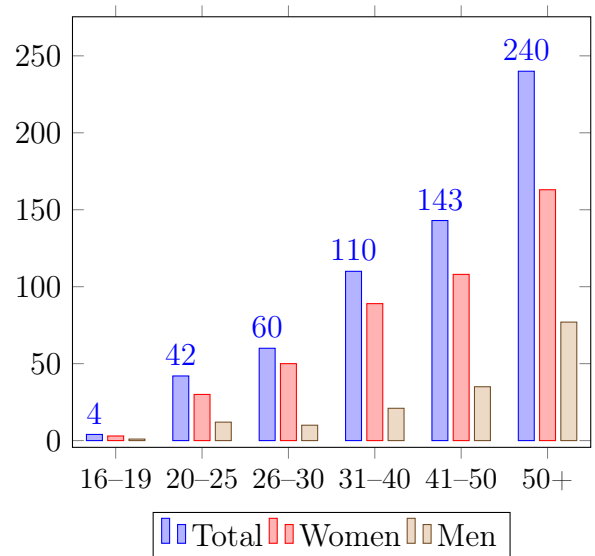
Figure 4.1: Survey 1 – age of the respondents

The majority of the people who took part in the survey were women aged over 40. Women dominated in every age category. The information about the participants' age and sex was collected only for statistical purposes, not in order to analyse whether women would answer differently than men or the other way around. Even though the differences between men's and women's language might be very subtle in Norwegian, this aspect was out of the scope of this project.

Education	Women	Men	Total
primary school	3	2	5
high school	43	17	61
university	397	137	533

Table 4.2: Survey 1 – education of the respondents
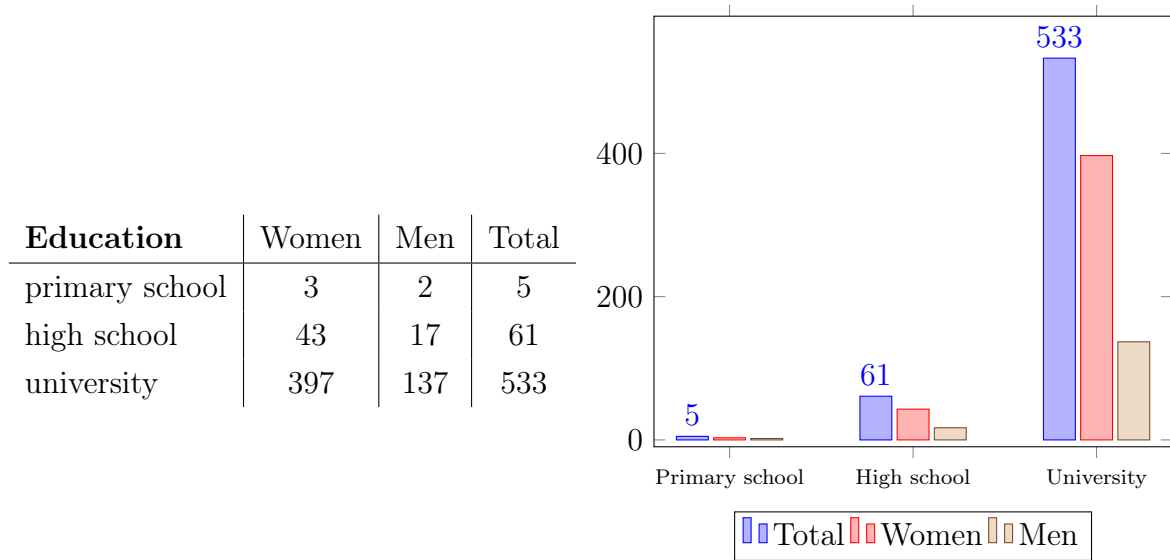
Figure 4.2: Survey 1 – education of the respondents

When it comes to participants' education, the majority chose the option 'university' (533 people in total, see table 4.2). This however did not indicate whether a person obtained a university diploma or not. In Norway, one can take a 1-year university course without obtaining a bachelor's or master's degree and this can still be considered as higher education. Again, this information was collected only for statistical purposes, not to check whether the level of education influences in any way the participants' language. Such aspect would surely be an interesting subject to study but was out of the scope of this pilot research.

The fact that the majority of the respondents have higher education is nevertheless interesting. It might just be that people primarily interested in the language would be more willing to take part in such a study than those who only read the posts on the forum but do not participate in the discussions. Another explanation is that education and interest in the language correlate but the results of this survey do not allow for such generalisations. It is safe to assume that people with higher education were generally more willing to participate in the survey.

Another question in the first part of the form concerned the origin of the respondents. In 2017, when the survey was carried out, there were 19 municipalities in Norway. All of

them, plus the option 'other' are listed in table 4.3 and on the accompanying map. The location 'other' was chosen only by two participants who also marked 'Norwegian' as their primary language. Other respondents who did not chose any of the regions of Norway as their place of origin also chose different languages in control questions (dreaming/counting in a given language). This has made the elimination of non-native speakers of Norwegian much easier and diminished the risk of a mistake.

The map 4.3 shows the regions which the participants came from. Most participants were from Oslo region (municipalities Oslo and Akershus), the North (municipality Troms) and the West (municipality Hordaland), with quite a high percentage of people from Trondheim region. All of these, apart from the North, are most densely populated areas of the country. The dialects spoken in those parts of the country belong to the main dialect groups which include: *østnorsk* (Eastern Norwegian), *vestnorsk* (Western Norwegian), *trøndersk* (Norwegian from Trondheim region) and *nordnorsk* (Northern Norwegian).

As all participants were informed that the survey was written in *bokmål*, we can assume that the dialects did not influence the results. What is more, the results did not show a large number of outliers that would suggest the influence of the dialects on participants' written language. However, such influence cannot be completely eliminated without further analyses (see the discussion in chapter 7).

Regions	Respondents
Akershus	57
Aust-Agder	10
Buskerud	26
Finnmark	21
Hedmark	23
Hordaland	41
Møre og Romsdal	30
Nordland	35
Nord-Trøndelag	21
Oppland	21
Oslo	99
Østfold	22
Rogaland	29
Sogn og Fjordane	11
Sør-Trøndelag	32
Telemark	26
Troms	56
Vest-Agder	13
Vestfold	24
Other	2

Table 4.3: Survey 1 – origin of the respondents
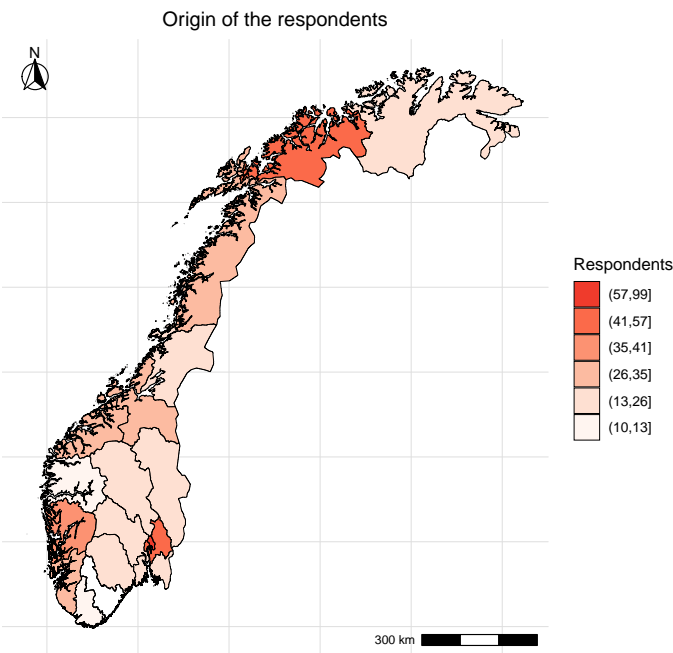
Figure 4.3: Survey 1 – origin of the respondents

4.3 Results

The table 4.4 shows general results of the pilot study for each of the nouns. The number of answers and the percentage is given next to each noun and each of the grammatical forms. For instance, BN of the first noun on the list was chosen by 589 respondents, indefinite singular by 5 and so on.

The first impression is that bare nouns are most frequent in generic contexts, whereas definite plural nouns are rather rare (cf. the analysis of the corpus data in chapter 5). The texts in the survey were not generic in the sense of Behrens (2005) (see chapter 2), but their function was to suggest the generic reading of the sentences with gaps to fill.

Noun / Answers (%)	Bare.Noun		Ind.Sg		Def.Sg		Ind.Pl		Def.Pl	
1- a strength training	589	(93.2)	5	(0.79)	24	(3.8)	13	(2.06)	1	(0.16)
2- an apple	132	(19.19)	191	(27.76)	3	(0.44)	362	(52.62)	0	(0)
3- an orange	374	(45.95)	19	(2.33)	46	(5.65)	373	(45.82)	2	(0.25)
4- a friend	0	(0)	203	(27.03)	0	(0)	530	(70.57)	18	(2.4)
5- a politician	0	(0)	56	(6.6)	17	(2)	447	(52.71)	328	(38.68)
6- a dog	64	(7.28)	93	(10.58)	379	(43.12)	293	(33.33)	50	(5.69)
7- a religion	547	(75.03)	47	(6.45)	112	(15.36)	14	(1.92)	9	(1.23)
8- a smartphone	272	(30.39)	98	(10.95)	392	(43.8)	103	(11.51)	30	(3.35)
9- a truck	0	(0)	494	(57.24)	26	(3.01)	266	(30.82)	77	(8.92)
10- a bicycle	360	(35.75)	131	(13.01)	300	(29.79)	177	(17.58)	39	(3.87)
(*)11- milk	555	(80.43)	0	(0)	135	(19.57)	-	(-)	-	(-)
(*)12- music	591	(94.41)	0	(0)	35	(5.59)	-	(-)	-	(-)
13- a child	245	(25.47)	21	(2.18)	54	(5.61)	331	(34.41)	311	(32.33)
14- a beer	78	(11.34)	2	(0.29)	508	(73.84)	82	(11.92)	18	(2.62)
(*)15- breadstuff	585	(99.83)	0	(0)	1	(0.17)	-	(-)	-	(-)
16- an audiobook	163	(17.3)	259	(27.49)	70	(7.43)	423	(44.9)	27	(2.87)
17- a backpack	433	(47.9)	92	(10.18)	203	(22.46)	146	(16.15)	30	(3.32)
18- a candle	329	(44.22)	26	(3.49)	2	(0.27)	382	(51.34)	5	(0.67)
19- a down jacket	133	(16.65)	457	(57.2)	1	(0.13)	207	(25.91)	1	(0.13)
20- a straw hat	210	(27.38)	520	(67.8)	4	(0.52)	29	(3.78)	4	(0.52)
21- a gut feeling	11	(1.68)	106	(16.23)	512	(78.41)	8	(1.23)	16	(2.45)
22- a green lamp	27	(3.46)	412	(52.82)	2	(0.26)	329	(42.18)	10	(1.28)
23- a white dress	389	(41.56)	85	(9.08)	142	(15.17)	261	(27.88)	59	(6.3)
24- a big touchscreen	309	(32.91)	194	(20.66)	75	(7.99)	278	(29.61)	83	(8.84)
25- a regular beer bottle	16	(1.86)	452	(52.56)	8	(0.93)	368	(42.79)	16	(1.86)
26- a white shirt	199	(25.42)	493	(62.96)	55	(7.02)	29	(3.7)	7	(0.89)
27- a ceramic stove	320	(34.08)	55	(5.86)	98	(10.44)	336	(35.78)	130	(13.84)
(*)28- gray hair	593	(98.34)	4	(0.66)	6	(1)	-	(-)	-	(-)
29- a red house	24	(3.16)	334	(43.95)	278	(36.58)	98	(12.89)	26	(3.42)
30- a white flag	223	(25.9)	476	(55.28)	78	(9.06)	81	(9.41)	3	(0.35)
Total	7771		5325		3566		5966		1300	
Total (%)	32.48%		22.25%		14.90%		24.93%		5.43%	
Countable (%)	25.43%		24.84%		15.82%		27.85%		6.07%	
(*) Uncountable (%)	92.77%		0.16%		7.07%		–		–	

Table 4.4: Survey 1 – results

The table 4.5 and the graph 4.4 show how the different NP types were distributed throughout the survey texts. Very low percentage of definite plural nouns contrasts with BNs which account for over 30% of all answers. However, the proportions differ slightly when we analyse countable and uncountable nouns separately.

Forms	Number	%
bare nouns	7771	32.48
indefinite sing.	5325	22.25
definite sing.	3566	14.90
indefinite pl.	5966	24.93
definite pl.	1300	5.43
Total	23928	100[a]

Table 4.5: Survey 1 – NP types

[a]The numbers given in the table are rounded to 0.01. The actual numbers sum up to 100%.
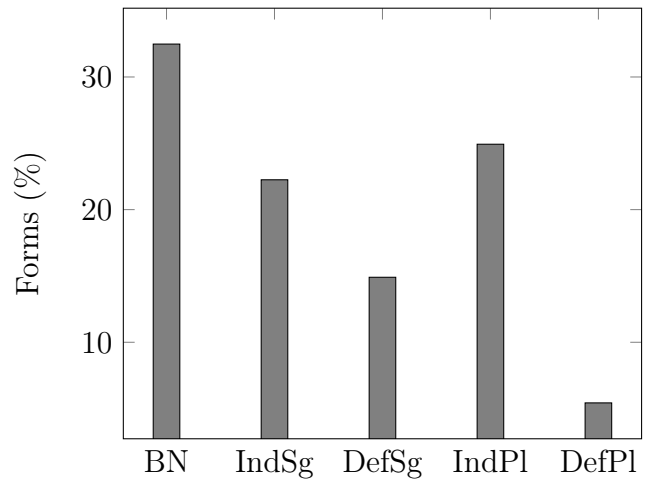
Figure 4.4: Survey 1 – NP types in generic contexts

4.3.1 Countable nouns – general results

The majority of the survey texts concerned countable nouns, as these show more variation when it comes to the use of the five NP types. 12 of all texts in the survey included whole phrases, such as 'a regular beer bottle' in line 25 in table 4.4. The other 14 countable nouns did not contain any modifiers. The general results for countable nouns can be seen in table 4.6 and graph 4.5.

Forms	Number	%
bare nouns	5447	25.43
indefinite sing.	5321	24.84
definite sing.	3389	15.82
indefinite pl.	5966	27.85
definite pl.	1300	6.07
Total	21423	100[a]

Table 4.6: Survey 1 – NP types, countable

[a]The numbers given in the table are rounded to 0.01. The actual numbers sum up to 100%.
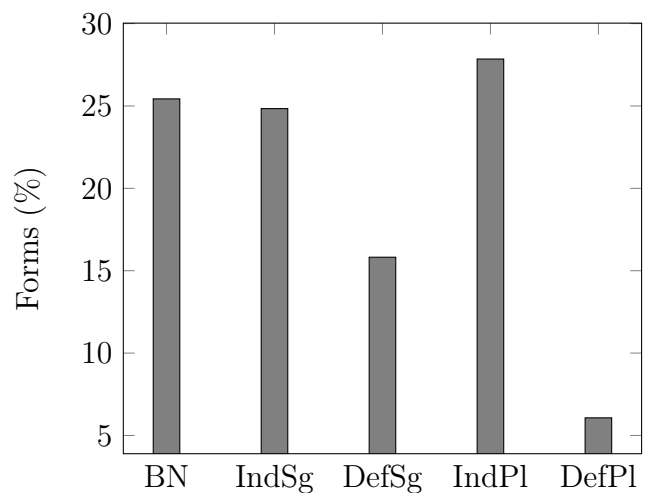
Figure 4.5: Survey 1 – countable, NP types

Three forms are surprisingly the most frequent ones, namely indefinite plural, BN and indefinite singular. With indefinite plural being a *default* generic form in many Germanic languages, including Scandinavian languages (see chapter 3 and e.g. Teleman et al. 1999 for Swedish), the survey has proved that this holds also for Norwegian. Indefinite singular is also considered very common e.g. in English generics, even though certain restrictions apply (see chapter 2).

What may be considered surprising is the high percentage of BNs in generic contexts. The distribution of BNs is much wider in Mainland Scandinavian languages than for instance in English and this was also observed in this pilot study, also when it comes to countable nouns. If we look at the results in the table 4.4 we will notice that BN was chosen as a correct form for numerous countable nouns, such as 'orange', 'religion' or 'backpack' among others.

(108) Stadig flere sliter med depresjon og angsttilstander som et resultat av stress. For mange kan plikter på jobb og i hverdagslivet være en hovedgrunn til økt stressnivå. Blant vanlige behandlinger og terapier, finner man også en del psykologer som anbefaler alternative behandlingsmetoder. For noen mennesker kan for eksempel a) religion b) en religion c) religionen d) religioner e) religionene være til hjelp.

More and more people suffer from depression and anxiety which are caused by stress. Too many obligations at work and in everyday life can be reasons for high levels of stress. Among standard cures and therapies, one can also find some psychologists who recommend alternative treatments. For instance, for some people a) religion b) a religion c) the religion d) religions e) the religions can prove helpful.

Certain nouns from the survey, such as 'religion' in (108), may be interpreted in two ways: as a concept and therefore a mass-like noun, or as a subtype, namely a particular religion (catholic, protestant etc.). In the example (108), the first interpretation seems more plausible. The use of certain countable nouns without articles, as for instance 'religion' from the example above, suggests that some nouns can be perceived as abstract or mass-like. The survey results show a high degree of acceptability of BNs, be it with countable or uncountable nouns.

4.3.1.1 Types of generic generalizations

As has been said in chapter 2, the material from the pilot study was analysed according to the cognitive model proposed by Leslie et al. (2011), where 6 main generalization types were given according to the predication type and truth-value. The predications used in such sentences are:

1. Quasi-definitional   true
2. Majority characteristic   true

3. Minority characteristic   true
4. Majority   true
5. Striking   true
6. False generalization   false

All predication types, except for false generalizations, are true in terms of truth-value of the generic. Quasi-definitional type does not allow for exceptions and is often used with kind-predicates. Majority characteristic is a quite common type of generic predication that concerns most members of a given kind but allows for exceptions, whereas minority generics render generalization over a whole kind by asserting it a feature characteristic only for the minority (e.g. laying eggs).

The predication type referring to the majority means that a given property is prevalent among the members of a given kind but it is not characteristic of the kind. For instance, we might state, that candles create nice atmosphere but it is not the main function or the distinctive feature of the object. Striking predication, on the other hand, occurs when a feature associated with a minority of the kind and is not very common among its members, is presented as generic. The difference between the minority characteristic and the striking predication is that the latter concerns a feature that is considered dangerous.[5]

The table 4.7 shows the classification of the nouns from the survey in terms of predications they occurred with. Each of the types will be discussed in greater detail in the following sections.

Type	NP	Number
Quasi-definitional	orange, truck, child, beer	4
Majority characteristic	strength training, apple, friend, dog, bicycle, green lamp, red house, white flag	8
Minority characteristic	politician, audiobook	2
Majority	religion, candle, down jacket	3
Striking	regular beer bottle	1
False generalization	smartphone, backpack, straw hat, gut feeling, white dress, big touch screen, white shirt, ceramic stove	8
Total		26

Table 4.7: Survey 1 – Types of generic generalizations, countable nouns

---

[5]The model was discussed in greater detail in section 2.4.

Quasi-definitional   Among the 26 countable nouns in the survey, four of them occurred in the texts with quasi-definitional predications. The text about trucks in (109) is an interesting example since none of the respondents chose BN among all NP types.

(109)   Antall bilulykker i Europa er på et relativt høyt nivå. Tusenvis av lastebiler som kjører gjennom kontinentet bidrar til stor trafikk på motorveier og i byer. For å kjøre trygt og unngå ulykker er det viktig å huske at **lastebil/en laste-bil/lastebilen/lastebiler/lastebilene** veier flere tonn og trenger derfor lang tid til å bremse.

The number of car accidents in Europe is on a relatively high level. Thousands of cars that drive through the continent contribute to the traffic on the highways and in the cities. In order to drive safe and avoid accidents one should remember that truck/a truck/the truck/trucks/the trucks weigh several tons and need(s) therefore longer time to brake.

The NP types used most often by the respondents in this example included the two indefinite forms (singular and plural), followed by the definite plural and definite singular. The most frequent NP type, chosen 494 times, suggests that 'a truck' was interpreted as a prototype. And since quasi-definitional generics do not allow for exceptions, a prototype with a given feature can be used when referring to the whole kind.

Interestingly enough, the definite plural form was preferred over the singular one in this example, proving once again that in Norwegian delimited generics are not limited to people, as is the case in English (cf. Radden 2009). The use of definite plural form can also suggest that trucks are understood as a subtype of a broader category 'vehicle'. This way the feature assigned to the subkind in the text applies to trucks only and not all vehicles in general.

In the texts about children and oranges, the choices were distributed between BN and indefinite plural. Additionally, the noun 'child' scored high also in the case of definite plural. The noun denotes people so that choice is not surprising. Another possible explanation may be that children, similarly as trucks, are perceived as subtypes of the general, bigger category. This dependency can be observed in a simple test: a subtype can also be described with the name of the main type, whereas the opposite is not possible. Namely, children are people but it is possible to say that people are children because that would be qualified as a false categorization.

The text about beer in (110),[6] even though falling into the same category as the previous nouns, gave slightly different results.

(110)   Det er blitt lettere å være entreprenør i Norge. Blant de nye prosjektene som støttes er det flere små bryggerier rundt omkring i landet. Nybegynnere i bransjen

---

[6]In Norwegian, 'beer' is used either as a general name for alcoholic drinks with the article *et* or in the meaning 'a glass of beer' with the article *en*. It was therefore clear that the example concerned beer in general, not a particular type or quantity of it.

ser det som en sjanse til å komme inn på ølmarkedet. For disse entreprenørene betyr **øl/et øl/ølet/øl (pl.)/ølene** de selger en liten, men fast inntekt.

It has become easier to be an entrepreneur in Norway. Among many projects that receive funding, there are many small breweries in the country. Beginners in the business see it as a chance to enter the beer market. For those entrepreneurs, beer/a beer/the beer/beers/the beers they sell mean a small but stable income.

The most frequently chosen NP type was definite singular, followed by indefinite plural and BN respectively. Nevertheless, the vast majority of the answers was given to the definite singular form suggesting that 'the beer' was perceived as a prototype. Similar, quasi-definitional generics were discussed by Pettersson (1976) in his account on definiteness and generics in Swedish. The analysed examples show that also in Norwegian prototypes can occur as definite singulars. However, BNs are also common in this function, as has been showed in previous examples.

Majority characteristic   Among all survey texts with countable nouns, 8 were classified as majority characteristic, namely generics allowing for exceptions but still referring to the majority of a given group. Such types of predications occurred in examples such as (111a) and (111b).

(111)   a. Det er mange faktorer som må tas hensyn til når man skal pusse opp et kontor. Gode og komfortable møbler må man selvfølgelig ha, men tilbehøret er også viktig. Flere undersøkelser har for eksempel vist at **grønn lampe/en grønn lampe/den grønne lampen/grønne lamper/de grønne lampene** hjelper til å fokusere bedre.

There are many factors that need to be considered when one wants to renovate an office. Good and comfortable furniture is a must but accessories are also important. Many studies have shown that for example green lamp/a green lamp/the green lamp/green lamps/the green lamps help(s) to focus better.

b. Norsk arkitektur og interiørdesign har gått gjennom flere endringer i de siste årene. Det som var typisk norsk for tjue år siden ikke nødvendigvis er det i dag. Men noen ting blir alltid knyttet til den norske kulturen som for eksempel /textbfrødt hus/et rødt hus/det røde huset/rød hus/de røde husene i skogen.

Norwegian architecture and interior design have changed much in recent years. What was typically Norwegian twenty years ago, is not necessarily today. But some things will always be connected to the Norwegian culture, as for example red house/a red house/the red house/red houses/the red houses in the forest.

The fact that green lamps might have a positive effect on people does not mean that any green lamp would have the same function. The example with red houses refers

to a noun phrase that could be considered as familiar or as a well-established kind – it is an element characteristic for the Norwegian culture and therefore recognisable for most people from this country. However, in certain regions of the country white wooden houses are considered typical, therefore this predication falls into the category majority characteristic.

Minority characteristic   Minority characteristic is the opposite of the previous predication category, namely it considers a feature shared by a minority of the population and projects it over the whole kind. Among the survey texts, only two texts were classified this way and those are the examples in (112).

(112)   a.   Verdenspolitikken har i det siste vært i krise. I mange land, blant annet Norge, minsker samfunnets tillit til politikere og offentlige organisasjoner. Flere og flere mennesker innrømmer at de ikke tror på det de store politiske partiene påstår. Bortsett fra det mener mange at **politiker/en politiker/politikeren/politikere/politikerne** har relativt stor makt i det moderne samfunnet.

Recently the international political situation has been in crisis. In many countries, also in Norway, society's trust to politicians and public institutions is deteriorating. Many people admit that they no longer believe what the big political parties claim. Apart from that, many say that politician/a politician/the politician/politicians/the politicians have relatively big power in the modern society.

b.   I en digitalisert verden der nesten alt foregår på nettet, står papirlitteraturen overfor den største krisen noensinne. Det er flere som velger å laste ned en app istedenfor å lese en bok i fritiden. Mange mener også at vanlige papirbøker er i ferd med å forsvinne helt fra markedet. I stedet kan imidlertid **lydbok/en lydbok/lydboken/lydbøker/lydbøkene** være en god litterær opplevelse.

In a digitalised world where almost everything happens on the internet, paper books face the biggest crises ever. Many people choose to download an app instead of reading a book in free time. Many people say also that regular books will copletely disappear from the market. Instead, audiobook/an audiobook/the audiobook/audiobooks/the audiobooks can be a good literary experience.

Minority generics may be difficult to recognise as such generalisations are often widely accepted, such as the example *Ducks lay eggs* provided by Leslie et al. (2011). In the survey, the texts identified as minority characteristic concerned the nouns 'politician' and 'audiobook'. The predications refer to minorities – in fact only a small number of politicians have real power and good literary experience can be also provided by e-books,

regular paper books or poetry readings.

Majority Majority predication, similarly as majority characteristic, also refers to a feature held by the majority of the population but the feature in question is not considered distinctive for the kind. An example of such predication can be observed in sentence (113), where the feature of candles is projected on the majority of the kind but is not essential.

(113) I høstmånedene er det mange som velger å kose seg hjemme om kvelden med et glass vin og en god bok. Lurer du på hvordan du kan skape den koselige høststemningen i ditt eget hjem? Det er relativt enkelt og krever ikke veldig mye tid. Først og fremst er det fint med **stearinlys/et stearinlys/stearinlyset/ stearinlys[pl.]/stearinlysene** hjemme.

In autumn months many people choose to stay cosy at home in the evening, with a glass of wine and a good book. Are you wondering how you could create that cosy autumn atmosphere in your own house? It is relatively easy and does not take much time. First of all, it is a good idea to have candle/a candle/the candle/candles/the candles at home.

The idea that candles create a cosy atmosphere is a popular opinion, especially in the Scandinavian context where such notions as *kos* and *hygge* are popular. However, the main characteristic of a candle is to provide light so the statement about its cosiness is very objective. Also, a candle that does not create a cosy atmosphere still fulfils all the essential criteria needed for it to be recognised as a candle (being made of certain material giving light etc.).

The majority of the answers in the text about candles were BNs and indefinite singular, again conforming to the main tendencies in the whole survey where those two forms were used most often. What is more, majority predications occur often in generic sentences where generalisations are made in reference to non-essential features of a kind, so such choice of the forms is not surprising. Referring to a prototype can be expressed with a BN, whereas kind-reference in the meaning 'the majority' or 'most' is often rendered with indefinite plural.

Striking Striking generics are based on predicates that are considered shocking and not typical of a given kind. In the pilot study, the text about a regular beer bottle was considered such predication, as can be observed in (114).

(114) Interiørdesign endrer seg stadig og kan noen ganger bli litt overraskende. Moderne designere bruker ofte gamle kopper og keramikk til å pynte kjøkkenet og stuen. Man kan for eksempel bruke **vanlig ølflaske/ en vanlig ølflaske/ den vanlige ølflasken/vanlige ølflasker/de vanlige ølflaskene** til å pynte kjøkkenet.

Interior design is changing constantly and can sometimes be surprising. Modern

designers often use old mugs and ceramic to decorate kitchens and living rooms. One can for instance use regular beer bottle/a regular beer bootle/the regular beer bottle/regular beer bottle/the regular beer bottles to decorate the kitchen.

An unusual and shocking use of an everyday object, often considered garbage, falls into the category of striking predications. Such predications refer to a minority of the kind and render improbable and/or surprising characteristics. It rather safe to say that beer bottles are not usually considered good decorations.

Among the NP types chosen in this text, indefinite singular and indefinite plural were most common. The adjective in the NP certainly influences the more frequent use of the indefinite than BN, as compared to the examples with nouns only. The indefinite plural form, second most chosen option, expresses reference to some beer bottles that represent the mentioned kind.

False generalization False generalisations are the only predicates in the model of Leslie et al. (2011) that have false values. This means that such generalisations are perceived as generic but they lack the truth-value, as can be seen in the examples in (115).

(115) a. Kjøkkenutstyr er noe alle må velge en gang i livet. Kommer du til å møblere ditt eget kjøkken snart, er det lurt å sammenligne forskjellige typer komfyr før du bestemmer deg. Blant alle de tilgjengelige komfyrene er det **keramisk komfyr/en keramisk komfyr/den keramiske komfyren/keramiske komfyrer/de keramiske komfyrene** som anbefales oftest.

   Kitchen equpiment is something everyone needs to choose once in life. If you are going to furnish your kitchen soon, it is smart to compare different types of stoves before you make a decision. Among all available stoves, ceramic stove/a ceramic stove/the ceramic stove/ceramic stoves/the ceramic stoves is/are recommended most often.

   b. I dagens verden er det viktig å være online døgnet rundt. Dette gjør at telefonprodusentene kommer med stadig nye funksjoner i produktene sine. Ifølge brukerne er det i dag bare **stor trykkskjerm/en stor trykkskjerm/den store trykkskjermen/store trykkskjermer/de store strykkskjermene** som teller.

   In today's world it is important to be online 24h a day. Because of that the producers invent new functions in their products. According to the users, nowadays only big touchscreen/a big touchscreen/the big touchscreen/big touchscreens/the big touchscreens counts/count.

Interpreting a given predication as a false generalisation requires that there's a corresponding counterpart that is of equal importance. For instance, the fact that ceramic stoves are recommended often is a generalisation based on current trends from a certain

place and time. There are many stove types that constitute a counterpart for ceramic stoves and that could be used when referring to current trends in other parts of the world.

The example about phones with big touchscreens also represents one aspect of mobile technology that can be easily contrasted with other important features of such devices. Even though it is believed that big touchscreens are essential nowadays, such a generalisation is false because a key feature of every mobile device is its ability to connect to different networks. However, this feature is taken for granted and often overlooked leading to such statements as seen in the example (115b).

When it comes to NP types used in the case of false generalisations, BNs and indefinite plurals dominated in 3 of the survey texts from this category ('white dress', 'big touchscreen' and 'ceramic stove'). In the text about the NP 'white shirt', indefinite singular form was chosen most often, followed by BN.

4.3.2 Uncountable nouns – general results

Uncountable nouns, even though a minority of all NPs in the survey, also show some diversity (see table 4.8 and graph 4.6). In the text about gray hair (line 28 in table 4.4) some respondents chose the indefinite NP type ('a gray hair'). However, such a low percentage (0.16%) does not really influence the general tendencies for mass nouns.

Forms	Number	%
bare nouns	2324	92.77
indefinite sing.	4	0.16
definite sing.	177	7.07
Total	2505	100

Table 4.8: Survey 1 – NP types, uncountable
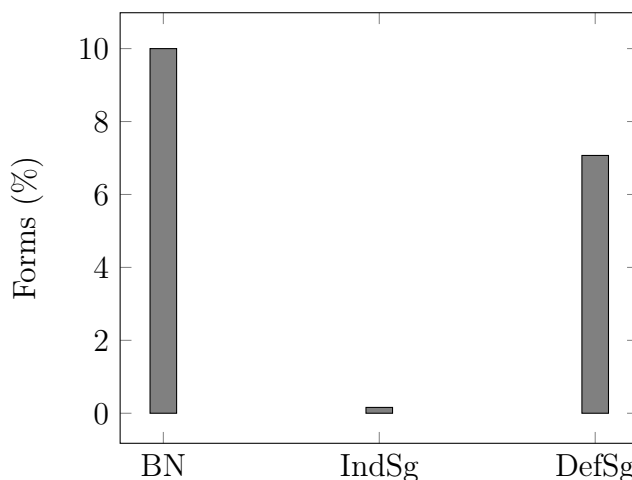
Figure 4.6: Survey 1 – uncountable, NP types

Let us look at the table 4.4 again. The nouns in 11, 12 and 15 follow a general pattern similar for all Germanic languages, namely they occur either as BNs or definite nouns. The noun phrase 'gray hair' however, was chosen 4 times with indefinite article. Only in 6 answers the definite form was chosen, whereas the majority opted for a bare noun. As has been said, all participants could choose more than one answer in each of the survey questions. Therefore the numbers of answers for each of the NP types in all test items are varied. What is more, due to the number of participants it is not possible

to analyse the answers of each and every one of them separately. In the analysis, main tendencies are discussed, as they give an overview of the phenomenon and this was the main point of this pilot study.

Even though the indefinite form accounts for less than 1% of all answers in the example (116) below, it is nevertheless an interesting case. It is rather unlikely that the respondents have chosen this answer by a mistake but this possibility cannot be ruled out completely due to the design of the survey. If we assume that the choice of indefinite singular was intentional, another question arises: what interpretation could such a sentence have? Apart from a specific reference to one particular strand of hair,[7] 'a gray hair' can also be understood as a prototype. The first gray hair symbolises certain age or the ageing process in general. Since the text in the example explicitly suggested ageing, the latter interpretation seems plausible. Interpreting 'a gray hair' as a prototype implies that the phrase may be understood as *familiar* by certain speakers (Borthen 2007, see also section 2.2.1.1).

(116)   Det blir stadig vanligere å endre litt på utseendet når man blir eldre. I dag tilbys plastiske operasjoner og ulike kurer som skal utsette aldringsprosessen. Mange velger friske farger når de begynner å få a) grått hår; b) et grått hår; c) det grå håret.

It is more and more and more common to modify one's image as one gets older. Nowadays, there are many plastic surgeries and different cures that can delay the ageing process. Many people choose hair dyes in vibrant colours when they start to get a) gray hair; b) a gray hair; c) the gray hair.

An argument against such a reading of the example is the presence of the aspectual verb 'to begin'. It suggests a whole process and not a certain point in time. Nevertheless, 'a gray hair' that symbolises ageing does not deny the fact that ageing is a process. Quite the opposite – 'a gray hair' is the first sign of the process beginning.

The other survey texts with uncountable nouns did not give surprising results as all respondents opted either for BNs, the definite form or both. BNs dominated in the case of all mass nouns, which was expected.

4.3.2.1   Types of generic generalizations

Among the uncountable nouns in the survey were: 'milk', 'breadstuff', 'gray hair' and 'music'. The texts about the first three nouns are written in a way that suggests the generic meaning of the last sentence which, when interpreted in isolation, could also have a specific reading. The generic sentence in the text bout music could be interpreted as generic also in isolation.

---

[7]The indefinite article *en* ('a') developed from the numeral 'one' and in Modern Norwegian both the article and the number are homophones.

Type	NP	Number
Quasi-definitional	–	0
Majority characteristic	gray hair	1
Minority characteristic	–	0
Majority	music, bread	2
Striking	milk	1
False generalization	–	0
Total		4

Table 4.9: Survey 1 – types of generic generalizations, uncountable nouns

Among the four nouns in this category, there were three different types of generic generalisations: majority characteristic, majority generics that occurred twice, and striking generics.

Majority characteristic   The example in (117) shows majority characteristic, namely gray hair is a biological sign of ageing but certain exceptions occur. For instance, some people might lose their hair early or they may get gray hair at an early age. Majority characteristic allows for such cases, being anomalies to the general tendency.

(117)   Det blir stadig vanligere å endre litt på utseendet når man blir eldre. I dag tilbys plastiske operasjoner og ulike kurer som skal utsette aldringsprosessen. Mange velger friske farger når de begynner å få **grått hår/et grått hår/det grå håret**.

   It is more and more common to change one's appearance when one gets older. Nowadays, there are many plastic operations and different treatments that can delay the ageing process. Many people choose bright colours when they begin to get gray hair/a gray hair/the gray hair.

Majority generics   The text in (118a) presents majority generics, where a certain behaviour (here: having bread as the main food in a cabin) is assigned to the majority of the members (here: people who have cabins/go on cabin trips). Such property, even though associated with most members of a kind, does not characterise the whole kind Leslie et al. (2011: 19).

(118)   a. Det er typisk norsk å dra på hyttetur. Uansett om du reiser alene eller med venner må du ikke glemme det som er viktigst på slike turer, nemlig maten! Taco og varme retter kan være fristende, men det er også greit med noe mer tradisjonelt. Derfor bør man alltid ha **brødmat/en brødmat/ brødmaten** på hytta si.

It is typically Norwegian to go on cabin trips. No matter whether you travel alone or with friends, you mustn't forget what's the most important in such trips, namelig food! Taco and warm dishes can seem tempting but something more traditional is also a good idea. Therefore one should always have breadstuff/a breadstuff/the breadstuff in their cabin.

b. Autistiske barn har det vanskelig med å tilegne seg grunnleggende kunnskaper som lesing og skriving. Gjennom årene har man jobbet med nye opplæringsteknikker og verktøy som kan brukes på skolen. Det har blitt bl.a. påvist at **musikk/en musikk/musikken** hjelper barn med autisme til å fokusere bedre på skolen.

Autistic children struggle with learning new skills such as reading and writing. For years people have been working on new teaching strategies and tools that could be used in schools. For instance, it has been proved that music/a music/the music helps autistic children to focus better at school.

In the text about music in (118b), the generic sentence with a gap is unambiguous and it could be interpreted generically also without a broader context. According to the classification proposed by Leslie et al. (2011: 19), the text presents majority generics. It means that a given property (here: the property of music to help people concentrate) is prevalent in a kind but it cannot be perceived as a principled connection for the whole kind. There may exist certain music genres that do not help people concentrate.

Striking generics   The text about milk is an example of striking generics Leslie et al. (2011: 19), where a quality that is commonly believed to be true of the whole kind (here: dairy strengthens the bones) is presented in a negative light in reference to certain members of the kind. Similarly as in the example provided by Leslie et al. ('Pit bulls maul children'; meaning: not all dogs and not always maul children), also here milk is only one type of dairy that may cause osteoporosis, whereas the same would not be stated abut cheese for instance. The feature of causing the bone disease is shocking and dangerous, as was classified by Leslie et al..

(119)   Bør man spise meieriprodukter for å få i seg nok kalsium? Og kan kalsium fra meieriprodukter styrke bein i kroppen? De nyeste undersøkelsene viser at i land der det spises mest meieriprodukter forekommer benskjørhet mye oftere enn vanlig. Er det faktisk slik at melk/en melk/melken forårsaker dette?

Should one eat dairy in order to get enough calcium? And can it be that calcium from dairy strengthens the bones? The latest studies show that in the countries where most dairy is consumed, osteoporosis is much more common. Can it in fact be that **milk/a milk/the milk** causes it?

The examples discussed above show that the classification provided by Leslie et al. allowed to divide the texts into categories based on the predications they occurred with. All of the predication types were represented in the survey, allowing for a relatively big diversity of texts in the pilot study.

Respondents' comments   As has been mentioned before, the participants could add comments to each of the survey texts. However, the majority of the comments did not concern the survey subject but for instance the use of commas.[8] This also means that the subject of the survey was unknown to the respondents and therefore their feedback concerned mostly spelling and general language use in the presented texts.

The very few comments that proved useful for the survey were the ones in which the participants listed other possible answers, without marking them in the test items. Even though it was possible to mark more than one answer in each of the survey questions, certain participants chose only one NP type and listed the other ones in the comment section. Such answers were manually corrected in order to count all of the possible NP types in the final results.

This pilot study was designed this way in order to obtain more detailed feedback on the tested phenomenon. Unfortunately, the chosen strategy did not prove successful and was therefore changed in the main survey of this project. Despite the lack of comments from the respondents in this survey, the results still show main tendencies in choice of different NP types in generic contexts.

4.3.3   Statistical analysis

The data collected in the pilot study was analysed statistically in order to test for significance of the results. First, all survey nouns were analysed with regard to data distribution, which can be seen in table 4.10. Minimal and maximal values for each of the forms were calculated, as well as median and mean.

The minimal value of 0 for definite singular means that there was at least one survey question where none of the respondents chose that option. If we look at the table 4.4 again, we'll see that in the text with the noun 'friend', none of the respondents chose the definite singular form. Comparing the minimal values of all NP types shows that the two definite forms were a lot less frequent than BNs and indefinite plurals.

The maximal values, as the name implies, show the maximal number of the answers for each of the NP types in all of the test items. For instance, the maximal value of 593 for BNs means that in at least one of the survey texts 593 respondents chose this NP type. As can be seen in the table 4.4, 593 people chose BN in the text about gray hair,

---

[8]All of the texts were proofread and edited by a native speaker of Norwegian so the comments concerning the use of commas or definite/indefinite NP types could be a result of variation within the *bokmål* standard itself.

when both countable and uncountable results are considered.

The division of the data into quartiles shows the distribution of each of the NP types throughout the survey texts. When it comes to plural forms, indefinite and definite, they were not available in 4 texts where mass nouns were used. These are marked as 'NA' in the table 4.10.

	BN	**IndSg**	**DefSg**	**IndPl**	**DefPl**
Min.	0.00	0.00	0.00	8.00	0.00
1st Qu.	67.50	22.25	6.50	86.00	5.50
Median	234.00	99.50	54.50	263.50	18.00
Mean	259.00	177.50	118.90	229.50	50.00
3rd Qu.	385.20	315.25	140.20	355.50	47.25
Max.	593.00	520.00	512.00	530.00	328.00
				NA's: 4	NA's: 4

Table 4.10: Survey 1 – descriptive statistics

Figure 4.7 shows a graphic representation of the data. Bare nouns, the most frequently chosen in the survey, have the biggest range between the minimal and the maximal values, followed by indefinite singular and indefinite plural. No outliers were observed with these three NP types, due to the wide range of values.

Outliers, presented as small circles on the boxplot, represent answers that were out of the main tendency. For instance, definite plural has a rather narrow range of answers (relatively low minimal and maximal values compared to the other NP types, with a few outliers. The division of the data into quartiles show low values in Q1 (5.50), Median [Q2] (18.00) and Q3 (47.25), compared to BN, IndSg and IndPl.

Similarly as in the case of definite plural, also the definite singular has a narrow range of answers (the minimal vs. the maximal value). It is slightly bigger than with DefPl but the division of the data into quartiles still shows low numbers compared to BNs, IndSg and IndPl (Q1 of 6.50, Median [Q2] of 54.40 and Q3 of 140.20).

The distribution of the data and descriptive statistics let us see which of the NP types were chosen most often by the respondents and how they compare between each other. Already such a simple comparison of the data shows where the differences between the forms are distributed. However, in order to analyse the collected material in even more detail, a set of statistical tests was conducted. The methodology of all tests utilised in this project was presented in chapter 1. In empirical chapters, only results and their interpretation will be presented.

The next step of analysing the data from the pilot study, consisted in verifying whether analysis of variance can be performed. In order to meet the requirements of ANOVA, the homogeneity of variances and normality of data need to be tested.

Levene's test was utilised to check homogeneity of variances and it resulted in p-value
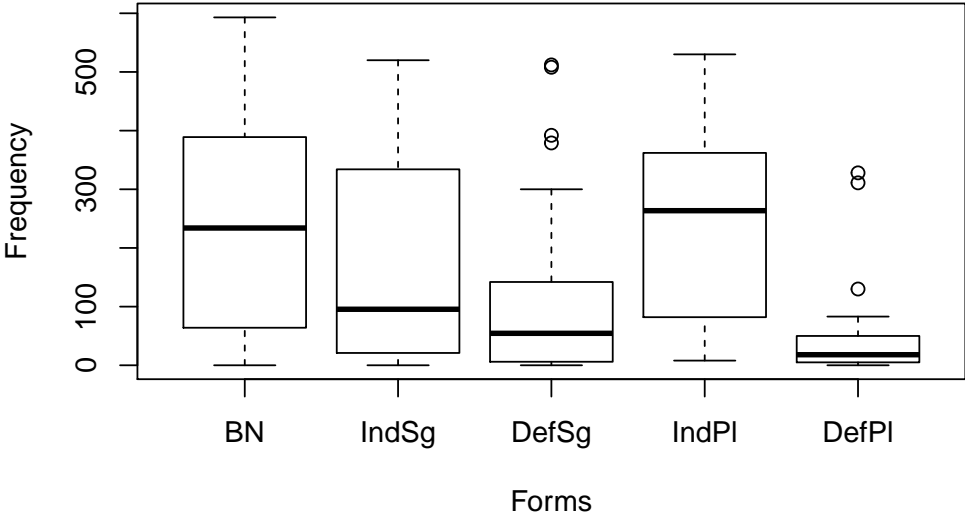
Figure 4.7: Survey 1 – boxplot

	Df	F value	Pr($>$F)	
Group	4	5.30	5.32e-04	***
	137			

Table 4.11: Survey 1 – Levene's test

under 0.05. The significance of this test is at 0.001 level, therefore a parametric test such as ANOVA could not be performed. The *p*-value of 6.68e-06 in Shapiro-Wilk test also confirmed that the data is not normally distributed which again suggested using non-parametric tests, such as Kruskal-Wallis.

Kruskal-Wallis resulted in *p*-value of 3.87e-05 and proved to be statistically significant, suggesting that there are differences between the tested groups (i.e. NP types). In order to locate the differences, Dunn's test was performed (see table 4.12).

mean comparison	BN	IndSg	DefSg	IndPl
IndSg	1.45			
p-adjust	1.00			
DefSg	2.69	1.24		
p-adjust	0.07	1.00		
IndPl	-0.11	-1.51	-2.71	
p-adjust	1.00	1.00	0.07	
DefPl	4.17	2.78	1.58	4.14
p-adjust	**0.00**	0.054	1.00	**0.00**

Table 4.12: Survey 1 – Dunn's test

Values marked in bold in table 4.12 designate differences between the groups. Significant differences were observed in two pairs of NP types: 1) BN – definite plural and 2) indefinite plural – definite plural. The differences between the mentioned NP types indicate that BNs were used in different texts (contexts) than definite plural nouns. When it comes to the difference between the two plural forms, indefinite and definite, one can expect that indefinite plural did function as a *default* kind-reference, whereas definite plural was considered as delimited generics. If we look at the table 4.4 again, we will see in which texts the definite forms appeared most often. The topic of those texts were 'politician', 'child' and 'ceramic stove'. The first two nouns, namely 'politician' and 'child' are in accordance with the theory of Radden (2009), where delimited plural generics are connected only with people. The example with 'ceramic stove' proves that in Norwegian, delimited generics may be extended also to *familiar* everyday objects, such as the stove mentioned in the example. Nevertheless, the data from this pilot study does not allow for such conclusions and only suggests that the statistically significant differences between the forms may in fact be a result of differences in interpretation of the examples.

Countable nouns Let us now look at countable nouns only, which constituted the majority of the survey material. As has been stated before, mass nouns do not exhibit such a diversity when it comes to NP types used in generic contexts so the main focus of this study remained on countable nouns. However, certain nouns in the survey could be perceived both as countable and mass nouns (e.g. 'beer' or 'strength training'). These were presented in the survey texts as countable nouns with a possible interpretation as mass nouns, allowing for a broader contexts in those texts.

Table 4.13 shows the distribution of the data in questions concerning countable nouns. Minimal and maximal values were counted and the material was also divided in quartiles. Since there were only 4 mass nouns in the survey, the results for countable nouns are not very different from the general results presented above. However, certain values did change and this is particularly visible in the case of the maximal value for indefinite singulars, which was expected since this form is very rarely used with mass nouns.

	BN	IndSg	DefSg	IndPl	DefPl
Min.	0.00	2.00	0.00	8.00	0.00
1st Qu.	36.25	55.25	10.25	86.00	5.50
Median	204.50	118.50	62.50	263.50	18.00
Mean	209.50	204.65	130.35	229.50	50.00
3rd Qu.	326.75	392.50	187.75	355.50	47.25
Max.	589.00	520.00	512.00	530.00	328.00

Table 4.13: Survey 1 – descriptive statistics, countable nouns

Boxplot 4.8 presents the distribution of each of the NP types. We can observe that only in the case of definite singular and plural forms there are some outliers, whereas other NP types are rather homogenous. Three NP types have a very wide range of values, namely BN, indefinite singular and indefinite plural, which is the reason for the lack of outliers. The forms with a smaller range of values, namely the two definite forms, are the only ones with answers that are out of the main tendency.

When it comes to the indefinite plural form (the *default* generic form), the minimal value of 8 is the highest among all NP types. Also in the quartiles, the values are relatively high, ranging from 86.00 for Q1, through 263.50 for the median and 355.00 for Q3.
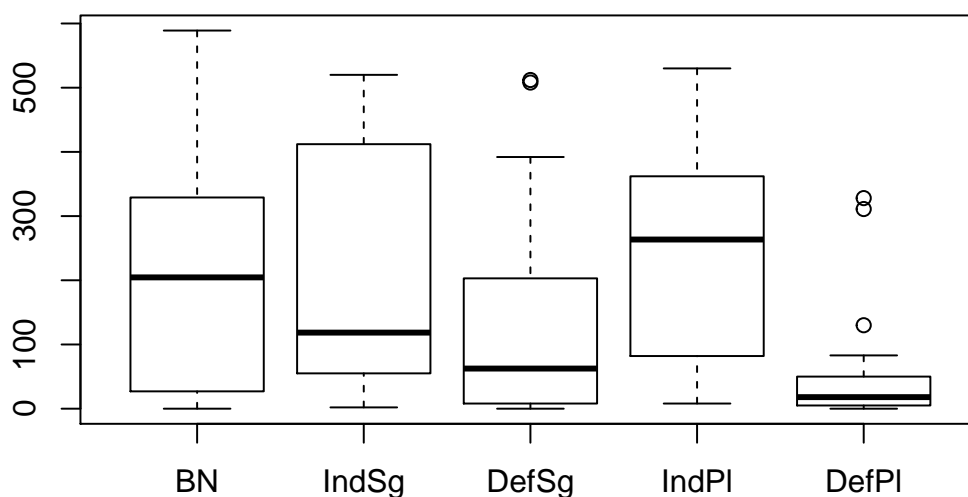
Figure 4.8: Survey 1 – boxplot, countable nouns

Further analysis of countable nouns was similar as with all nouns, namely homogeneity of variances and normality of data were checked. The assumption of normal distribution of the data was not met, therefore Kruskal-Wallis test was conducted, resulting in p-value below 0.05. This indicates that there are differences between the groups of countable nouns.

The post hoc test indicated where the differences lie (see table 4.14). Main differences were observed in the following pairs of grammatical forms: 1) BN – definite plural, 2) indefinite singular – definite plural and 3) indefinite plural – definite plural.

In each of the pairs, delimited generics (definite plural) were different from other NP types that are often used in generic contexts, namely BNs and the two indefinite forms – singular and plural. These differences are in accordance with the data presented in the part on general statistics above.

mean comparison	BN	IndSg	DefSg	IndPl
IndSg	-0.21			
p-adjust	1.00			
DefSg	1.68	1.89		
p-adjust	0.94	0.59		
IndPl	-0.81	-0.60	-2.49	
p-adjust	1.00	1.00	0.13	
DefPl	3.55	3.76	1.87	4.36
p-adjust	**0.004**	**0.002**	0.61	**0.00**

Table 4.14: Survey 1 – Dunn's test, countable nouns

Uncountable nouns The data concerning uncountable nouns from the survey was separated and analysed in a similar manner as countable nouns above. Due to the fact that mass nouns occur usually as BNs or as definite nouns, little diversity in the data was observed, as is shown in the table tab:survey1-uncountable and the boxplot 4.9.

	BN	IndSg	DefSg
Min.	555.00	0.00	1.00
1st Qu.	577.50	0.00	4.75
Median	588.00	0.00	20.50
Mean	581.00	1.00	44.25
3rd Qu.	591.50	1.00	60.00
Max.	593.00	4.00	135.00

Table 4.15: Survey 1 – Descriptive statistics, uncountable nouns

The minimal and maximal values of the data were estimated, for each of the three NP types. The values in the table 4.15 show that there are major differences especially in the use of indefinite singular – its minimal value falls way below the maximal one, resulting in a wide range that can also be observed in the boxplot 4.9. The very few answers with indefinite singular occurred in the text about 'gray hair'. As has been discussed before, such an interpretation would suggest that the respondents understood it as 'the first gray hair' which symbolises ageing. However, those answers were in the minority and in the other three texts about 'milk', 'music' and 'breadstuff', only BNs and the definite form were used.

The distribution of the data in the quartiles was also relatively consistent, due to the very limited diversity of the data. The values of Q1, Median and Q3 lie rather close in the case of BN (555.0, 588.0 and 591.50 respectively), whereas the definite form shows slightly bigger differences between the values with Q1 of 1.00, Median of 20.50 and Q3 of 60.00, with the maximal value of 135.00.

Further statistical analyses of the uncountable nouns from the survey were not possible, due to the limited amount of data. What is more, mass nouns in generic contexts
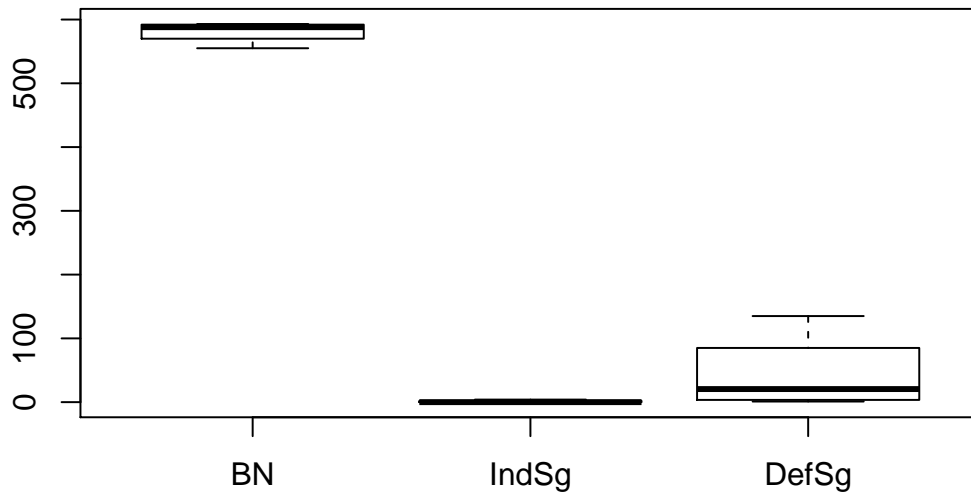
Figure 4.9: Survey 1 – boxplot, uncountable nouns

do not show such a high degree of diversity as countable nouns, which can be observed already in the table 4.15 – indefinite form has very low values, whereas BN and definite form have much higher values. Also, in the case of BN, both minimal and maximal values are very high which implies that the form is in a way *default* when it comes to generics of mass nouns.

4.4 Conclusions

The results collected in this pilot research have shown that the all five NP types can indeed be used generically in Norwegian but the forms are not always freely interchangeable. The general results indicate that BN, indefinite plural and indefinite singular are the NP types most often chosen to render generic meaning. When it comes to mass nouns, as expected BN and the definite forms are preferred with a very limited use of the indefinite form.

The results were interpreted with the use of the cognitive model proposed by Leslie et al. (2011), where different predication types used in generic texts were described. In the pilot survey, all predication types were present and the distribution of the NP types in each of the groups was analysed. Similarly as in the case of general results, also the cognitive analysis has shown that BN and indefinite plurals can be considered *default* forms in the case of Norwegian generics. However, high percentage of other forms chosen in a few of the survey texts indicates that other factors influence the choice of the forms. These factors can be texts' topics, the predications present in the texts, as well as the predicates in individual sentences and interpretation of the NPs (familiar or WEK). These

aspects will be analysed further in the following empirical parts of this project.

The statistical analyses performed in this study also show that there are significant differences between certain NP types. The analyses were performed in two ways: first on all of the survey texts and then on the texts with countable nouns only. Mass nouns constituted only a small part of the material as their distribution is rather regular in generic context. Therefore only simple descriptive statistics were performed on mass nouns.

5 Corpus data

5.1 The choice of the source texts

As has been said in the previous chapters, genericity is a language phenomenon that can
be challenging to define. One does not only need to decide whether or not a given noun
phrase, a sentence or a text is generic or not, but also the way of expressing genericity
differs depending on the context, text type and sometimes the noun itself. While studying
genericity, the choice of research material can be a challenge and it can influence the results
if not done properly.

Whereas most of the existing studies rely on sentence analysis where the studied
material consists of sentences created for the sake of research, there is a number of works
that include corpus research.[1] The choice of this type of material, namely longer texts
instead of individual sentences, has obvious advantages. One can not only observe how
the phenomenon functions in a language but also how a broader context influences the
way genericity is expressed. Certain drawbacks of this method should also be mentioned:
How to choose a corpus/corpora for the study? How to manage the data? Can a given
corpus be downloaded and tagged manually for genericity? What genres should be taken
into account? In this section I shall try to answer those questions and present the reasons
for choosing to build a specialised corpus of generic texts.

The choice of the study material when it comes to the Norwegian language was one
of the biggest challenges in the project. Norwegian has two written standards, *bokmål* and
nynorsk[2], the first one being used by the majority of the Norwegian population. The two
variants of the language are equal and have the same status in politics, culture, education
and media. What is more, the Norwegian government has introduced many programs in
order to maintain the two languages and to ensure that both are present in the public
domain. Choosing only one of the variants for analysis can therefore be controversial.
There are however reasons for such choice of the research material.

First of all, as has been mentioned above, *bokmål* is a widely used language variant

---

[1]For further discussion see chapters 2 and 3.

[2]Spoken language will not be considered in this dissertation. It is only worth mentioning that there
is no official standard of spoken Norwegian and that the majority of the population uses one or more of
many Norwegian dialects.

and a vast majority of the texts are written in it. Second of all, the two languages are not very distant from each other when it comes to grammar and vocabulary. Main differences between *bokmål* and *nynorsk* concern morphology and not necessarily sentence structure (cf. Faarlund 1988). Nevertheless, one cannot assume that genericity functions in exactly same way in both variants since *bokmål* belongs to Eastern Scandinavian languages (closely related to Danish) whereas *nynorsk* shares many features with Western Scandinavian languages. Choosing the more common variant of the language maximises the chances of describing main tendencies in the language, while still leaving a possibility to repeat the study on *nynorsk* later.

Once the dilemma concerning the written standard of the language has been solved (or at least diminished), the choice of corpus material needs to be considered. There exist a number of Norwegian corpora that are either freely available or that have limited access to language researchers, for instance: 1) Oslo-korpuset, 2) norsk aviskorpus or 3) NoWaC. The majority of the Norwegian corpora comprise texts in *bokmål* but it is of course possible to find texts written in *nynorsk*.

One of the main challenges with already existing corpora is data management. Genericity is not tagged in any of the corpora, neither Norwegian, nor English nor in any other language, since it is a context-dependent phenomenon. Tagging semantic data can be automatised only to a certain degree and even manual tagging does not guarantee success in 100% of the cases (some of the generic sentences can be ambiguous and therefore tagging them as generic or not, depends heavily on language intuition of the person performing the task of tagging).

What is more, none of the corpora that could have been used for the study on genericity was accessible for downloading. Editing a corpus and tagging texts for genericity was therefore impossible this way. Given the resources, scope of the PhD project and available tools, I have decided not to rely on the online corpora and to create a specialised data set that could be sorted, stored and tagged with the use of Open Source programs and packages.

Specialised corpora are rather small, comprising fewer words than general corpora which can contain billions of words. A specialised corpus is built for a given purpose, it can for instance contain texts on a given topic, of a given genre or focusing on a given aspect of the language.

A specialised corpus can be annotated and analysed *by hand and eye* as it contains a smaller amount of data and is much more manageable than general corpora (McEnery and Hardie 2012: 3). What is more, one can test hypotheses and study phenomena that cannot be automatically annotated. Genericity is a type of language feature that cannot (so far) be automatically recognised and tagged, as it requires understanding of the text and not only the ability to recognise a limited number of grammatical forms or parts of speech.

In order for the study to be representative, one needs to consider text genres that should appear in the corpus. Even though there were studies conducted on works of literature (see e.g. Behrens 2005 for the account on genericity from a cross-linguistic perspective based on translations of 'The Little Prince' by Antoine de Saint-Exupéry), finding texts with many generic references in it is a hopeless task. First of all, the present study concerns only the Norwegian language so repeating Behrens' study on the Norwegian translation of 'The Little Prince' would not show how genericity is expressed in everyday language. What is more, if we considered several literary works, even in the form of ebooks, data management would be a challenging task. Popular ebook formats, such as Mobi and Epub, are not easily converted into .txt format that could then be uploaded into Open Source tools.

Should we eliminate *belles-lettres*, another questions arises: why not use newspaper texts or Internet texts? There are tools that enable automatic collection of Internet texts (that is how the NoWaC corpus was built) but controlling the type of texts collected can prove challenging and is rather a task for a project in IT than in linguistics. Manual copying of newspaper articles available online is also a time-consuming task that may in addition not provide the desired amount of data.

Considering all of those aspects, I have decided to look for generic texts in a place that for sure contains a lot of them. Similarly as Carlsson (2012), I have opted for specialised literature, namely encyclopaedic texts. Carlsson analysed 36 texts about animals, retrieved from two books, whereas the idea behind the corpus created for this project was to collect texts in few different categories, not only botanical or biological ones.

What Carlsson's study shows is that it is fully possible to test the functioning of a given language phenomenon with the use of a rather small, specialised data set. Her analysis is descriptive and focuses mainly on text and discourse analysis[3], whereas the study conducted for this project is both quantitative and qualitative, as shall be discussed in section 5.3 of this chapter. Such a combination of methods and a variety of topics in each of the corpus texts makes the study innovative, even if the collected data is rather homogenous.

Since genericity expresses generalisations and truths about kinds or groups of objects, the texts in the corpus needed to be scientific and e.g. describe the history and features of a given tool, piece of furniture or even tasks of a given profession. Therefore the collected data was divided into five main categories. All of the categories and technical details of the corpus will be described in section 5.2 below. Let us now briefly consider the advantages and drawbacks behind the decision to base the whole corpus on encyclopaedic texts only.

The first and most obvious advantage of an encyclopaedia (even in its online edition, as it is the case with *Store norske leksikon* [SNL] utilised in this project) is that the texts

---

[3]For further discussion on the study, see section 3.3.

are of high quality, compared to texts collected randomly on the internet. Each of the texts is written by a renowned scientist or a specialist in a given field, as opposed to e.g. Wikipedia, which also is a valuable source of information but lacks the credentials about the authors. Other texts that could be retrieved automatically from the Internet include forums, blogs and articles published on various websites, usually without any editing. In addition to that, the texts from SNL contain a fair amount of data about a given object or phenomenon which makes the samples representative – each of the texts contains at least one generic reference. The texts were chosen in a way that guarantees the diversity of topics, as well as a proper length of each of them. That would not be possible in the case of random text collection for a bigger corpus.

The decision to use online encyclopaedic texts in the corpus construction was also dictated by practical reasons. Utilising any paper source such as journals and books would mean the necessity to transcribe the texts first and then create a database that could be processed automatically. Considering the availability of computer tools and linguistic packages for Open Source programs, such a method would be a waste of time and would not necessarily give better results. Encyclopaedic texts are nowadays easily found online and some of them have even been transcribed from paper editions. Online encyclopaedia edition is also regularly updated and corrected, which can be traced at the bottom of each page of SNL where the date of publication and corrections is given. That cannot be said of most of the Norwegian journals that could have been utilised for data collection (for instance *Illustrert vitenskap* or *Historie* among others).

A drawback of such a choice of the material is the fact that the data was chosen based on the research question itself. This is in fact true but considering the points mentioned earlier, namely limited online resources, difficulty in manual tagging of genericity and the scope of the whole project, the solution seemed to be the most reasonable one and the one that would provide enough material to analyse. A possible risk connected to newspaper texts or literary works is that even a big amount of data might not provide as many instances of generics as shorter texts from an encyclopaedia. One can easily imagine that looking for generics in modern literary works or newspapers would take much more time than analysing the material itself. An optimal solution to all of these issues was therefore to use texts that are easily available, up to date and that are written by a renowned authors and specialists.

Another problem connected to the genres is that opting for encyclopaedic texts only means that the data is highly homogenous. Even though each of the texts has a different author and is therefore written in a slightly different manner, the genre remains the same. The homogeneity of data can be both an advantage, as well as a disadvantage of the study. Focusing on one particular text genre, which in addition contains many generic expressions, makes the collected samples relevant for the study. On the other hand, a set of data that is very homogenous can also put limitations to the analysis –

we can only analyse genericity in encyclopaedic texts, not in written language generally. In order to avoid that problem, an AJT survey based on the corpus data was designed (see chapter 6 for further discussion) and it comprises judgements based on language intuitions. Referring to the language intuition can illustrate how a language's grammar works (Devitt 2010). Combining the two methods can therefore diminish the homogeneity of the collected data.

As has been stated before, genericity does not occur in the language as often as for instance genitive form or any other strictly grammatical phenomenon. Generics, being rather a semantic aspect of the language, depend on the context but also on the topic of a given text. Searching for generic expressions in texts that concern a few domains seems therefore more reasonable than searching for it in modern literary works or in the news. One can of course find a generic sentence or two in the news or in a collection of short stories but such texts may also be devoid of a broader generic context. Relying on texts that concern for instance one animal species or a particular tool guarantees that generic expressions will occur in a context but also that the generic meaning will not be as ambiguous as it can be in other sources.

Even though many linguistic works are based on sentence analysis, context plays an important role in this study. Tagging the texts for genericity was already a challenging task and generic context of the material was certainly of great help. Creating individual sentences for the analysis lacks that context and can therefore lead to misunderstandings. What an author of a study thinks of as generic and unambiguous, can be a specific reference or an ambiguous sentence for another native speaker of a studied language.

Last but not least, building a specialised corpus implies that a given phenomenon or phenomena will be tested based on the collected material. Specialised corpora are homogenous as they contain texts of one genre, texts concerning the same topic or even texts written by the same author (e.g. specialised corpora of literary texts). Therefore an argument that the data is too homogenous and cannot show all the aspects of genericity is invalid. Genericity tends to occur in specific types of texts, such as scientific texts, so it only seems reasonable to collect and annotate the texts that are very likely to be generic and to contain a lot of generic expressions, rather than looking for genericity in a place where it cannot be found.

As has been discussed in section 2.2.3, generic texts differ from texts that only contain generic sentences. In short, a generic text is a text that concerns one topic and makes generalisations and assumptions about it. A text containing generic sentences is the one where such sentences can occur but they do not need to be frequent. An example of a generic text can be a description of a plant found in a botanical book, whereas 'The Little Prince' analysed in Behrens' study is not a generic text – the novel's topic does not concern only foxes or roses.

Building and annotating a corpus of texts can be a complex task, not only because of

the amount of the material that needs to be compiled but also because of the theoretical questions one needs to face. The choice of the material and text genres can influence the analysis and the results. In the case of this study, choosing an online encyclopaedia as a source for all of the corpus texts proved to be a good decision. All of the samples contained generic expressions and all of them were generic, which allowed to analyse the material on three levels, as mentioned by Behrens (2005). The structure of the corpus and data analysis will be presented in the following sections of this chapter.

5.2   Corpus structure and tagging

As has been mentioned before, the corpus of generic texts is based on the texts retrieved from SNL, a Norwegian online encyclopaedia. The whole data set consists of 170 texts in total (27 761 words) and is divided into five categories. The choice of texts, as well as categories of the corpus was inspired by the research question, namely it was necessary to include the texts that concern both animate and inanimate objects, countable and uncountable nouns, as well as nouns that describe concepts, activities and professions.

Corpus data was collected during six months: from January to June 2018[4]. The texts retrieved from SNL were divided into five main categories: 1) people (20 texts), 2) animals (25 texts), 3) plants (25 texts), 4) tools (25 texts) and 5) other (75 texts).

The division of the texts into particular categories was done based on a few criteria. First of all, the texts had to concern both well-known nouns such as e.g. 'a dog', 'a horse', as well as less known members of a given group such as 'a rhino' in the case of animals. Second of all, the category 'other' comprises 75 texts, as opposed to 25 and 20 in the rest of the categories. The reason for this is that, apart from main categories such as animals and plants, the nouns could be classified in many different groups (e.g. vehicles, weapons, everyday objects etc.) which could be problematic for the data management. In order to avoid diving the corpus into many small categories of topics, one big category of nouns was created.

The category people comprises only names of groups, such as *Kuvlunger* ('Kuvlung'), and functions such as *konge* ('king') and *rabbi* ('rabbi'). Professions were not included in this category – they are a part of the category 'other'. The reason for this is that professions are rather considered as names of certain jobs and in most cases they do not designate people who perform those tasks. If the references were specific, then the nouns would have been included in the group 'people'. Since the references are generic and therefore describe generalisations about the professions, they are not considered as designating people. The case with category 'people' is somehow different – when we talk about an ethnic group or people with aristocratic background or who have some special

functions in the society (e.g. kings, priests etc.), we perceive them as individuals, not institutions. Such interpretation was applied throughout the corpus.

Such a wide variety of nouns in the corpus can also allow for a more detailed analysis than the ones conducted so far (e.g. Carlsson's works on biological terms in Swedish, and Borthen's and Halmøy's research on Norwegian nouns) but also provide data that has not been collected in this way before. One of the aspects that is of particular interest is the division of studied texts into particular categories. This can be done in two ways: 1) according to the text's topic (5 categories of texts), 2) according to the animacy factor (animate vs. inanimate). A third possible factor can be the category of well-established kinds (WEK, Carlson and Pelletier 1995). Some of the nouns, especially the ones that are somehow typical for a given culture, occur most often in definite form or as BNs. This can suggest that they are so well known that they function almost as proper names or abstract concepts. Identifying nouns as WEKs can be ambiguous though. There are no formal criteria concerning such a division of nouns, therefore they were not tagged during the collection of the data and only briefly discussed as a subcategory of the corpus.

The texts were saved in an .xlsx file, without any formatting or hyperlinks in it. The file with 170 cells (one cell per text) was then uploaded into the R software in order to perform the annotation of the material. The first step was to automatically mark all of the nouns in the corpus. This was possible with the use of an open source package, UDPipe.

After all nouns in the corpus were filtered out, the tagging was done manually. Each of the texts was saved in a separate .xlsx file. The tags added included:

1. GEN-tag
2. NP type
3. position in the sentence
4. function in the sentence

Tagging was verified two times in order to minimise the risk of mistakes occurring in the material. A random sample of tagged texts was then consulted with a native speaker of Norwegian who confirmed the tagging. Because of the available resources and time constraints, it was not possible to hire another annotator, preferably a native speaker of Norwegian, who would check the whole material.

After all of the texts were annotated and double-checked, generic sentences were filtered out. Each of the texts contained at least one generic reference and in the majority of the cases it was in average 3-4 generic expressions. Some of the texts were short, for instance those concerning well-known and widely used tools or objects. Some of the animal species were described in much greater detail than for instance exotic animals, that are not present in the Scandinavian climate. Even though it was impossible to find 170 diverse texts that would be of exactly same length, the collected material has proved to contain enough of instances of genericity to conduct the analysis (see the statistical

data in section 5.3.

As has been mentioned above, each of the corpus texts was saved in a separate .xlxs file. When the tagging was completed, the texts were merged with the use of R software. This was necessary for the analysis to be carried out. The division of the texts into 170 files had also another advantage, namely it made the tagging process easier and more precise, since each of the files concerned only one noun or one noun phrase. Had all the texts been saved in one .xlxs file, the length of the file could have made it much more difficult to notice and tag correctly all instances of genericity. What is more, choosing a random sample of texts for the native speaker to control was also feasible thanks to this tagging method.

In order to annotate and then filter out generics, all nouns and pronouns in all the texts had to be correctly identified. Thanks to the linguistic package mentioned above, the task was possible but it is worth mentioning that in few cases manual annotation of the nouns and pronouns was necessary. A possible explanation to why the package did not recognise some of the nouns as nouns is that they were compound nouns or loan words. However, the number of such instances was rather small and necessary corrections were performed manually and double checked afterwards. In the case of pronouns, the problem did not occur.

5.3   Collected data

The texts collected in the corpus consist of 27 761 tokens[5] in total. 870 nouns and noun phrases are annotated as generic, making it 9,34% of all tokens. Each of the texts contained at least one generic reference (most of the time a lot more than that), which would not be the case should newspaper texts or literary works be analysed. What is more, generic nouns and phrases in the collected material were rather straightforward because of the texts' topics, namely they focus on general information about a given object, species, profession etc.. When it comes to everyday speech and written language of a different genre, understanding of the context and its interpretation would play a much greater role in annotating generics.

The nouns and noun phrases annotated as generic were sorted according to the form they occurred in. Norwegian generic nouns can take on all five NP types, namely bare form, indefinite and definite singular forms, as well as indefinite and definite plural forms. As has been discussed in section 3.2, generic expressions can be created with the use of either of the forms but with slight differences in meaning. These differences, however, depend on many factors, one of them being the context, as the one of the corpus texts. In the collected material, the most frequently occurring form of generic nouns was BN, followed by indefinite plural and definite singular forms, as shown in the table and the

---

[5]Throughout the thesis the notion of 'token' is used interchangeably with a more colloquial 'word'.

graph below.

Forms	Number	%
bare nouns	345	39.66
indefinite sing.	31	3.56
definite sing.	185	21.26
indefinite pl.	230	26.44
definite pl.	79	9.08
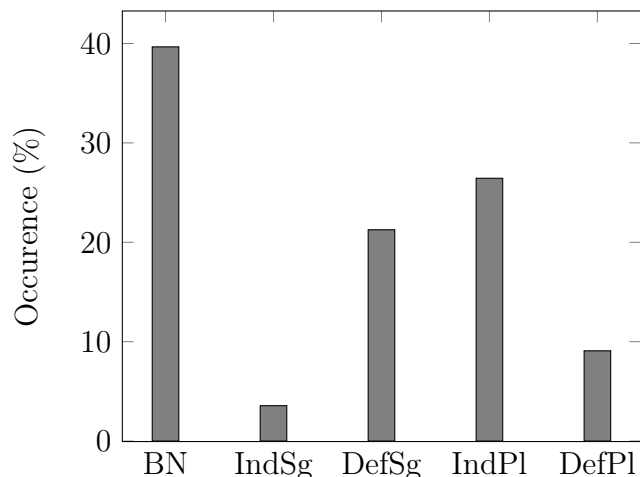
Table 5.1: Corpus – general results

Figure 5.1: Corpus – general results

According to the theory on the matter, bare nouns are not very likely to occur in generic contexts, especially when it comes to countable nouns (see e.g. Borthen 2003 and Halmøy 2016). This may prove to be true in the analyses of individual sentences, where BNs are acceptable either in the case of mass nouns or when they were used with WEK (see section 2.2.1.1 of this chapter).

One of the instances where the use of BNs is the most obvious is in the case of encyclopaedic key words. Such words were often both titles of the texts, as well as first mention of a given subject in the text. Norwegian bare nouns in the corpus were not only key words but, interestingly enough, the form occurred very often in subsequent mentions, following another forms used earlier in the texts. This can suggest that Norwegian generic BNs are not only mass nouns or nouns used in the encyclopaedia as entries. Some of them do qualify into the WEK category but a great number of nouns used as BNs throughout the corpus were *regular* nouns (see the examples in 120).

(120) a. Sølvsmed   er fra   1995 ett lærefag.
          silversmith-Ø is from 1995 a   subject
          Silversmith is a subject from 1995.

       b. Mausoleum   er et monumentalt gravmæle.
          mausoleum-Ø is a monumental gravestone
          A mausoleum is a monumental gravestone.

       c. Magnet   er en gjenstand som gir   opphav til et magnetisk felt.
          magnet-Ø is an object   that gives origin   to a magnetic   field
          A magnet is an object that produces magnetic field.

The two other NP types that occurred most often in generic contexts are indefinite plural and definite singular forms. The first one is a form typically connected with genericity in Germanic languages (see e.g. Radden and Dirven 2007 and Radden 2009). Also

in Norwegian bare plurals refer to an unspecific group of objects or animals, making a reference generic.

(121)   a. Skater er god matfisk.
              rays   are good food-fish
              Rays are good food.

          b. Krokodiller er en orden av store krypdyr.
              crocodiles   are an order of big   reptiles
              Crocodiles are an order of big reptiles.

          c. Skruer har mange anvendelser og former.
              screws have many uses   and forms
              Screws have many uses and forms.

When it comes to definite singular form in generic contexts, it occurred in 21,26% of the sentences marked as generic.

(122)   a. Gitaren   kom opprinnelig fra   Midtøsten.
              gitar-DEF came originally   from Middle.East
              The guitar came originally from Middle East.

          b. Mikroelektronikken har revolusjonert kalkulatorteknikken.
              microelectronic-DEF have revolutionised calculator.technique-DEF
              Microelectronics have revolutionised calculator techniques.

          c. I naturen   finnes sjiraffen   i dag   bare i Afrika
              in nature-DEF exists giraffe-DEF in today only in Afrika.
              Nowadays, giraffes live only in Africa.

The last of the examples shows that generic nouns do not need to occur at the beginning of a sentence. The position of generic nouns in the corpus sentences was in most cases initial but the difference was not significant. When it comes to NPs' grammatical functions in the texts, one can observe certain disproportion, as presented in table (5.3).

Position	Number	%
initial	443	50.92
non initial	427	49.08

Table 5.2: Corpus – NPs' position in a sentence
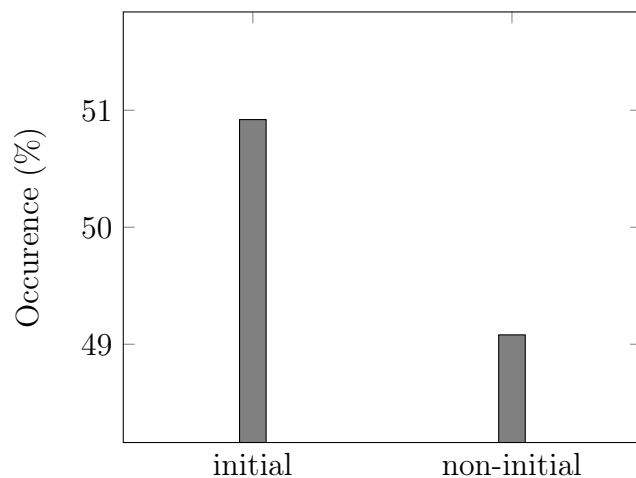
Figure 5.2: Corpus – NPs' position in a sentence

Functions	Number	%
subject	577	66.32
object	146	16.78
genitive modifier	34	3.91
other	113	12.99

Table 5.3: Corpus – NPs' function in a sentence
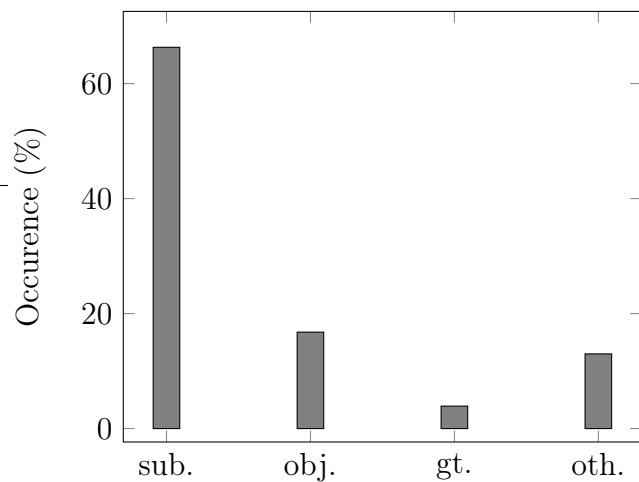
Figure 5.3: Corpus – NPs' function in a sentence

Genericity can occur on different levels. One differentiates therefore between 1) NP-level generics, 2) sentence-level generics and 3) text-level generics. NP-level genericity can co-occur with sentence-level and text-level genericity but, in contrary to the latter types, it always contains a generic noun. Sentences with a generic NP in a subject position are regarded in scholarly literature as the most common type of generics (see chapters 2 and 3). These type of sentences were the most frequent ones in the corpus, constituting 66.32% of all generic sentences. However, this applies to generic sentences that include a generic NP. As has been mentioned before, certain types of generic sentences, e.g. habituals, do not need to include generic NPs at all. The generic reading might be provided by a kind-predicate or context. The former is typical for kind-referring sentences, whereas the latter applies to all types of generics. In the corpus, the majority of generic sentences were kind-referring and therefore they had generic NPs.

In the following sections of this chapter thematic categories, mass nouns and well-established kinds will be analysed statistically. Similarly as with the pilot survey, ANOVA variance analysis and post-hoc tests were conducted on the collected data. The goal of such method was to determine whether there were statistically relevant differences in use of each of the NP types and if so, where these differences were.

5.3.1 Thematic categories

The nouns in the corpus (texts' topics) were divided into 5 thematic categories ('people', 'animals', 'plants', 'tools' and 'other'). One of the factors taken into account when dividing the texts into categories was animacy which I shall now briefly describe.

All the texts in categories 'people' and 'animals' are animate nouns, whereas the nouns in other categories are generally regarded as inanimate. Animacy however is a rather complicated notion as it can be seen from both biological and linguistic point of

view. Biological animacy refers, as the name implies, to biological processes that only living organisms can perform. In this sense plants but also bacteria and fungi are animate. In the linguistic sense, only truly animate nouns can occur with control predicates (cf. Comrie 1989: 59-62), which express action taken with a will. Plants could occur with certain predicates that express actions (e.g. 'to give shadow', 'open the door' as in 'the wind opened the door' etc.) but they cannot occur with control predicates.

Classifying people or animals as animate usually does not pose problems but when it comes to plants the case is not as simple. Certain languages have a very clear-cut classifications and differentiate e.g. between humans (animate) and other creatures and objects (inanimate), whereas in other languages the system of identifying as (inanimate) is a lot more complex (Comrie 1989: 185).

Comrie (1989) proposes also the general degree of animacy that applies to a great majority of languages:

<div align="center">human > animal > inanimate</div>

A more detailed model was proposed by Rosenbach (2008: 153), where animate nouns are classified in the following manner:

<div align="center">human N > animal N > collective N</div>

Including collective nouns in this classification allows for a broader interpretation of certain NPs, especially those occurring as BNs in Norwegian (for further discussion see section 5.3.1.1).

Comrie (1989: 62) claims also that *a high degree of animacy is necessary for a noun phrase to be interpreted as having a high degree of control or as an experiencer, but is not a sufficient condition.* According to this description, only creatures that are considered highly animate, namely not only living but also breathing, moving and so on, have high degree of control (can perform actions but also cause others to perform them). Plants clearly do not conform to this description, even though they might be considered biologically animate (=living/alive). The ability of such nouns to occur as agents was discussed by Becker (2014: 170) who says that

> plants would be described by adults as "alive" but they bear none of the prototypical characteristics of animate things: they do not move on their own (save for subtly orienting toward the sun, which small children are unlikely to have witnessed), they lack prototypical animate features like faces, and they lack intentions as far as we know. In short, plants are alive but inanimate, so the label "alive" does not equate with "animate".

Plants can by all means be interpreted cognitively as animate in the sense "living/alive" but they cannot be perceived as such in a purely linguistic sense. Other researchers go as far as to simply say that *plants are always inanimate* Yamamoto (1999: 48), whereas

others refer to *folk biology* or *folk taxonomy* (Becker 2014: 288) according to which plants could be perceived as animate in certain contexts. For the sake of this study, plants are considered linguistically inanimate but cognitively and biologically animate, which is also mirrored in the corpus data.

The animate nouns in the corpus occurred mainly in two categories: people (ethnic groups and official functions) and animals (names of animal species). As has been stated before, professions were regarded as category 'other' because they imply something else than descriptions of groups of people. However, the nouns could be considered animate when interpreted as people performing certain professions, not professions *per se*. These examples will be analysed and discussed separately from the categories 'people' and 'animals'.

5.3.1.1 Category 'people'

Category 'people' consists of 20 texts that describe ethnic groups and official functions that are not considered professions in the narrow sense, such as 'king' or 'priest'. The category is the smallest among all five groups of topics since many of the people groups and ethnic minorities were described in SNL with only one or two sentences. Since each of the corpus texts had a length of at least one paragraph, the choice of the source material in this category was limited. However, the difference is not significant and it does not seem to influence the general results.

In the category 'people', there were 98 generic expressions. The nouns appeared in all five forms, however the use of indefinite singular was somehow limited (10.20%). The most frequently utilised NP types were definite singulars, indefinite and definite plural forms, as well as bare nouns. The definite singular form, used in 31.63% of the cases, might suggest that a number of NPs in this text category function as well-established kinds or that they can be perceived as prototypes of a given social group or people belonging to this class.

As can be seen in table (5.4) and figure (5.4), the use of definite plural generics is almost at the same level as BNs and indefinite plural forms. It may suggest that definite plural generics are in fact most often used when talking about people, similarly as in English. However, they are not limited to this use, contrary to English plural generics (see section 2.4 and e.g. Radden and Dirven 2007 for further discussion).

Forms	Number	%
bare nouns	19	19.39
indefinite sing.	10	10.20
definite sing.	31	31.63
indefinite pl.	20	20.41
definite pl.	18	18.37
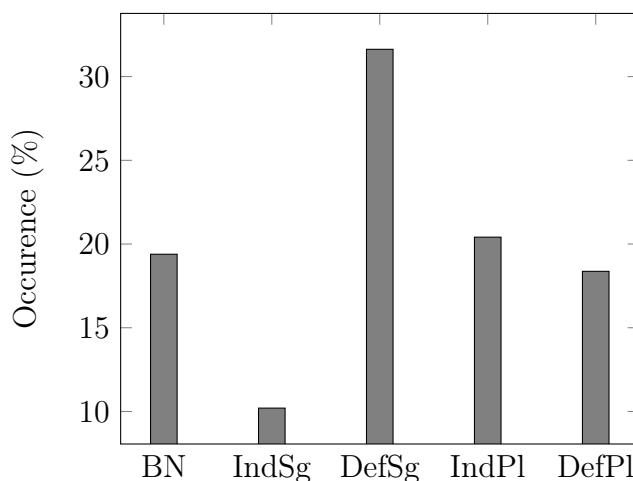
Table 5.4: Corpus – generic NP types (people)

Figure 5.4: Corpus – generic NP types (people)

The sentences in (123)[6] show the use of the three most common NP types in the category 'people'. Among these, BN and definite plural seem to play a slightly different role than definite singular, namely they refer rather to whole social groups, not people of a particular profession. What is more, 'aristocracy' can be regarded as a collective noun (see the classification of collective NPs in the previous section), referring to a certain social background, whereas 'farmers' clearly refers to a group of people.

(123)   a.   **Adel** er en samfunnsklasse med lovfestede, arvelige politiske og sosiale forrettigheter; forekommer fra de eldste tider i de fleste kulturer.

Aristocracy is a social class with legal, hereditary and political privileges; occurs from the oldest times and in most cultures.

b.   **Prester** [priests] får sine oppgaver tildelt enten ved arv eller ved personlig valg (kall).

Priests receive their tasks either by heritage or by a personal choice (call).

c.   I Tyskland øst for Elben og i de slaviske land ble **bøndene** [farmers-DEF] til og med livegne, og godseierne la helt urimelig arbeidsplikt på dem.

In Germany, East for Elben, and in the Slavic countries, farmers were indentured and the land owners obliged them to do an unimaginable amount of work.

The difference between the reference to an official function and to people who do a certain type of work as a result of the social class they belong to, is shown in sentences (123b) and (123c). One might argue whether or not 'priest' denotes a profession, an official function or even a call. Nevertheless, the indefinite form 'priests' is unambiguously generic, whereas the definite plural in 123c might suggest both specific and habitual reading of the sentence.

---

[6]The glosses in English are added in square brackets and to generic nouns only.

On the one hand, the NP refers to farmers from the given region and on the other, it denotes a whole social class by giving an example of farmers from Germany and Slavic countries.

(124) a. **En biskop** [a bishop] er den øverste kirkelige lederen i et bispedømme.

A bishop is the highest church leader in a diocese.

b. **Fylkesmannen** [county.governor-DEF] har også en rekke direkte oppgaver i forhold til helsetjenesten, dels med hjemmel direkte i lovgivningen og dels for eksempel på vegne av Sosial- og helsedirektoratet, særlig innenfor fore-byggende helsearbeid og relatert til helsepolitiske satsninger.

The County Governor has a number of tasks concerning health care, which are partially stated in the law and partially being an obligation given by the Norwegian Directorate for Health and Social Affairs, especially concerning preventive health work and health policies.

The two singular forms, indefinite and definite in (124) show the reference between what could be considered a call or a life role and an official function, perceived almost as a profession. In (124a), 'a bishop' refers rather to the tasks such a person performs rather than people being in this position, whereas the definite form of 'the County Governor' in (124b) implies that the profession is somehow unique. Being a county governor is not a very common profession and it is therefore seen as a particular official function, similarly as a king or a politician.

When it comes to the place and function of generic NPs in the category 'people', the majority of the nouns were subjects (71%) in initial position (53.06%). However, in a number of examples one could observe generic NPs in object position (14.29%). Genitive and other grammatical forms occurred only in 13 sentences.

The position and function of the nouns in generic sentences depends on the writing style, as well as the topic of the text. However, in all text categories, generic nouns are subjects in the majority of sentences, with other functions being a lot less common. It may mean that in *classic* generic sentences and texts, generic NPs are subjects, irrespectively of their position in the sentence (cf. Lyons 1977).

Figure 5.5: Corpus – noun position (people)

Position	Number	%
initial	52	53.06
non initial	46	46.94

Table 5.5: Corpus – noun position (people)
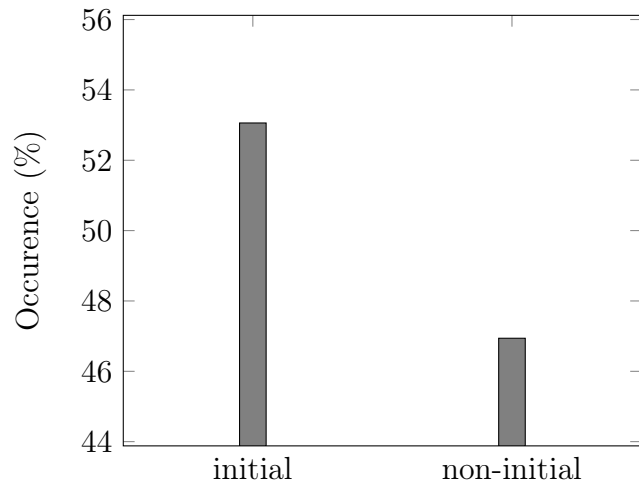
Functions	Number	%
subject	71	72.45
object	14	14.29
genitive modifier	7	7.14
other	6	6.12

Table 5.6: Corpus – noun function (people)
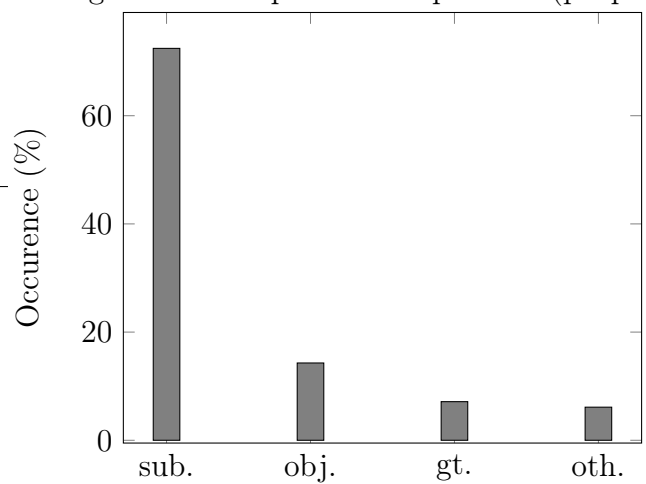
Figure 5.6: Corpus – noun function (people)

5.3.1.2   Category 'animals'

The other text group that contains animate nouns is the category 'animals' consisting of 25 texts. The description of animal species are one of the most typical generic texts (cf. the study of Swedish generics by Carlsson 2012). Amon all NPs in this text category, 11.20% were generic references.

Nearly half of all generic nouns denoting animal species were bare plurals (49.66%). Only indefinite singulars were not used in generic contexts to refer to animals, whereas definite singulars, bare nouns and definite plurals appeared in a number of sentences (21.77%, 17.01% and 11.56% respectively, see table 5.7).

Forms	Number	%
bare nouns	25	17.01
indefinite sing.	0	0.00
definite sing.	32	21.77
indefinite pl.	73	49.66
definite pl.	17	11.56
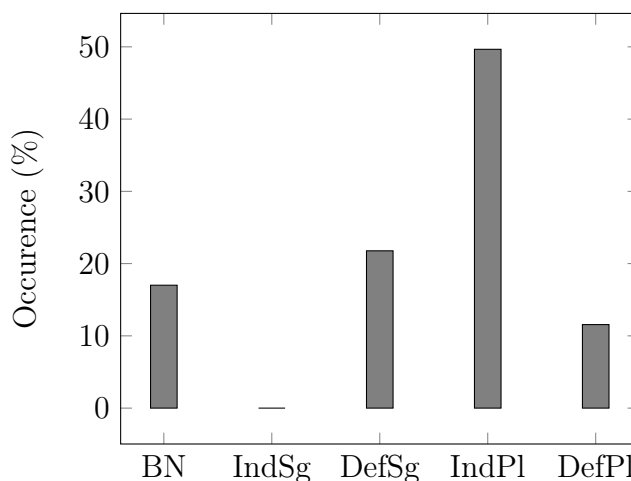
Table 5.7: Corpus – generic NP types (animals)

Figure 5.7: Corpus – generic NP types (animals)

The sentences in (125) show examples with each of those forms. Bare noun in (125a) is used as a subject in a classic generic sentence that describes a kind. The predicate used in the sentences cannot be interpreted as kind-referring. Weighing a certain amount of kilograms can be a characteristic feature of a species but not in the same sense as e.g. being a mammal. One can easily imagine dodo birds weighing more than 25 or less than 20 kg.

(125)   a. **Dronte**, også kalt **dodo** [dodo-Ø], Raphus cucullatus, på Mauritius veide 20–25 kg og fantes i store mengder da europeerne kom dit i 1598.

Dodo, Raphus cucullatus, on Mauritius it weighed 20-25 kg and occurred in large quantities when the Europeans arrived in 1598.

b. **Skater** [skates] er god matfisk.

Skates are good edible fish.

c. **Oteren** [oter-DEF] er nattaktiv.

The oter is a nocturnal animal.

d. **Humlene** [bees-DEF] danner samfunn og betegnes sammen med honningbiene som sosiale bier.

Bumblebees create societies and together with honey bees are regarded as social bees.

The sentence in (125b) is a model kind-reference, where the NP occurs as a bare plural – a *default* generic form (cf. 2 and 3). 'The otter' combined with the predicate 'to be a nocturnal animal' in (125c) is a generic sentence but not a kind-reference in a classic understanding of the term. The last example in (125d) utilises a definite plural form 'the bumblebees' when referring to the species. The predicate is truly generic – it characterises the kind's behaviour.

What is more, the animal species that occurred as definite plurals in the corpus were either small animals perceived almost as mass (e.g. insects) or animals considered gregarious, mainly as a result of folk taxonomy. These nouns included such species as seagulls, crocodiles, bats and mice, among others. On the other hand, certain animals such as giraffes, which are gregarious animals, occurred in the corpus either as BNs or definite singulars. It is a rather *exotic* animal species from the European point of view, which might be one of the reasons for the use of singular NP in this case.

When it comes to generic nouns' positions in this category, they were mostly non-initial nouns (53.74%). The majority of the nouns were subjects (63.95%) and nearly 30% of all NPs in this category had other grammatical functions than subjects, objects and modifiers.

Position	Number	%
initial	68	46.26
non initial	79	53.74
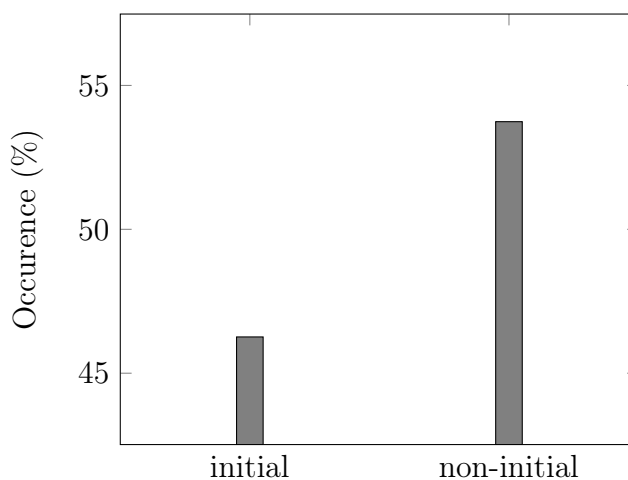
Table 5.8: Corpus – noun position (animals)

Figure 5.8: Corpus – noun position (animals)

Functions	Number	%
subject	94	63.95
object	6	4.08
genitive modifier	4	2.72
other	43	29.25
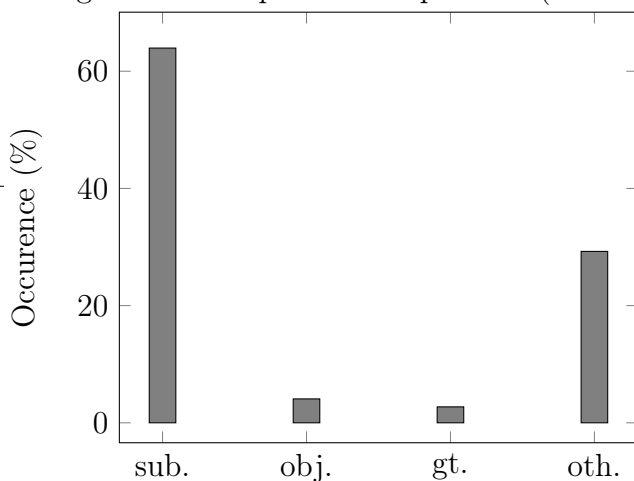
Table 5.9: Corpus – noun function (animals)

Figure 5.9: Corpus – noun function (animals)

5.3.1.3   Category 'plants'

The category 'plants' consisted of 25 texts, each of them denoting one plant. 142 sentences in this category were annotated as generic. Despite the fact that the plants are considered linguistically inanimate, the nouns in this category occurred in similar forms as animals, namely there was no instance of indefinite singular (see table 5.10). However, the use of BNs was much higher in the category plants (73.24%), with minor use of definite singular (9.86%), and definite (8.45%) and indefinite plural forms (8.45%).

Forms	Number	%
bare nouns	104	73.24
indefinite sing.	0	0.00
definite sing.	14	9.86
indefinite pl.	12	8.45
definite pl.	12	8.45

Table 5.10: Corpus – generic NP types (plants)
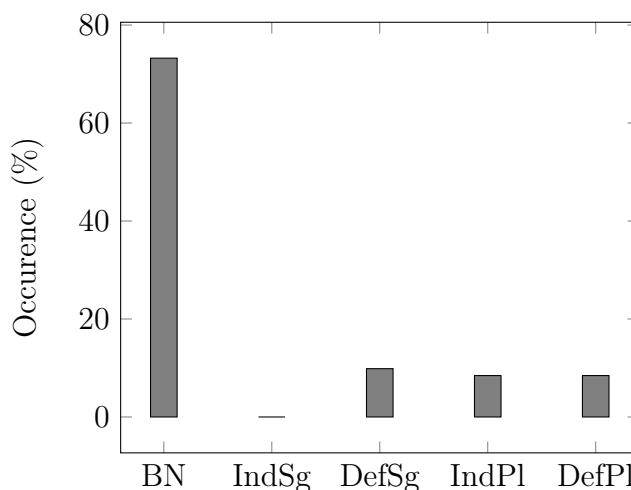
Figure 5.10: Corpus – generic NP types (plants)

The examples in (126) demonstrate the use of each of the NP types in descriptions of plants. 'Fennel' in (126a) can be understood as a mass noun, which explains the use of bare form. It is not a particular fennel or a number of fennels that are used as edible vegetables.

The definite singular in (126b) refers to a single tree as a prototype of all oaks. The predicate 'to be spread' can be considered a kind-predicate, as one cannot say that a single tree is spread on some area.

The plural forms in (126c) and (126d) express a generic reference in two ways: bare plural in (126c) means that ferns usually occur in tropical areas of the world but they can also grow in other areas (see the generic overgeneralisation effect in 2.4.3.1), whereas 'the agarics' in (126d) expresses a reference to muschrooms of this type. Definite plural generics in this sentence seem to access subkinds – there are many types of agarics and all of them have lamellae.

(126)   a. **Fennikel** [fennel-Ø] brukes også som grønnsak, rå eller kokt.

            Fennel is also used as a vegetable, raw or cooked.

        b. **Eika** [oak-DEF] hadde da en større utbredelse enn nå.

The oak was then more spread than now.

c. **Bregner** [ferns] er utbredt over hele Jorden, de fleste i tropene.

Ferns are spread in the whole word, mostly in tropical climate.

d. **Fluesoppene** [agarics-DEF] er skivesopper med hvite skiver.

The agarics are mushrooms with white lamellae.

The positions and functions of generic NPs in this category are depicted in tables 5.11 and 5.12. The majority of the nouns were subjects, with quite a high percentage of object function (22.54%). When it comes to initial and non-initial positions of generic nouns describing plants, they were equally distributed throughout the texts. A similar regularity was also observed in the category 'tools' which will be discussed below.

Position	Number	%
initial	71	50.00
non initial	71	50.00

Table 5.11: Corpus – noun position (plants)
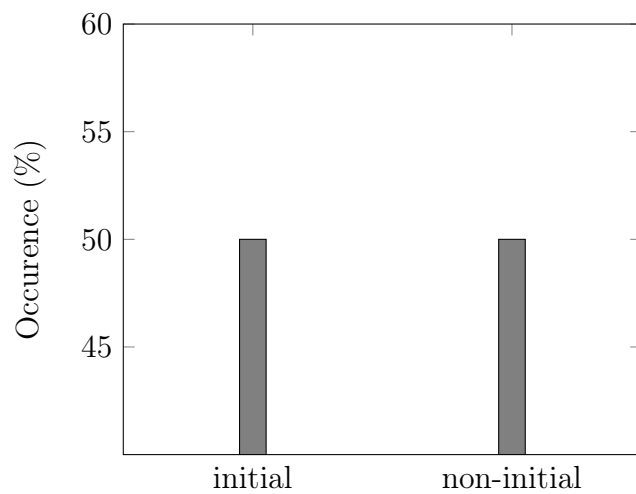
Figure 5.11: Corpus – noun position (plants)

Functions	Number	%
subject	89	62.68
object	32	22.54
genitive modifier	2	1.41
other	19	13.38

Table 5.12: Corpus – noun function (plants)
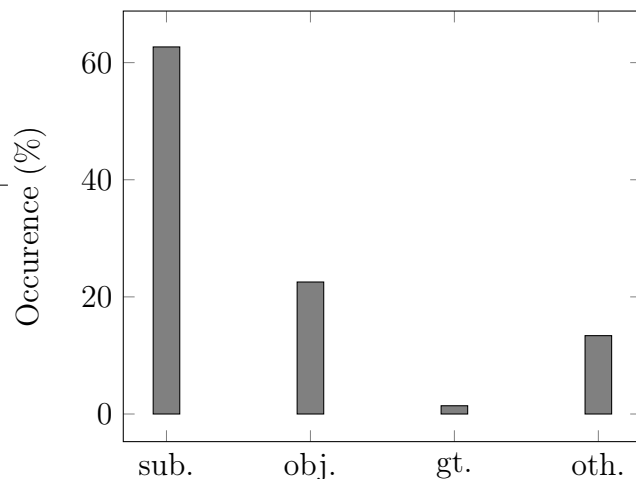
Figure 5.12: Corpus – noun function (plants)

5.3.1.4   Category 'tools'

The fourth category of corpus texts, namely 'tools', contained 25 texts about different utensils, both everyday objects and specialised devices. The category contained also one mass noun describing an action, namely welding. Among all nouns and pronouns in the texts, 122 nouns were tagged as generic.

Tools were described with the use of all five NP types available in Norwegian (see table **??**). Similarly as in the previous category of texts, BNs were most often used when referring to different objects stating 52.45% of nouns, with definite singular form being next most used (27.87%). Surprisingly enough, indefinite plural which is said to be a somehow *default* generic form in Germanic languages, was used only in 12.30% of the cases, followed by indefinite singular (4.10%) and definite plural forms (3.28%).

Forms	Number	%
bare nouns	64	52.46
indefinite sing.	5	4.10
definite sing.	34	27.87
indefinite pl.	15	12.30
definite pl.	4	3.28
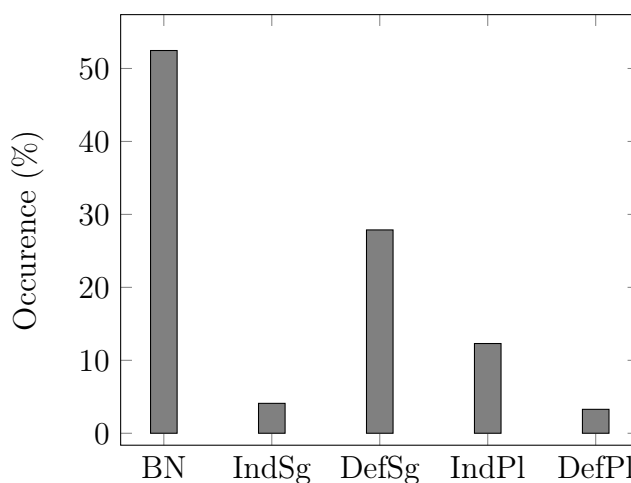
Table 5.13: Corpus – generic NP types (tools)

Figure 5.13: Corpus – generic NP types (tools)

The examples in 127 show how different NP types function in some of the corpus texts. The use of BN occurs in a very typical encyclopaedic sentence where the function of an object is explained. This way the noun is interpreted generically but also has a function of a prototype, similarly to indefinite singular in (127b) and definite singular in (127c). However, the difference between these three forms is that one cannot perceive BN as a singular, contrary to what certain researchers claim (cf. Borthen 2003). Following the view of Halmøy (2016) we shall interpret Norwegian BNs as unmarked in terms of definiteness and number. This makes the noun 'hand net' in (127a) not only a prototype of the device but also a concept of it (cf. Lyons 1977 and Pettersson 1976).

The examples in (127d) and (127e) are plural nouns. The sentence with the indefinite noun is a generic sentence, giving main characteristics of screws in general and therefore applying to most (if not all) objects from this category. The definite noun 'the gun', referring to spray guns used for different types of painting, refers to a subkind – guns designed particularly for powder coating. Similarly as in the category 'plants', definite plural generics can access subkinds.

(127) a. **Høvel** [hand.net-Ø] er et verktøy til utjevning av treoverflater.

A hand plane is a tool used to even wooden surfaces.

b. **En plateskrue** [a plate.screw] er spiss og laget av herdet stål.

A plate screw is sharp and made of hardened steel.

c. **Tangen** [pliers-DEF] kan brukes: til å holde fast et arbeidsstykke, til å gjennomskjære det, til å forme det ved pressing.

Pliers can be used: to hold an object one is working on, to cut through it, to form the piece by pressing it.

d. **Skruer** [screws] har mange anvendelser og former. Man skjelner mellom **skruer** beregnet for stål og metalldeler, og for arbeid i tre.

Screws have many uses and forms. One differentiates between screws meant for steel and metal parts, and those meant for working with wood.

e. **Pistolene** [guns-DEF] kan også være konstruert spesielt for elektrostatisk påsprøyting.

The guns [powder coating guns] can also be specially made for powder coating.

As has been said before, exactly half of the generic nouns in the category 'tools' occurred in initial position and the other half in non-initial one. The majority of the nouns, namely 57.38%, were subjects, 25.42% objects and the rest occurred in genitive constructions or had other grammatical functions.

Position	Number	%
initial	61	50.00
non initial	61	50.00

Figure 5.14: Corpus – noun position (tools)
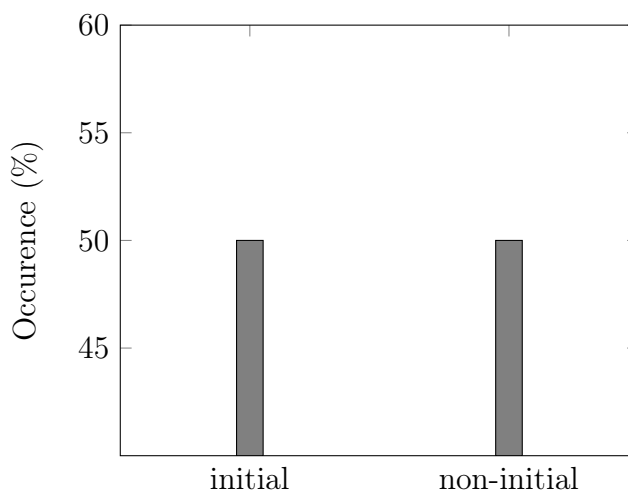
Table 5.14: Corpus – noun position (tools)

Functions	Number	%
subject	70	57.38
object	31	25.41
genitive modifier	7	5.74
other	14	11.48
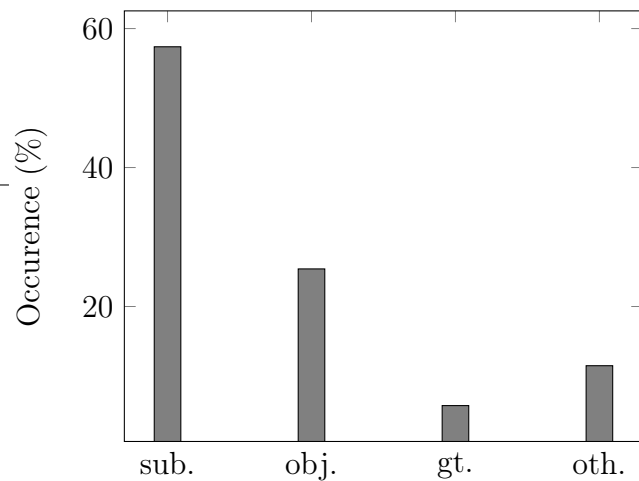
Table 5.15: Corpus – noun function (tools)

Figure 5.15: Corpus – noun function (tools)

5.3.1.5   Category 'other'

The last and the biggest category of corpus texts is the category 'other'. It comprises 75 texts denoting everyday objects, buildings, notions and phenomena, as well as professions, which will be discussed in a separate subsection in the following part of the chapter. The nouns collected in this category were so diverse that it would be challenging to create a small subcategory for each of them. It is also therefore this group of texts is much bigger than the other ones. The category contains also mass nouns, which will be discussed further later in the chapter.

Genericity was tagged in 361 sentences. Generic nouns in this category occurred in all five forms, BN being most frequent and stating 36.84% of all generic nouns. The indefinite plural form appeared in 30.47% and the definite singular in 20.50%, making those three forms dominant in this group. Definite plural and indefinite singular nouns constituted 7.76% and 4.43% respectively.

Forms	Number	%
bare nouns	133	36.84
indefinite sing.	16	4.43
definite sing.	74	20.50
indefinite pl.	110	30.47
definite pl.	28	7.76
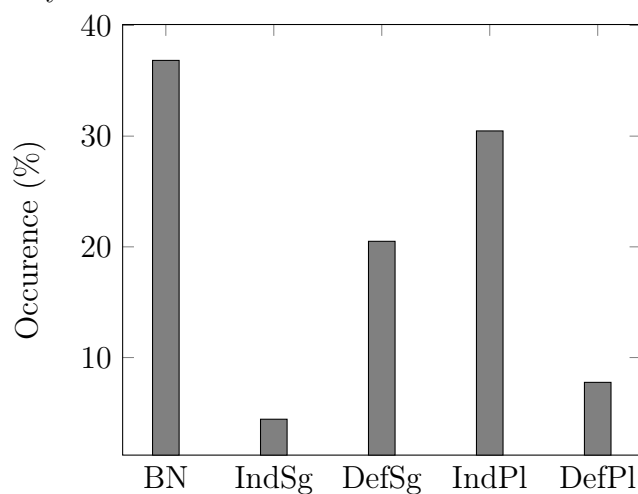
Table 5.16: Corpus – generic NP types (other)

Figure 5.16: Corpus – generic NP types (other)

The examples retrieved from this part of the corpus were divided into those concerning objects such as those in (128), and those describing professions in (129).

Similarly as in the previous categories, all five NP types were utilised in generic contexts. The two forms that are op particular interest are BN and definite plural, as those have slightly different functions than e.g. in English. The bare noun 'stavkirke' in (128a), combined with what could be considered as a well-established kind, is not only an encyclopaedic entry but also a concept of this type of church. The sentence describes the characteristics of stave churches, allowing for a truly generic reading.

The definite plural form in (128e) occurs with past tense and the adverb 'often'. Again, the definite plural form accesses a subkind (a type of tent), whereas the predicate and the adverb make the whole sentence into a characterising expression. The use of past tense in a generic sentence is not very common but possible and it does not influence the generic reading (see Lyons 1977). A similar structure can be observed in 128c, where the past tense is used with definite singular.

The sentences in (128b) and (128d) are again typical generic sentences – they describe functions and features of some objects – axles and fish hooks respectively.

(128)   a. **Stavkirke** [stave.church-Ø] er en høyt utviklet kirketype, oppført i reisverk av tre, kjent fra norsk kristen middelalder.

   A stave church is a highly developed type of churches, constructed with wooden staves, originating in Christian medieval Norway.

   b. **En aksel** [an axle] kan også være en stillestående maskindel hvor de roterende deler er lagret, f.eks. foraksel på en bil, og en vanlig vognaksel.

   An axle can also be a still part of a machine in which the rotating parts are stored, e.g. a front axle in a car and a regular wagon axle.

   c. **Badstuen** [sauna-DEF] var normalt innredet med en oppmurt røykovn av naturstein midt på gulvet uten røykavtrekk.

   A sauna was normally furnished with a furnace oven made of natural stone, and without a smoke extractor.

   d. **Fiskekroker** [fish.hooks] produseres i 12 000 forskjellige størrelser og varianter, fra 4–5 mm opptil 30 cm.

   Fish hooks are produced in 12 000 different sizes and variants, from 4-5 mm up to 30 cm.

   e. **Tipiene** [tipis] var ofte dekorert med symboler og tegninger som fortalte om eierens bragder.

   Tipis were often decorated with symbols and drawings depicting the owner's feats.

Professions   The difference between professions and words describing official functions is that in order to become e.g. a psychologist one needs to fulfil certain level of education in a given subject, whereas being e.g. a county governor does not imply that one studied at a given faculty (this is illustrated particularly well in example 129a). It might be desired but it not obligatory. The division between professions and expressions describing people groups (not only ethnicity but also class background etc.) may seem vague but it was possible to draw the line between the two for the sake of this study. This has allowed to illustrate how different those two categories are.

The examples listed in (129) show descriptions of some of the professions that were included in the corpus. The BN in 129a describes a person who performs this type of job, whereas indefinite singular used as a genitival modifier in (129b) directs the main focus on the work connected with the profession. The work of a psychologist is also described with the use of bare plural in (129d), where a reference to a group of people is expressed. Definite singular form in (129c) has a two-fold function: on the one hand it directs the focus on the profession and on the other, it shows what tasks band leaders are responsible for.

(129)   a. **Psykolog** [psychologist-Ø] er en person med utdanning i psykologi.

A psychologist is a person with education in psychology.

b. **En bakers** [a baker-GT] arbeid er nær beslektet med det som utføres av en konditor, og mange yrkesutøvere i bransjen har avlagt svenneprøve i begge disse håndverksfagene.

A baker's work is closely related to work of a confectioner, and many other workers in this line choose a certification exam in both these rofessions.

c. **Dirigenten** [band.leader-DEF] instruerer orkesteret eller koret, bestemmer tempo og dynamikk og sørger for at musikkverket blir forsvarlig innøvd.

A band leader instructs an orchestra or a choir, controls the tempo and the dynamics, and makes sure a musical piece is performed right.

d. **Psykologer** [psychologists] arbeider som helsepersonell.

Psychologists work as health staff.

Interestingly enough, none of the professions occurred as a definite plural. The assumption that the professions function differently than notions that describe people was initially confirmed by the corpus data. However, certain examples from this category such as (129a) and (129d) do in fact refer rather to people performing the job than to the profession itself. Nevertheless, texts about groups of people had a slightly different structure than those in the subcategory 'professions', namely a text about e.g. the Bagli party would concern the origin of that group, their history etc., whereas describing a profession focuses more on the work and only in a smaller degree on people performing

it.

Having looked at different nouns from this category, let us now see how the nouns were distributed within the texts. The majority of the nouns in this category were in initial position (52.91%). Most of the NPs were subjects (70.08%) and only some were objects (17.45%). Genitive constructions and other grammatical functions were tagged in 12.48% of the nouns. Even though the diversity of topics in this category was much greater than in other text groups, the results are not very different from the ones discussed earlier in the chapter.

Position	Number	%
initial	191	52.91
non initial	170	47.09

Table 5.17: Corpus – noun position (other)
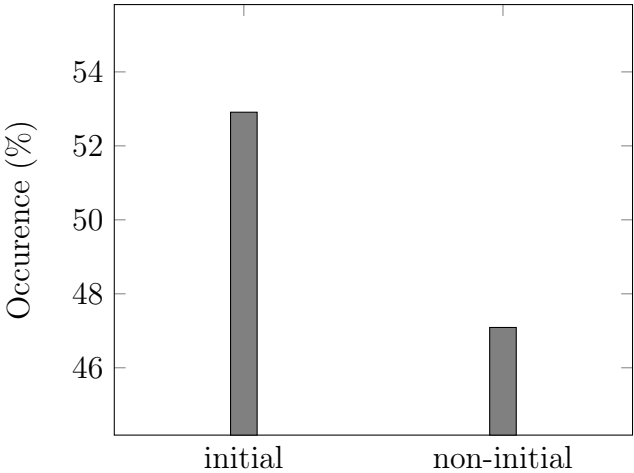
Figure 5.17: Corpus – noun position (other)

Functions	Number	%
subject	253	70.08
object	63	17.45
genitive modifier	14	3.89
other	31	8.59

Table 5.18: Corpus – noun function (other)
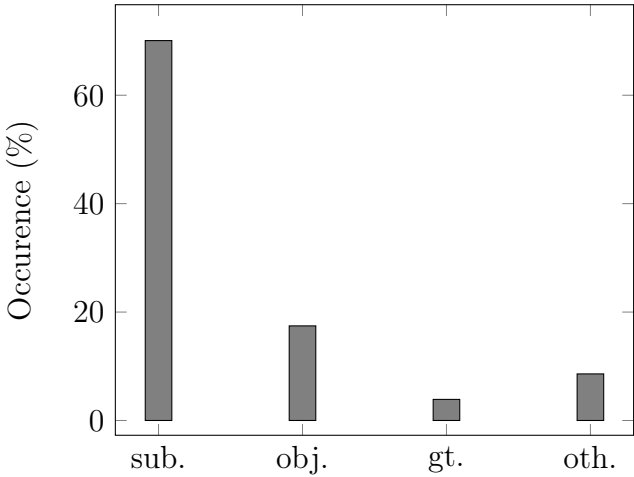
Figure 5.18: Corpus – noun function (other)

Mass nouns   As has been said in chapter 4, generic mass nouns in Norwegian and other Germanic languages can in fact appear in two forms: definite or as bare nouns. However, it is possible to perceive certain mass nouns as *semi-countable*, e.g. 'beer' designates a type of drink, whereas 'a beer' is perceived as a type of beer or one unit of the drink (a glass, a can etc.).

Since the pilot research included a number of mass nouns and the results were in accordance with the assumptions, only a few uncountable nouns were chosen for this analysis. The nouns included:

- sveising – welding
- fjernstyring – remote control
- språkvitenskap – linguistics
- cellulose – cellulose
- demokrati – democracy

As expected, the nouns occurred either as BN or definite nouns. Germanic languages do not really allow for other forms to occur with mass nouns. In (130) the noun 'linguistics' is used only as a BN, as it is an abstract noun.

(130)   a.   **Språkvitenskap** [linguistics-Ø] er det vitenskapelige studiet av språk. Språk kan studeres fra mange forskjellige innfallsvinkler. Noen språkvitere er opptatt av språk som et typisk trekk ved mennesket, altså språk som kognitivt system som skiller mennesket fra andre dyr på jorda. Innen denne formen for **språkvitenskap** er man mer opptatt av hva som er felles for alle språk enn hva som er forskjellig, fordi man studerer det som gjør mennesket til en art som skiller seg vesentlig fra andre arter.

Linguistics is a scientific study of language. Language can be studied from many different perspectives. Some linguists are interested in languages as a typical feature of humankind, namely language as a cognitive system which differentiates people from other animals in the world. In this form of linguistics, one is interested in what is common for all languages, not what is different, as one studies factors that make people different from other species.

Other mass nouns that were tagged in the corpus also occurred as BNs and some of them, such as 'welding' took on a countable form as a definite plural compound noun *sveisemtodene* ('the welding methods').

Familiar nouns and well-established kinds   As has been mentioned in the section 2.2.1.1, certain nouns and noun phrases are sometimes recognised by native speakers as well-established kinds or the so-called familiar nouns. Both notions are synonymous to a certain degree, as has been claimed by (Borthen 2007: 144) for instance. In the literature on genericity, WEKs are mentioned more often than familiar nouns but the two refer to the same concept, namely reference to an object or an animal that can easily be recalled by the hearer based only on the discourse.

What one determines as a well-established kind varies from speaker to speaker but certain tendencies were observed in the corpus material. Apart from the types of nouns

mentioned above, the category WEK can also include:

1. everyday objects,
2. animals kept as pets,
3. objects characteristic for a given culture (here: the Norwegian and Scandinavian/Nordic cultures),
4. objects that, even though countable from a grammatical point of view, are sometimes perceived and/or interpreted as mass nouns (certain plant types, materials, everyday objects etc.),
5. objects that are considered familiar as a result of frequent occurrence (e.g. white wedding dress, Cola bottle etc.).

Some of the nouns considered to be well-established kinds in the corpus included: 'dog', 'fjord', 'knife', 'stave church', 'cottage', 'guitar' and 'telephone'.

(131)   a. **Hunden** var det første husdyret vårt, og er i dag den arten som viser størst variasjon i størrelse, farge, fasong og atferd. (...) Hvordan **ulv** ble til **hund** finnes det flere teorier om, men hypotesen om selvdomestisering står sterkest.

The dog was our first domestic animal and today it is the species that shows the biggest diversity when it comes to size, colour, type and behaviour. (...) There are many theories about how a wolf became a dog but the most popular one concerns self-domestification.

b. **Fjord** brukes om en større, forgrenet innskjæring av havet, oftest lang og smal og omgitt av fjellsider. Noen ganger blir uttrykket også brukt om langstrakte innsjøer, for eksempel Tunnhovdfjorden. I vitenskapelig terminologi er **fjord** betegnelsen på en havbukt dannet ved at en bre har formet og fordypet en tidligere dal. **Fjordene** har bratte, u-formede sider og kan være meget dype, men har oftest en grunnere terskel ved munningen.

[The notion] 'fjord' is used when talking about bigger split of the shore; most often long and narrow, surrounded by mountains. Sometimes the expression is used when referring to long lakes, for instance Tunnhovdfjorden. In scientific terminology, 'fjord' is a name of a bay, created when a glacier formed and deepened a valley. The fjords have steep, U-shaped sides and can be very deep but most often then have a shallow threshold by the opening.

In the example (131a), the noun 'dog' appears both in definite form and as a BN. Also the noun 'wolf' could be perceived as a WEK, even though it is not the main topic of this text. What is more, 'wolf' is also strongly rooted in the Norwegian culture which can influence its interpretation in generic contexts.

The sentence in (131b) describes characteristic properties of this land form, using

both BNs and definite forms. Bare nouns can be interpreted as concepts, in a way *model* fjords, whereas definite plural generics give actual fjords as examples when talking about their characteristic features.

At first sight, WEKs do not differ from *regular* nouns when it comes to the grammatical forms they occur in. The difference is mainly semantic and pragmatic. The issue has been discussed in several works on genericity (see e.g. Carlson and Pelletier 1995, Mari et al. 2013a) but the judgment whether or not a given noun is familiar always depends on speakers' language intuitions.

The nouns proposed as WEKs in this section were consulted with a native speaker of Norwegian but their interpretation is by no means final. This aspect of the study is only an attempt to identify nouns that do not seem to follow the same pattern as other nouns in generic contexts. Nevertheless, the study would have to be designed in a different way in order to show the diversity of WEKs and their interpretations.

The presented examples are meant to show that the phenomenon exists in Norwegian but is highly context dependant. What is more, certain predicates and sentence structures might occur solely with BNs which then obtain a WEK-like reading (e.g. *spille gitar*, 'to play guitar').

The table 5.19 shows an overview of all results. When we compare all of the NP types that appeared in the corpus, we can see that BNs occurred in every category and with very different values, ranging from as low as 19.39% (category 'people') to as high as 73.24% (category 'plants'). BNs are often used when the generic reference is made with a prototypical object.

Also indefinite plural presents a wide range of usage across the texts, ranging from 8.45% (category 'plants') to 49.66% (category 'animals'). The results confirm that, especially in the case of kind reference in such categories as 'animals', this form remains *default* in terms of classic kind-reference.

Forms	Animals		Plants		Tools		Other		People		Total	
	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%
bare nouns	25	17.01	104	73.24	64	52.46	133	36.84	19	19.39	345	39.66
indefinite sing.	0	0.00	0	0.00	5	4.10	16	4.43	10	10.20	31	3.56
definite sing.	32	21.77	14	9.86	34	27.87	74	20.50	31	31.63	185	21.26
indefinite pl.	73	49.66	12	8.45	15	12.30	110	30.47	20	20.41	230	26.44
definite pl.	17	11.56	12	8.45	4	3.28	28	7.76	18	18.37	79	9.08
Total	147		142		122		361		98			
Total (%)	16.90		16.32		14.02		41.49		11.26			

Table 5.19: Corpus – categories, NP types
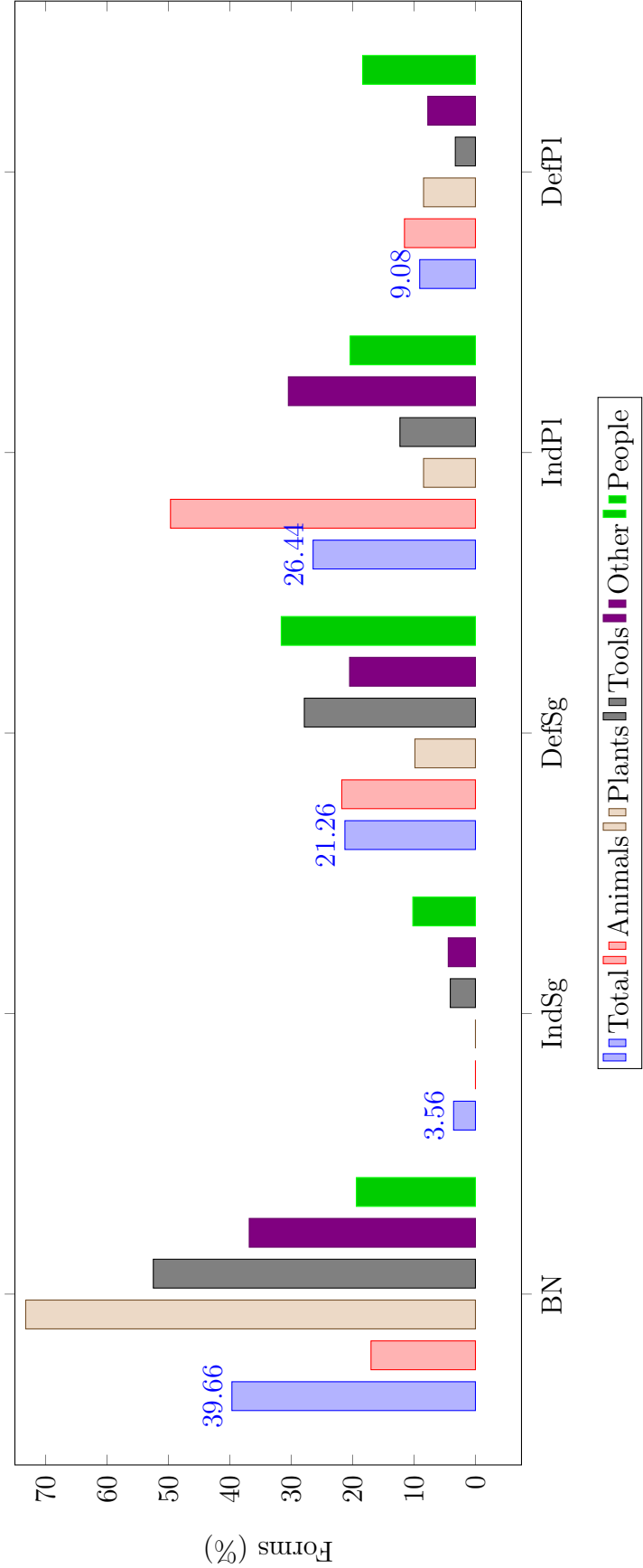
Figure 5.19: Corpus – categories, NP types

5.4 Cognitive analysis

The results from the corpus were analysed according to the cognitive models proposed by Radden and Dirven (2007) and Radden (2009). Types of generic references discussed by the authors do not account for BNs, therefore the model designed by Pettersson (1976) for Swedish was used and modified according to the gathered data.

The types of generic references proposed in the cognitive grammar of English by Radden and Dirven and Radden, include:

1. indefinite singular generic,
2. definite singular generic,
3. indefinite plural generic,
4. definite plural generic.

When we consider a similar model designed for Swedish, bare nouns occur as one of the possible choices in such contexts (Pettersson 1976). Also in this study, the use of BNs was frequent and occurred in every text category. It was therefore possible to differentiate between different meanings of BNs used throughout the corpus.

The table 5.20 presents a simplified matrix of NP types and contexts in which they occurred most often in the corpus.

BN	IndSg	DefSg	IndPl	DefPl
1/ mass nouns	1/ stand-alone objects	1/ mass nouns	1/ kinds	1/ ethnic/social groups
2/ WEKs, familiar NPs	2/ professions	2/ WEKs	2/ groups of people	2/ gregarious animals
3/ quasi mass nouns	3/ prototypes	3/ official functions		3/ subkinds
4/ concepts		4/ kinds		4/ hyperonyms
5/ professions				

Table 5.20: Corpus – generic NPs, a simplified matrix

The list of contexts in which the different NP types were used, is not exhaustive. However, based on the collected material and the conducted analyses, those were the most common contexts, aligning partially with the cognitive models presented in the theoretical chapter.

Bare nouns, not taken into account in the two models of Radden and Dirven (2007) and Radden (2009), were identified in multiple contexts. According tothe graph presented by Pettersson (1976: 121), in Swedish bare nouns are used in *unlimited* contexts (namely indefinite) and with non-specific nouns (generic). A similar statement can also be said about Norwegian where reference to kinds can in certain cases be rendered with BNs, as in the example (125a), repeated here as (132), or in the example provided by Borthen (2003: 31) (repeated here as 133).

(132) **Dronte**, også kalt **dodo** [dodo-Ø], Raphus cucullatus, på Mauritius veide 20–25 kg og fantes i store mengder da europeerne kom dit i 1598.

Dodo, Raphus cucullatus, on Mauritius it weighed 20-25 kg and occurred in large

quantities when the Europeans arrived in 1598.

(133)   a. **Bil** er et kjøretøy.
             car is a vehicle
             A car is a vehicle.

         b. **Datamaskin** er et nyttig hjelpemiddel.
             computer   is a useful tool
             A computer is a useful tool.

The examples below show two different contexts in which a kind can be referred to with BN. The noun 'dodo' in (132) functions as an encyclopaedic entry, a keyword of some sort, and therefore its use without any article is dictated by the text's genre. However, the kind reference is clear.

In the examples in (133), BNs have only one function, namely they denote kinds by referring to unspecific and *unlimited* entities. In this regard, Norwegian follows the Swedish model of Pettersson. In the matrix presented above, such cases are identified as references to concepts. 'Car' and 'computer' in (133) do not refer to any particular entities that would represent the whole kind but rather they denote kinds through concepts of those two objects. A similar use of BNs was observed in the case of professions, where BNs referred to concepts, rather than actual people being teachers, psychologists etc.

When it comes to other uses of BNs in the corpus texts, namely with *quasi* mass nouns, WEKs and familiar nouns, the examples were numerous. *Quasi* mass nouns are nouns that can be perceived as uncountable, as for instance 'moss'. In Norwegian, the noun is countable and it is therefore possible to say 'the mosses'. However, BNs were used frequently in such cases.

The last context in which familiar nouns and WEKs occur as BNs, was also discussed in the scholarly literature on Norwegian nominal system (see e.g. Borthen 2003, 2007; Rosén and Borthen 2017 and Halmøy 2016). This context depends on speakers' interpretation of a given sentence or text but certain tendencies can be observed. For instance, with predicates such as 'to have', 'to get' or 'to buy', familiar nouns occur as BNs. One can for instance say *Jeg skal kjøpe bil* ('I will buy a car'), where the noun 'car' would be interpreted as a concept of a car, not a particular entity.

A similar notion to the concept of an object is its prototype, which is most often rendered with indefinite singular (Radden and Dirven 2007: 108). In such a reference, one entity is chosen by the speaker as a prototype of the whole kind. This form was also used when talking about professions and stand-alone objects such as 'a knife' or 'a screwdriver'.

The definite singular reference occurred most often with mass nouns, WEKs, as well as kinds and official functions, especially those understood as unique (for instance being a kind, a priest etc.). As has been presented in the model of Radden and Dirven and Pettersson, definite singular reference concerns prototypes or refers to distinctive features

of a given kind. Such contexts were indeed present in the corpus.

When it comes to indefinite plural, all models used for this analysis describe the form as *default* when it comes to genericity. Indeed, many *classic* generic references, namely kind-references, are rendered with this form. Also when talking about groups of people, the indefinite plural was used. What is more, in the corpus one more context was observed, namely general truths. As has been mentioned in chapter 2, genericity has two senses: it can refer to kinds or it can express regularly occurring actions in the form of habitual sentences. In the corpus, the second of the senses was also rendered with indefinite plural, making this form the generic default.

The last of the forms depicted in the table 5.20 is definite plural and it occurred in four main contexts. Reference to people was most often rendered with this form, aligning with what Radden (2009) calls for 'delimited generic'. However, in Norwegian definite plural is not at all limited to only this context and was also used when referring to animals perceived as gregarious[7] and to subkinds. This use of the definite plural contrasts with the cognitive model for the English language, where subkinds are usually accessed by definite singular.

Also hyperonyms were described with definite plural. This can be interpreted in a way that a hyperonym is the main class of certain objects or animals, for instance the hyperonym 'vehicle' includes cars, trucks and many others. The same applies to hyperonyms used when describing kinds – 'horse' can have many subkinds and both can be accessed with definite plural.

Despite a much wider use of definite plural generics in the corpus, the models proposed by Radden and Dirven (2007) and Radden (2009) also apply to Norwegian. Delimited generics remain the default form when talking about groups of people, with a few other contexts such as subkinds and hyperonyms.

The cognitive models proposed in this section and the matrix of the NP types used in different contexts in the corpus show that Norwegian generics are slightly more complex than generic references in English. Nevertheless, the forms were used rather consequently in the corpus texts, proving that even though all five NP types can be used generically, they are not interchangeable in each and every context.

In the next section, a statistical analysis of the results will be provided. The goal of the analysis is to show main tendencies when it comes to distribution of the data and the occurrences of the different forms in the corpus. The statistical tests do not answer the questions concerning distribution of the forms in particular contexts and should not be interpreted this way.

---

[7]Some animal species are only *perceived* as gregarious by the speakers' of the language, even though the biological classification might state differently.

5.5   Statistical analysis

Apart from the qualitative description of the corpus data, a number of statistical tests was performed. In order to see how the NP types are distributed throughout the corpus, the data was analysed with the use of descriptive statistics. First, the material was divided into quartiles and minimal and maximal values were calculated, as well as the median and the mean. The results are depicted in table 5.21.

The minimal value of 0 for each of the forms indicates that there are texts in the corpus where one or many of the forms do not appear at all. For instance, there is at least one text that does not have any BNs, at least one that does not have any indefinite singular nouns and so on. The examples of texts with minimal values of 0 for indefinite singular were seen in the category 'animals', where none of the texts contained that form.

The maximal value on the other hand, as the name implies, shows the maximal values of each of the forms in the corpus texts. For instance, there was at least one text where BN occurred 15 times, at least one where indefinite singular occurred three times etc.

The data was then divided into quartiles which show a more precise distribution of each of the forms throughout the corpus. The values from the corpus were ordered from the minimal to maximal and the standard deviation (SD) was calculated. SD shows the dispersion of the data which differs significantly between the forms. For instance, BNs are a lot more dispersed than indefinite singulars. This may be due to the fact that BNs were generally most frequently used throughout the corpus, irrespectively of the text category. The distribution and dispersion of the corpus data can also be seen in figure 5.20.

	BN	IndSg	DefSg	IndPl	DefPl
Min.	0.00	0.00	0.00	0.00	0.00
1st Qu.	1.00	0.00	0.00	0.00	0.00
Median	1.00	0.00	0.00	0.00	0.00
Mean	2.029	0.182	1.088	1.353	0.465
3rd Qu.	3.00	0.00	2.00	2.00	0.00
Max.	15.00	3.00	9.00	12.00	6.00
SD	2.26	0.51	1.58	2.18	1.07

Table 5.21: Corpus – descriptive statistics of generic NP types

The values depicted in the boxplot 5.20 show the descriptive statistics such as minimal and maximal values, mean, median and the outliers, namely the values that do not follow the main tendencies.

The next step consisted in conducting statistical tests in order to compare data distribution for each of the forms. Similarly as in the previous chapter, the assumptions of ANOVA were verified. One-way ANOVA could not be performed as the assumptions of normality of data and homogeneity of variance were not met. The normality of data was
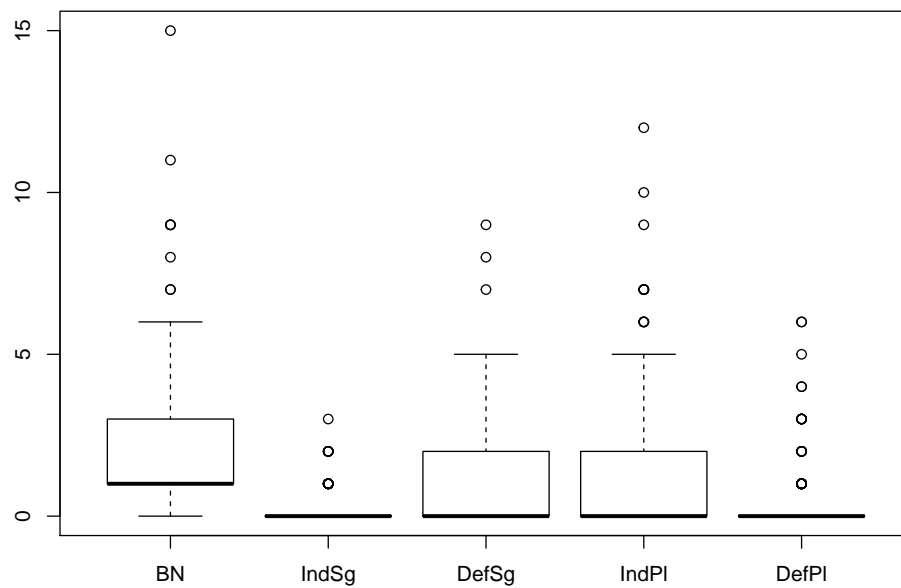
Figure 5.20: Corpus – boxplot of generic NP types

checked with Shapiro-Wilk test, which resulted in p-value of $< 2.2e\text{-}16$. As the p-value falls below 0.05, the data is not normally distributed and the necessary assumption is not met.

The second assumption concerning the homogeneity of variance was verified with Levene's test which also resulted in p-value below 0.05 (see the table 5.22). ANOVA was therefore not suitable for this analysis (Rasinger 2013: 140).

	Df	F value	Pr(>F)	
Group	4	19.58	5.99e-15	***
	845			

Table 5.22: Corpus – Levene's test

Another way of analysing the data in such case is performing the Kruskal-Wallis test, namely a non-parametric test that is used when the assumptions for ANOVA are not met. The analysis resulted in p-value $< 2.2e\text{-}16$ (below 0.05), indicating that there are differences between the tested groups. The groups in the analysis are the five NP types. However, the Kruskal-Wallis test does not show where exactly the differences are. In order to verify that, a post-hoc test is needed. As has been mentioned in section 2.5.1, the post-hoc test of choice in the case of Kruskal-Wallis is Dunn's test. The results of Dunn's test are depicted in the table 5.23.

The Dunn's test is a multiple comparison test that shows the differences of means between the groups, here – the NP types. The values marked in bold designate differences

mean comparison	BN	IndSg	DefSg	IndPl
IndSg	12.47			
p-adjust	**0.00**			
DefSg	5.83	-6.64		
p-adjust	**5.60e-08**	**3.17e-10**		
IndPl	5.93	-6.53	0.11	
p-adjust	**2.86e-08**	**6.70e-10**	1.00	
DefPl	10.49	-1.98	4.66	4.55
p-adjust	**0.00**	0.48	**3.12e-05**	**5.32e-05**

Table 5.23: Corpus – Dunn's test

between the forms. As can be seen, there are no statistical differences only between two pairs, namely 1) definite singular – indefinite plural and 2) indefinite singular – definite plural. This indicated that all other pairs of NP types are different from each other, from a statistical point of view. This was also observed in section 5.4 above, where different contexts were given for each of the NP types. The cognitive interpretation of the results was also confirmed in the statistical analysis, which makes the findings more reliable.

5.6   Conclusions

Generic texts allow to observe the phenomenon in a wider context, as opposed to sentence analysis. As has been shown in this chapter, numerous examples confirm that in Norwegian not only are all NP types allowed in generic contexts but also BNs are quite frequent in generic expressions with countable nouns. Both qualitative and quantitative analyses have shown the diversity of generic nouns and NPs in the corpus.

When it comes to animacy, it is a factor that seems to play a role in expressing generics. In the category 'people', definite generics constituted a great amount of the material, with rather limited use of BNs compared to other text categories. The category 'animals' differed slightly when it comes to the most frequently used NP types. There were no instances of indefinite singular and definite forms were in the minority of all generic expressions. Bare plurals were the most frequent form which supports the thesis that kind-reference is expressed with indefinite plural. Interestingly enough, the text group 'plants', even though grammatically inanimate, was very similar to the category 'animals'. Indefinite singulars did not occur in texts about plants. However, the most frequent NP type was BN which can suggest that some of the plants are considered mass, even though they are grammatically countable.

The matrix of NP types in generic contexts presented in this chapter is based on the models designed by Radden and Dirven (2007); Radden (2009) and Pettersson (1976). In order to account for the Norwegian data, the model was adjusted, especially when it

comes to BNs and definite plural (the so-called delimited generic). Most of the functions presented in the models applied also to Norwegian, such as treating indefinite singular as a concept of an object and definite singular as its prototype. What is more, a few more contexts were observed in the corpus and those were added to the model. Those contexts included for instance using BNs with animals, plants and objects perceived as mass (*quasi*-mass nouns) or describing subkinds with definite plural.

The last part of the corpus analysis consisted in conducting a number of statistical tests. The tests were performed in order to identify the dispersion of the data and differences between the tested groups (NP types). However, the results of the statistical part of this chapter should not be treated on par with the cognitive analysis. The tests allowed to see how the data is distributed and where the main tendencies lay when it comes to occurrences of the forms. The tests do not answer the main research question of this project and are a quantitative element of the analysis.

6   Conducted surveys

6.1   Method

After building and analysing the corpus for the project, an AJT survey was conducted among native speakers of Norwegian. The idea of testing speakers' intuitions and contrasting it with the corpus material was inspired by the study of Oosterhof (2008).[1] This way, the corpus served as a basis for grammaticality judgements utilised in the survey, providing in a way *model* generic sentences.

Grammaticality judgements in a form of a survey are widely used in linguistics. They can be utilised not only to check speakers' intuitions but also to confront the theory from normative grammar books with the actual use of a given language. However, many researchers have pointed out that AJT surveys come with a number of issues. One of them is the way a given task is designed. For instance, asking participants for their judgement on a given matter can give promising results, whereas asking them to justify *why* they judge a given sentence (in)correct seems futile (Schütze 2016: 56-57). The idea behind most grammaticality judgements is that the answers are very much intuition driven which means that it might be difficult for the participants to explain why they think a given sentence is correct or not.

One of the solutions one can apply in order for grammaticality judgements to give reliable results is to introduce ranking of the test items (Schütze 2016: 58). Such rankings can have different scales, more or less detailed, but their goal is always the same – to provide the participants with a wider range of answers. This way the participants are able to grade their answers according to the provided scale but without the need to explain why they made a given choice.

Many surveys that do employ sophisticated rating scales are conducted on smaller groups of participants (cf. the studies of Oosterhof 2008 and Ionin et al. 2011 among others). Since the first survey was conducted on a large number of participants and it did not include a grading scale, the second survey was designed in a slightly different way.

The main idea behind this survey was to recruit a smaller number of participants and to be able to prevent non-native speakers from taking part in the survey. A form

---

[1]For further discussion on survey methodology see section 2.5.1.

published online can be answered by anyone and it is therefore impossible to state with absolutely certainty whether each and every answer is an actual answer to the questions or whether it was done haphazardly by someone not even interested in the study. Recruiting participants in person minimised that risk significantly.

The analysis model used in this study Similarly as in the pilot survey, the test sentences were analysed according to the model proposed by Leslie et al. (2011). In each of the sentences, a type of generic predication was identified. The types are then analysed together with the judgements from the survey.

Since the corpus texts were analysed according to the models proposed by Radden and Dirven (2007); Radden (2009) and Pettersson (1976), also here the model will be utilised but only in order to develop the matrix of generic NP types and the contexts they are used in. This way, the models developed originally for English (Radden and Dirven; Radden) and Swedish (Pettersson) can be adjusted to Norwegian and based on empirical data that illustrates actual language use.

6.1.1 Survey structure and used tools

The survey based on the corpus material was modified in comparison to the pilot study. First of all, the number of participants was limited to 100 people who were chosen in person and not through a website. The first survey was conducted in order to verify whether or not genericity in Norwegian is as complex as it seems to be, whereas the second survey was meant to test native speakers' intuitions in reference to the generic source material.

In the first survey conducted for this project, the respondents received short texts where generic sentences contained gaps. The gaps were filled with a noun in one or more grammatical forms. In the second survey however, the test items were generic sentences retrieved from the corpus. The test items and filler items were chosen from different text categories. In addition, 5 filer items were placed in the survey not to suggest the actual topic of the study. The test items included the following NPs:[2]

1. Birkebeinere – the Birkebein Party
2. diakon – deacon
3. måke – seagull
4. skate – batoidea
5. flaggermus – bat
6. oter – otter
7. sjiraff – giraffe
8. krokodille – crocodile

---

[2]All NPs are listed as BNs. For full survey see the Appendix.

9. hai – shark
10. humle – bumble bee
11. ål – eel
12. torvmose – peat moss
13. kniv – knife
14. meander – meander
15. stavkirke – stave church
16. kulepenn – ballpoint pen
17. gen – gene
18. fiskekrok – fish hook
19. skrue – screw
20. gitar – guitar

The sentences with NPs 1-2 were retrieved from the category 'people', 3-11 from the category 'animals', 12 from the category 'plants', 13 from the category 'other' and the 14-20 from the category 'other'. As has been mentioned before, the sentences chosen for the survey had to be generic without a wider context and their forms needed to be diverse, for instance with generic nouns both in initial and non-initial positions. The subjects of the sentences were therefore not the main priority.

The filler items in the survey included 5 sentences with the following NPs:

1. sjefdirigent – leading conductor
2. husmann – cotter
3. knivspiss – knife tip
4. sysselmann – governor
5. telefon – telephone

The NPs in the test items and filler items occurred both in initial and non-initial positions in oder to provide more diversity and not to suggest the subject of the survey to the respondents.

When it comes to the software used to create and publish the survey, open source tools were used. The core of the software was R, whereas the graphical interface was designed with the use of Shiny Applications. All survey results were saved directly on Dropbox. The choice of such tools over for instance Google Forms that was utilised to build the first survey, was caused by the need to adjust the design of the form. Google does provide certain options when it comes to customisation but they are rather limited. Also, the answers from Google Forms are saved in one file and without the division into columns which makes the analyses a lot more complicated. The way of saving the data was customised with R and Shiny package.

Using R and Shiny allowed also to adjust the survey interface for all screen sizes.

Grading scale for the sentences was always displayed horizontally, which made it consistent with the provided grading instruction and therefore more intuitive for the participants. The participants could fill out the survey on the provided laptop and tablet or use a link or a QR-code to open the survey on their own electronic devices.

The survey was divided into two main parts. The first one contained questions about the participants, such as their sex, age and education. The second part consisted of a list of sentences with all possible NP types. Each of the forms had a grading scale next to it, ranging from 1 to 3, where 1 meant 'correct', 2 'acceptable' and 3 stood for 'incorrect/awkward'. The grading scale was simplified in comparison to Oosterhof's study mentioned before. The reason for this was the length and complicity of the survey. Since the survey was relatively long, adding a complex grading scale could discourage the participants or even result in random answers. It is also worth mentioning that because of the available resources the survey could not be conducted in a laboratory or other isolated area where the participants would not be distracted or interrupted. Adjusting the design of the survey to the conditions was therefore necessary. Nevertheless, two participants resigned from fulfilling the form, judging it too complex despite the provided instruction. Other respondents did not report similar issues.

6.2 Respondents

The second survey, similarly as the one utilised in the pilot study (see chapter 4), also contained a number of questions about the participants. The respondents had to specify their age, sex, education, origin and the language they usually swear in. The control question about the language used for swearing was added in order to make sure that Norwegian was the first language of all participants. In the first survey, many participants understood the control question as 'In what languages *can* you count', listing often all foreign languages they spoke in. What is more, swearing is an emotion-driven behaviour and it is therefore a lot more plausible that a speaker swears in her native language. However, especially in the younger generation, it is common to swear in English and the possibility of such answers to the question was also taken into account. This was verified by the fact that none of the respondents chose the location 'other' as their place of origin so potential answers about swearing in other languages were not critical in identifying native speakers of Norwegian.

Age	Women	Men	Total
18–19	13	2	15
20–24	32	27	59
25–30	11	12	23
31–40	0	2	2
41–50	1	0	1
Total	57	43	100

Table 6.1: Survey 2 – age of the respondents
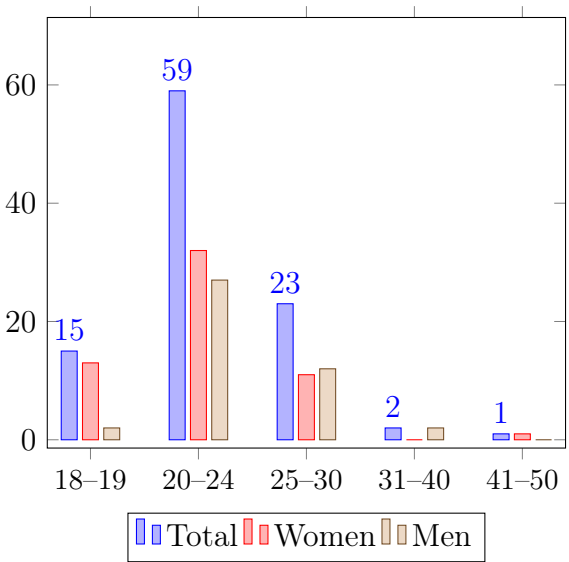
Figure 6.1: Survey 2 – age of the respondents

The table 6.1 and the graph 6.1 show age and sex of the respondents. 100 respondents took part in the survey, aged mostly 18-30. People from other age categories were rather reluctant and did not agree to devote their time to the study.[3] Similarly as in the first survey, also here women dominated in many age categories, except for the groups aged 25-30 and 31-40.

Education	Women	Men	Total
primary school	3	0	3
high school	31	15	46
university	23	28	51

Table 6.2: Survey 2 – education of the respondents
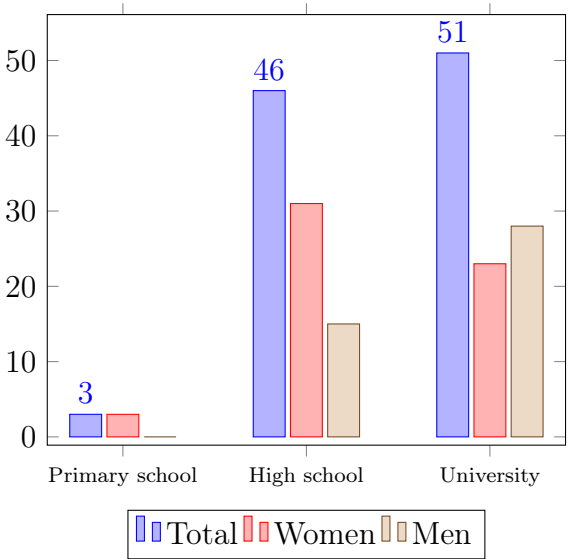
Figure 6.2: Survey 2 – education of the respondents

---

[3] All participants were informed about the duration of the survey (15-20 minutes) and that it concerned Norwegian language.

As can be seen in the table 6.2 and the graph 6.2, 51 respondents stated that they had higher education, 46 marked high school education and 3 people chose primary school from the answers. Given the high number of young people who most probably were studying at the University of Tromsø at that time, the choice of 'high school' in so many cases is not surprising – participants who had not yet completed a university degree chose this option.

The table 6.3 shows the origin of all participants. As can be seen on the map 6.3,[4] the majority of the people who took part in the survey were from Northern parts of the country. This was not surprising, given that the survey was conducted in Tromsø. However, there was a number of people from other regions of the country, e.g. Trøndelag, Vestfold and Østfold. Nevertheless, the representation of different regions of Norway was much less pronounced than in the pilot survey.

Similarly as with the pilot survey, also in this study the sentences chosen for the survey were written in *bokmål* and all respondents were informed about it beforehand. A few people asked whether they should answer the way they would in their dialects and they were asked not to do so but rather opt for more standardised answers. It shows that many participants were aware of the differences between *standard* Norwegian[5] and the dialects.

Compared to the first survey, the number of participants was significantly smaller. Since the respondents were recruited in person and not through an online forum, it was possible to verify beforehand whether they were native speakers of Norwegian. This was then confirmed by the places of origin that the participants had chosen.

---

[4]In 2019, when the survey was conducted, there were 18 municipalities in Norway as the country is in the process of merging certain municipalities; cf. the map in chapter 4.

[5]I use this term in its colloquial meaning as there is no such thing as 'standard Norwegian'. There are written standards of the language but those do not correspond to e.g. English 'RP', German 'Hochdeutsch' or Swedish 'rikssvenska'. For further discussion see chapter 7.

Regions	Respondents
Akershus	2
Aust-Agder	1
Buskerud	0
Finnmark	12
Hedmark	1
Hordaland	2
Møre og Romsdal	1
Nordland	12
Oppland	2
Oslo	2
Østfold	4
Rogaland	2
Sogn og Fjordane	1
Telemark	2
Troms	45
Trøndelag	7
Vest-Agder	0
Vestfold	4

Table 6.3: Survey 2 – origin of the respondents
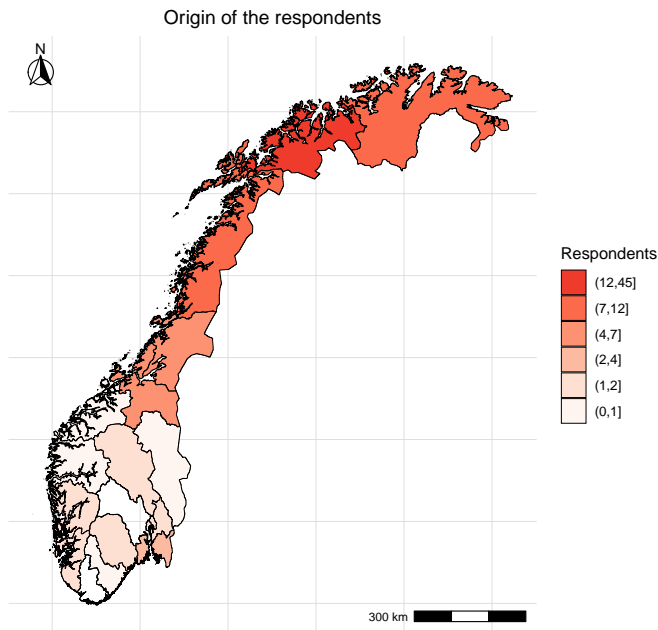
Figure 6.3: Survey 2 – origin of the respondents

6.3   General results

The sentences from the corpus chosen for this survey were in a way *model* generic sentences, namely their forms were not significantly different from what the previous analyses have shown. One of the nouns chosen for the AJT survey was a well-established kind, namely 'stave church'. However, stave churches are deeply rooted in the Norwegian culture so interpreting the noun as familiar is not context or intuition dependent in this case. The nouns chosen for the study and the forms they originally occurred in are presented in the table 6.4.

Noun	Form
birkebeinere	indpl
diakonen	defsg
måkene	defpl
skater	indpl
flaggermusene	defpl
oteren	defsg
sjiraffen	defsg
krokodiller	indpl
haier	indpl
humlene	defpl
ålen	defsg
torvmosene	defpl
kniver	indpl
meandere	indpl
stavkirke	BN
kulepennen	defsg
gener	indpl
fiskekroker	indpl
skruer	defpl
gitaren	defsg

Table 6.4: Survey 2 – chosen nouns

As can be seen, only one noun among the twenty chosen for the analysis occurred as BN and it is in fact 'stave church', considered as a *classic* WEK (see section 2.2.1.1). The majority of the sentences are indefinite plurals which, as has been said in previous chapters, can be considered default generics in most Germanic languages. Also in Norwegian, indefinite plurals are very often used when referring to whole kinds (see chapters 4 and 5). A number of example sentences occurred originally in the corpus with definite plural NPs, such as 'the seagulls' or 'the peat mosses'.

The respondents in the survey were supposed to fill out an online form where all test items and filler items were listed. All sentences had gaps with possible NP types listed below with the grading scale. The general results can be seen in the table 6.5 and the graph 6.4.

Relatively high percentage of plural forms judged as correct corresponds to the previous analyses in this project. The status 'acceptable' was chosen rarely in all cases which can mean that the respondents were opting rather for the 'correct-incorrect' scale, ignoring the 'acceptable' judgement. This can be observed in the graph 6.4, where the sum of all judgements in a given category are marked with blue.

Forms	1-Correct		2-Acceptable		3-Incorrect	
	Number	%	Number	%	Number	%
bare nouns	565	5.65	222	2.22	1213	12.13
indefinite sing.	767	7.67	260	2.60	973	9.73
definite sing.	1065	10.65	344	3.44	591	5.91
indefinite pl.	1279	12.79	348	3.48	373	3.73
definite pl.	1058	10.58	467	4.67	475	4.75
Total	4734		1641		3625	
Total (%)	47.34%		16.41%		36.25%	

Table 6.5: Survey 2 – NP types
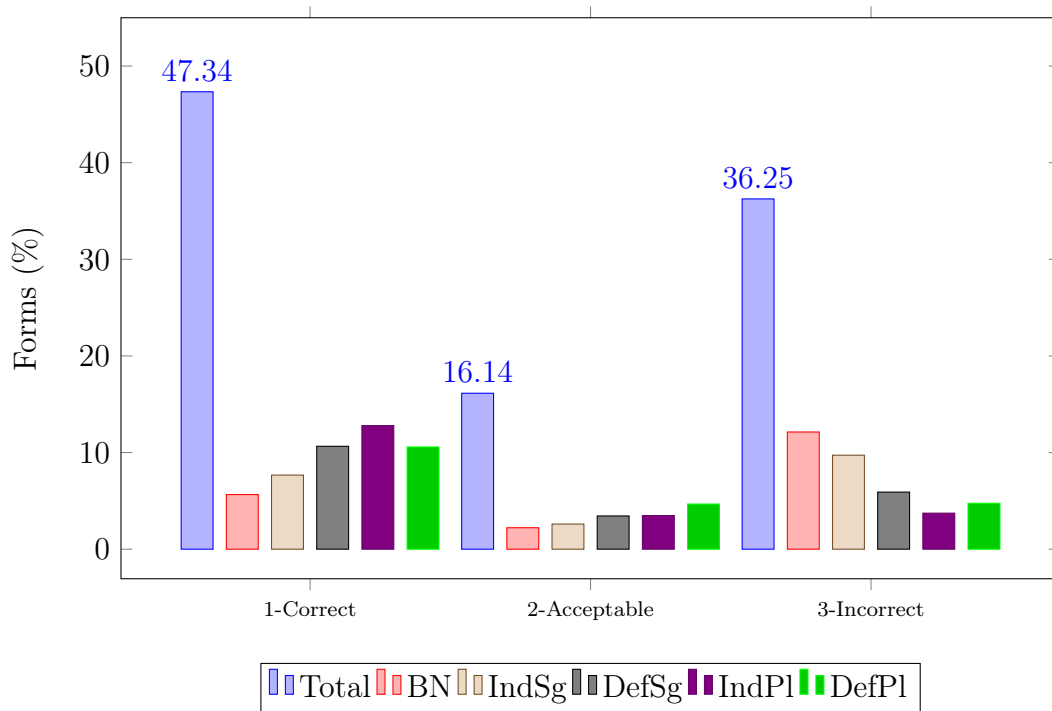
Figure 6.4: Survey 2 – NP types

In the majority of the examples, the participants judged the *original* corpus forms as correct and/or acceptable, providing also other options (see the table 6.6). In just a few of the sentences, the answers differed significantly from the corpus material (see the discussion below). Such a high agreement suggests that even though Norwegian generics may be expressed in a number of different ways, certain tendencies can be found. For instance, limited use of indefinite singular applied mainly to animate nouns, even though certain participants listed this form as acceptable or even correct. Certain nouns such as plants perceived as mass-like and animals considered gregarious, appeared often as definite plurals in generic context – similarly as in the corpus.

	BN			IndSg			DefSg		
	1-Correct	2-Acceptable	3-Incorrect	1-Correct	2-Acceptable	3-Incorrect	1-Correct	2-Acceptable	3-Incorrect
1- måke	3	6	91	51	24	25	49	25	26
2- skate	48	13	39	30	25	45	38	25	37
3- meander	39	17	44	46	18	36	27	33	40
4- kniv	14	10	76	28	21	51	24	19	57
5- fiskekrok	15	4	81	20	9	71	36	16	48
6- krokodille	10	11	79	46	21	33	69	15	16
7- hai	20	5	75	20	12	68	47	21	32
8- stavkirke	42	17	41	71	12	17	61	15	24
9- kulepenn	9	7	84	89	7	4	82	13	5
10- gen	12	19	69	60	7	33	36	15	49
11- birkebeiner	21	9	70	15	7	78	26	8	66
12- diakon	18	10	72	80	11	9	75	15	10
13- flaggermus	49	14	37	73	9	18	68	13	19
14- oter	44	13	43	20	5	75	74	15	11
15- sjiraff	35	12	53	9	10	81	69	21	10
16- torvmose	52	10	38	26	16	58	64	10	26
17- humle	19	12	69	26	12	62	51	22	27
18- skrue	21	11	68	32	10	58	14	9	77
19- gitar	22	16	62	14	10	76	84	13	3
20- ål	72	6	22	11	14	75	71	21	8
Total	565	222	1213	767	260	973	1065	344	591
Total (%)	5.65%	2.22%	12.13%	7.67%	2.60%	9.73%	10.65%	3.44%	5.91%

	IndPl			DefPl		
	1-Correct	2-Acceptable	3-Incorrect	1-Correct	2-Acceptable	3-Incorrect
1- måke	66	22	12	73	21	6
2- skate	47	24	29	25	26	49
3- meander	65	25	10	30	34	36
4- kniv	80	9	11	58	19	23
5- fiskekrok	77	11	12	37	31	32
6- krokodille	93	7	0	69	25	6
7- hai	88	4	8	70	21	9
8- stavkirke	46	25	29	41	29	30
9- kulepenn	34	24	42	30	28	42
10- gen	72	13	15	62	18	20
11- birkebeiner	62	21	17	79	13	8
12- diakon	18	21	61	24	20	56
13- flaggermus	53	16	31	67	23	10
14- oter	69	24	7	53	25	22
15- sjiraff	82	14	4	66	26	8
16- torvmose	38	23	39	52	21	27
17- humle	79	17	4	84	13	3
18- skrue	90	7	3	42	27	31
19- gitar	56	24	20	47	25	28
20- ål	64	17	19	49	22	29
Total	1279	348	373	1058	467	475
Total (%)	12.79%	3.48%	3.73%	10.58%	4.67%	4.75%

Table 6.6: Survey 2 – results

Bare nouns, both in the corpus and in the survey, represent an interesting status. In the corpus, many encyclopaedic entries were BNs, as well as nouns considered somehow familiar to native speakers of Norwegian. In the survey, bare nouns were referred to as 'correct' in a number of examples, indicating that it might in fact be another *default* generic form, together with indefinite plural. The nouns judged correct and/or acceptable as BNs included for instance 'bat', 'otter', 'eel' and 'peat moss'. 'Stave church' is an example of a typical well-established kind occurring without any article, both in the original texts and in the respondents' answers.

The majority of the examples were judged in a way that corresponds to the model sentences from the corpus. For instance, a number of NPs used in the survey occurred as indefinite plurals and the respondents also graded this form very high, suggesting often that definite plural could be used as well. Even though indefinite singular did not appear at all in the texts about animals, the participants listed this form often as acceptable or even correct. The examples will be discussed in greater detail according to the category they occurred in. The categories with most nouns in the survey will be discussed first.

Category 'people' Only one sentence was chosen from the text category 'people' and it was the example presented in (134). The original NP type used in the corpus was definite plural, being a classic delimited generic used when talking about groups of people, as has been shown in the model by Radden (2009).

(134) Birkebeinere [Birkebein.party-PL] var en politisk gruppering i Norge i høymidde-lalderen.

The Birkebein Party was a political group in Norway in high Middle Ages.

What is interesting in this example is that both plural forms were rated very high as correct and acceptable, whereas BN and the singular forms scored much lower and were considered incorrect/awkward by many respondents.

The use of both plural forms and low acceptability of BN and the singular forms aligns with the models proposed for English and Swedish, as well as with the model modified for the Norwegian data in the previous chapter. Talking about groups of people in Norwegian, in contrast to professions for instance, is usually connected with plural forms, referring to collectivity of the subject in question. This was presented in the animacy model by Rosenbach (2008) and discussed in detail in chapter 5.

Category 'animals' Among the nouns that originally were indefinite plurals in the corpus, certain were judged as correct and/or acceptable also with other forms, e.g. definite singular (correct with the noun 'crocodile' and 'shark'). The results show also that a lot of participants opted rather for *extreme* grades (correct-incorrect), rather than more neutral 'acceptable' in many of the examples.

The examples in (135) show two cases where the respondents provided more forms as correct and acceptable than those in the original sentences. The noun 'crocodile', originally a BP, was judged as correct in this form 93 times and 69 'correct' grades were given to definite plural. None of the participants judged the indefinite plural to be awkward or incorrect. The noun 'shark' was graded in a similar manner, with BP scoring 88 answers as correct. Definite plural 'the sharks' was considered correct 70 times.

(135)   a. Krokodiller [crocodiles] lever i vann og finnes i tropiske områder over hele kloden.

          Crocodiles live in water and they can be found in tropical climate in the whole world.

    b. Hos haier [sharks] som lever av fisk og lignende er det utviklet skjære- eller gripetenner.

          Sharks that live on fish and similar food have developed flattened or needle-like teeth.

Both sentences in (135) that are considered correct as definite plural generics, concern animal species. As has been stated in chapter 5, this NP type, when used generically, may suggest that the animals in question are seen as gregarious. Sometimes this is true from a biological perspective and sometimes it is just a result of folk taxonomy.[6]

The survey sentence about bats in (136) contains definite plural in the original sentence and two other forms judged as correct, namely indefinite and definite singular. In its original form, the sentence can be interpreted in a way that bats as gregarious animals, are often seen in groups. However, other NP types proposed as correct by the respondents suggest conceptual and prototypical readings of the noun (indefinite and definite forms respectively).

(136)   Flaggermusene [bats-DEF] har et enormt næringsbehov.

      The bats have an enormous appetite.

A similar interpretation of animals being gregarious was also observed in the example with bumble bees. The two plural forms were judged most often as correct and the original sentence contained definite plural .

Reference to kinds seems to have the widest range when it comes to the NP types chosen in such contexts. Another interesting noun from this category is the sentence in (137). The species 'eel' is described with definite singular in the original sentence. This NP type also scored high in the survey and was graded 71 times as correct and 21 times as acceptable. Also BN was chosen many times as correct, whereas the two plural forms scored slightly lower. Nevertheless, almost all NP types, except for indefinite singular, was chosen either as correct and/or acceptable.

---

[6]In the study of Jacoby et al. (2014) social behaviour of certain shark species was described.

(137)   Ålen finnes i hav og ferskvann fra Nord-Afrika og middelhavslandene i sør til Kvitsjøen i nord.

  The eel lives in seas and fresh waters, from Northern Africa and Mediterranean countries in the South to the White Sea in the North.

As has been mentioned in previous chapter, predicates used in generic sentences influence the choice of possible NP types. The predicate 'to live in seas' from the example (137) suggests a kind-reference. Therefore all forms, except for the BN, were judged as possible in this context. Reference to a kind can be expressed either through a reference to a concept or a prototype (singular forms) or by referring to a certain number of members from the kind (plural forms). Either way, the sentence is unambiguous because of its generic predicate.

Many examples with animals were judged as correct and/or acceptable with plural forms, whereas singular forms were chosen in fewer cases. This can refer to the interpretation of a given sentence. As has been, singular forms can be understood as concepts or prototypes of a given species, making it awkward/incorrect with certain predicates. The lack of singular forms in the category 'animals' in the corpus might have been the result of the predicates chosen by the authors of the texts.

Category 'plants'   The texts from this category followed similar patterns as the texts about animals in the corpus, namely there were no uses of indefinite singular. In the survey, one example from this category was chosen and it concerned the NP 'peat moss'. The fact that 'moss', even though grammatically countable in Norwegian, is considered almost as an uncountable noun can be seen in (138a), where definite plural generics are used.

(138)   a.   Torvmosene [peat.moss-DEF-PL] har en eiendommelig bygning.

      Peat moss has a homogeneous structure.

'Moss' is a plant of homogeneous structure can be difficult to count, despite its grammatical countability. On the other hand, the NP from the sentence is a subkind, namely a certain type of moss which in the model sentence is rendered with definite plural. As has been shown in the cognitive model presented in the previous chapter, definite plural generics can in certain cases access subkinds.

When it comes to the respondents' answers, no clear tendencies were observed. None of the proposed NP types was chosen as correct or incorrect more often than the other. This may suggest different interpretations of the example. On the one hand, the plant could have been perceived as a prototype or a concept of 'peat moss' and therefore quite a few respondents opted for BN and definite singular. On the other hand, the plant was perceived as consisting of many small elements or even as mass-like by the respondents who chose the plural forms. The predicate used in the sentence, namely 'to

have a homogeneous structure' can only refer to a whole kind, not one particular plant. Therefore, singular generic references are interpreted as conceptual or prototypical, as in the models of Radden and Dirven (2007) and Pettersson (1976).

Category 'tools'   Only one sentence from the category 'tools' was chosen for the survey and it concerned the noun 'knife'. As can be observed in (139), the generic reference in the sentence is understood as a reference to a subkind, namely certain knives produced by different companies. In the original sentence from the corpus, both references are rendered with indefinite plural.

(139)   I Norge har kniver fremstilt ved småindustri på Toten vært av anerkjent kvalitet, likeså kniver fra Geilo.

   In Norway, knives fabricated by a small company in Toten have been know for their quality, similarly as knives from Geilo.

   Many respondents graded BN and the singular forms as incorrect/awkward, whereas the two plural forms were preferred by the majority. Indefinite singular was chosen most often, confirming the original interpretation of the sentence. Definite plural form was graded correct 58 times and acceptable 19 times, suggesting either specific reference or reference to a subkind, namely particular types of knives produced in Toten and Geilo. As has been discussed in the previous chapter, in Norwegian subkinds can be accessed by definite plural. However, predicates used with definite plural and a wider context of a sentence might influence the interpretation.

Category 'other'   7sentences in the survey were retrieved from the category 'other' in the corpus. The only noun that occurred as BN in the original material was 'stave church', considered a well-established kind in Norwegian. In most cases, WEKs take on definite forms or they are used as bare nouns in generic texts. However, in the survey many answers of 'correct' and 'acceptable' were given to all five NP types with indefinite and definite singular scoring the highest (71 and 61 answers respectively). Bare noun was graded as 'correct' 42 times and 17 times as 'acceptable'. The fact that a lot of the participants opted for the definite singular (and partially BN) confirms that the noun is recognised as familiar.

(140)   Stavkirke [stave.church-Ø] er en høyt utviklet kirketype, oppført i reisverk av tre, kjent fra norsk kristen middelalder.

   A stave church is a highly developed type of churches, constructed with wooden staves, originating in Christian medieval Norway.

   The case of 'stave church' is an interesting example of a well-established kind that is characteristic for a given culture. Its use is very often conceptual (BN) or prototypical

(definite singular), confirming the existing theories on WEKs and familiar nouns (see e.g. Borthen 2007; Rosén and Borthen 2017).

6.3.1   Cognitive model

Similarly as in the previous chapter, a matrix of generic NP types from the survey was created, based on the cognitive models of Radden and Dirven (2007) and Radden (2009). Due to a smaller amount of the material in this survey, there are fewer contexts available for each of the NP types. However, some of them are the same as in the matrix based on the corpus (see table 6.7 below).

BN	IndSg	DefSg	IndPl	DefPl
1/ WEKs, familiar NPs	1/ prototypes	1/ prototypes	1/ kinds	1/ ethnic/social groups
2/ concepts	2/ *kinds	2/ WEKs	2/ ethnic/social groups	2/ gregarious animals
3/ *kinds		4/ *kinds		3/ mass-like plants

Table 6.7: Survey 2 – generic NPs, a simplified matrix

One of the interesting aspects of this model is that, according to the native speakers, reference to kinds can be expressed with every NP types. However, the types of kind-references differ slightly between the forms – this was marked in the table with an asterisk. Reference to kinds can be achieved through concepts and prototypes, namely with BN and the two singular forms. In the model described in Radden (2009: 224), kind generic is assigned only to definite singular and interpreted as a *prototypical subtype of a well-established type*. In Norwegian however, this function can be rendered with two singular forms. BNs have a similar role with the difference that they do not rely on prototypes but concepts. Concepts, same as BNs, are neutral when it comes to number marking. Therefore we cannot say that a kind reference expressed this way is done by referring to an example entity of a kind or a number of such entities.

Kind reference rendered with indefinite plural is a classic generic reference and it was often chosen by the respondents in sentences with kind-predicates. When it comes to definite plural generics, kind-reference is also possible but with certain restrictions. As has been observed both in the corpus and in the AJT survey, reference to gregarious animal species was most often expressed with this form. Interestingly enough, such classification of certain animal species is done often in accordance with folk taxonomy, not a biological classification.

Another interesting aspect that can be observed in the model, are NP types used with familiar nouns and WEKs. First of all, BNs are used often in this context and then they function as concepts, as has been seen for instance in the example (6.3) about 'stave church'. Such reference is neither singular nor plural from a strictly grammatical point of view. When it comes to cognitive interpretation however, one may argue that a concept of a stave church consists of a singular entity that is utilised in a generic reference.

The use of definite singular with WEKs and familiar nouns is a bit different. On the one hand, it can be connected to prototypes (cf. Pettersson 1976), where a singular entity denotes the whole kind. On the other hand, WEKs rendered with definite singular are sometimes perceived as proper names (cf. Carlson 1982; Carlson and Pelletier 1995). WEKs such as 'stave church' in Norwegian, are well-rooted in the language and culture.

The biggest change from the original model of Radden (2009) concerns definite plural form. In the model designed for English, delimited generics can be used only when talking about people. In Norwegian however, delimited generics, even though still limited, can be used in a wider range of contexts. One of the contexts remains the same as in English and it concerns social and ethnic groups. The two other contexts that are possible in Norwegian concern animals and plants. Similarly as in the corpus texts, also here the respondents judged sentences about gregarious animals and mass-like plants as correct when uttered with definite plural. Both these concepts assume that the subject in question contains more than one entity, e.g. gregarious animal species imagined as a flock of animals or moss consisting of many strands and forming therefore a mass-like plant.

The answers provided by the respondents were in most cases the same as in the original sentences from the corpus texts. Additional contexts provided by the respondents allowed to develop the proposed model further. A discussion of both model designed for Norwegian, namely one based on the corpus texts and the other on the AJT survey, will be discussed in greater detail in chapter 7.

6.3.2 Statistical analysis

The statistical analyses in this chapter are divided into two sections. In the first one, all the answers marked as 'correct' are analysed, and in the second section the answers marked as 'acceptable'. The statistical tests can provide a broader perspective on the values (namely the number of answers in each of the survey questions) but they cannot give answer to the main research question – which NP types are used in certain generic contexts. Similarly as in the previous chapter with the corpus data, quantitative analyses can indicate whether there are statistically significant aspects of the analysed data such as differences between the NP types.

6.3.2.1 Correct

The first step of analysing the data was to do conduct descriptive statistics which include the division of the data into quartiles, calculating minimal and maximal values, as well as mean and median. The results are presented in table 6.8 and boxplot 6.5.

The minimal and maximal values show the extreme scores for each of the NP types – the lowest and the highest values respectively. For example, the minimal value of 3 for BN means that in at least one of the survey questions only three people judged this form

	BN	IndSg	DefSg	IndPl	DefPl
Min.	3.00	9.00	14.00	18.00	24.00
1st Qu.	14.75	20.00	36.00	51.50	40.00
Median	21.00	29.00	56.00	65.50	52.50
Mean	28.25	38.35	53.25	63.95	52.90
3rd Qu.	42.50	53.25	69.50	79.25	67.50
Max.	72.00	89.00	84.00	93.00	84.00
SD	18.19	24.70	21.27	20.07	18.22

Table 6.8: Survey 2 – data 'correct'
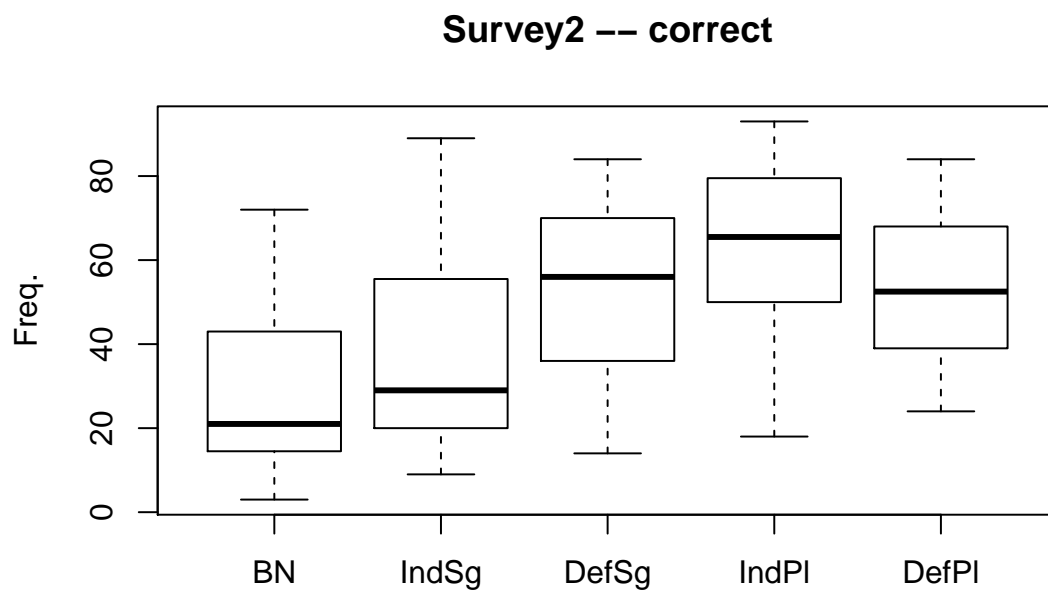
Figure 6.5: Survey 2 – boxplot 'correct'

as correct. When it comes to the maximal value of 72 for BN it indicates that there was at least one survey question where 72 respondents judged BN as correct. The values for other NP types were presented in the same manner.

In the table 6.8, Standard Deviation (SD) shows the dispersion of the data in the analysed groups. As can be observed, indefinite singulars marked as 'correct' demonstrate a slightly bigger dispersion than other NP types. Simply put, the values of indefinite singular rated as 'correct' are more diverse, namely they have a wider range in relation to the mean.

Similarly as in the previous chapters, the data was also tested for homogeneity of variances and normality in order to verify which statistical tests could be performed.[7] The first step was to test for homogeneity of variance and this was done with the use of Levene's test (see table 6.9). Since the p-value was above the significance level of 0.05, there was no evidence that the variances across the tested groups were significantly different from each other.

	Df	F value	Pr(>F)
Group	4	0.62	0.65
	95		

Table 6.9: Survey 2 – Levene's test, 'correct'

After conducting the Levene's test, the normality of data was verified with Shapiro-Wilk test. It resulted in p-value at 0.17, proving that the requirement concerning the normality of the data was not violated. In such a case a parametric analysis, namely ANOVA, could be performed.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Forms	4	15740	3935	9.244	2.37e-06	***
Residuals	95	40439	426.0			

(***) significant at 0.001 level.

Table 6.10: Survey 2 – ANOVA, 'correct'

ANOVA have shown statistically significant results with p-value below 0.001. This indicates that there are differences between certain of the tested groups, namely NP types. In order to verify where the differences lay, a post hoc test had to be conducted. The post hoc test of choice used throughout this study in the case of ANOVA, is Tukey HSD.

Tukey HSD is a multiple pairwise comparison where the groups are compared to each other in pairs, namely BN vs indefinite singular, BN vs definite singular and so on. Differences between four pairs of NP types were found – they are marked with bold in the table 6.11, whereas on the graph 6.6 the differences are indicated by the bars that do not cross the reference line at 0.

---

[7]For statistical methodology see section 2.5.1.

	BN	IndSg	DefSg	IndPl
IndSg	-10.00			
p-adjust	0.53			
DefSg	-25.00	-14.90		
p-adjust	**0.002**	0.16		
IndPl	-35.70	-25.60	-10.70	
p-adjust	**0.00**	**0.002**	0.48	
DefPl	-24.65	-14.55	0.35	11.05
p-adjust	**0.003**	0.18	1.00	0.44

Table 6.11: Survey 2 – Tukey HSD, 'correct'

95% family–wise confidence level

Differences in mean levels of Forms
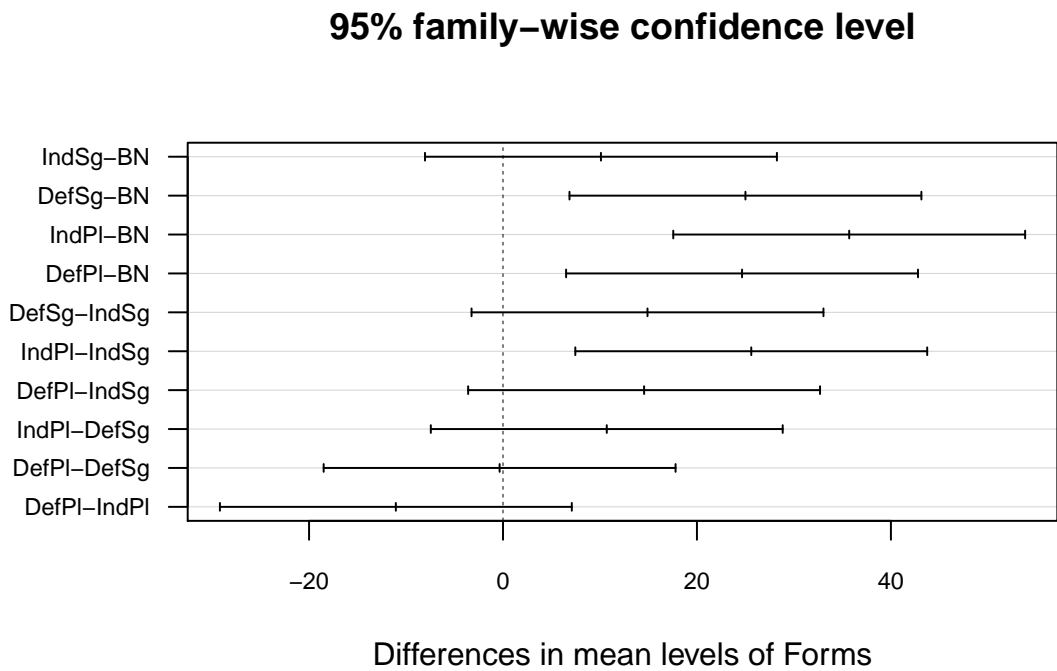
Figure 6.6: Survey 2 – Tukey HSD survey2, 'correct'

The fact that there are statistically significant differences between the tested groups does not tell us in exactly what way they are different – whether some of the NP types are used more often with animate nouns, mass nouns etc.. Statistical analyses take into account the values (here: the number of answers), not actual NPs from the test sentences.

The results provided show nevertheless that statistically significant differences are in many cases similar to those presented in the cognitive analysis. For instance, BNs are significantly different from definite singular and two plural forms. Even though certain generic contexts are similar between those forms (e.g. referring to WEKs is possible both with BN and definite singular), most of them are different such as *classic* true kind reference with indefinite plural or delimited generics of definite plural which is not compatible with BN.

6.3.2.2 Answers 'acceptable'

The answers marked by the respondents as 'acceptable' were analysed in the same manner, namely first descriptive statistics were performed and then a number of more complex tests.

The table 6.12 shows descriptive statistics of the data ranked as 'acceptable'. The material was first divided into quartiles. Then, minimal and maximal values were calculated, together with mean, median and SD. The boxplot fig:survey2-boxplot-acceptable shows the graphic representation of the data.

	BN	IndSg	DefSg	IndPl	DefPl
Min.	4.00	5.00	8.00	4.00	13.00
1st Qu.	8.50	9.00	13.00	12.50	20.75
Median	11.00	11.50	15.00	19.00	24.00
Mean	11.10	13.00	17.20	17.40	23.35
3rd Qu.	13.25	16.50	21.00	24.00	26.25
Max.	19.00	25.00	33.00	25.00	34.00
SD	4.22	5.92	6.17	6.92	5.39

Table 6.12: Survey 2 – data 'acceptable'

As can be observed, the values are much lower than in the case of correct answers. Nevertheless, certain tendencies can be observed. For instance, SD of all NP types is rather homogeneous, meaning that the range of the values are not very spread out in reference to the mean. What is more, the differences between minimal and maximal values are much lower for each of the NP types. One of the reasons for this is certainly a smaller data set. However, it is also possible that the answers 'acceptable' are not as clear-cut as 'correct'.

The next step of the analysis consisted in verifying the assumptions needed for ANOVA. First, homogeneity of variances was tested with the use of Levene's test. It

Survey2 –– acceptable
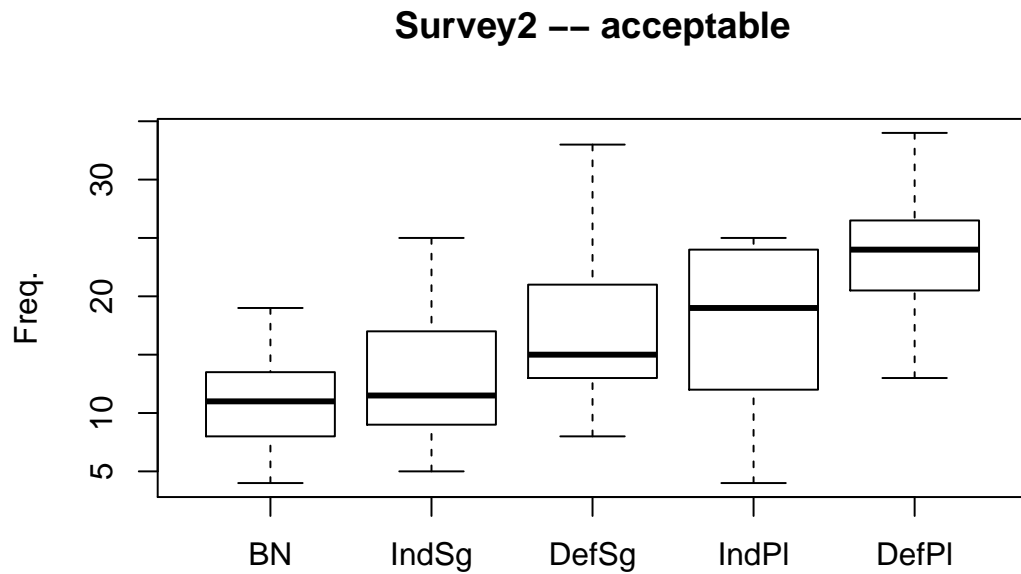
Figure 6.7: Survey 2 – boxplot 'acceptable'

resulted in p-value of 0.28 (see table 6.13), proving there was no violation of the first ANOVA requirement.

	Df	F value	Pr($>$F)
Group	4	1.28	0.28
	95		

Table 6.13: Survey 2 – Levene's test, 'acceptable'

The second requirement needed for a parametric test, namely normality of data, was checked with Shapiro-Wilk test. The result of p-value=0.69 means that the data is normally distributed and the analysis of variance can be performed.

	Df	Sum Sq	Mean Sq	F value	Pr($>$F)	
Forms	4	1792	448	13.35	1.14e-08	***
Residuals	95	3188	33.60			

(***) significant at 0.001 level.

Table 6.14: Survey 2 – ANOVA, 'acceptable'

As can be seen in table 6.14, ANOVA showed that there were differences between the tested groups so a post hoc test needed to be conducted. The post hoc test of choice was Tukey HSD.

The multiple pairwise comparison test has shown that there are significant differences between six pairs of NP types – these are marked with bold in the table 6.15 and

	BN	IndSg	DefSg	IndPl
IndSg	-1.90			
p-adjust	0.84			
DefSg	-6.10	-4.20		
p-adjust	**0.01**	0.16		
IndPl	-6.30	-4.40	-0.20	
p-adjust	**0.008**	0.12	1.00	
DefPl	-12.25	-10.35	-6.15	-5.95
p-adjust	**0.00**	**0.00**	**0.01**	**0.014**

Table 6.15: Survey 2 – Tukey HSD, 'acceptable'

95% family–wise confidence level

Differences in mean levels of Forms
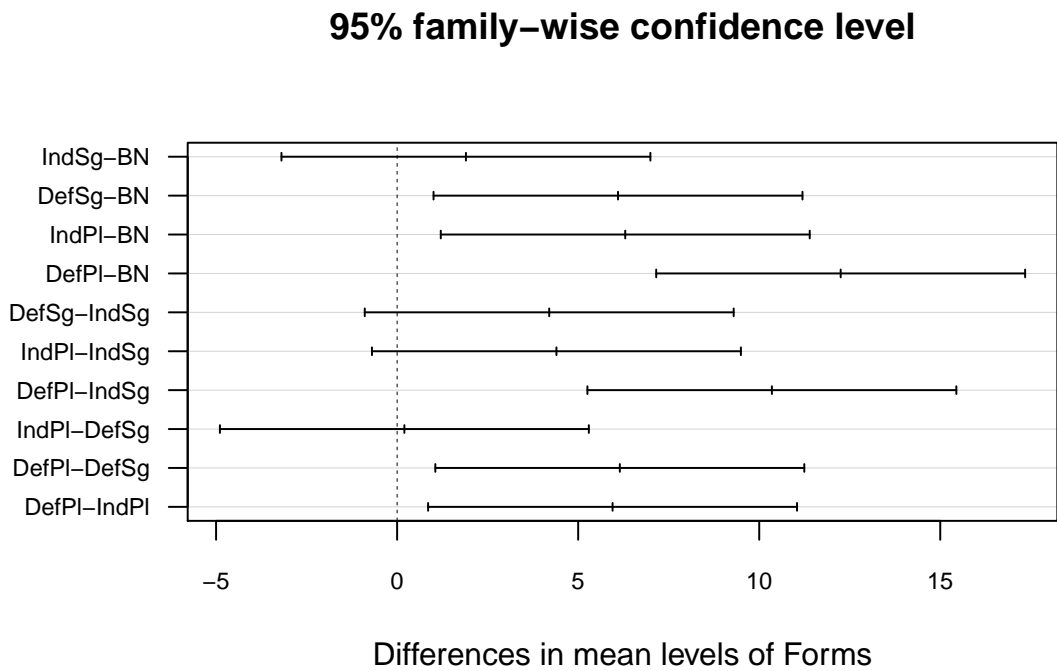
Figure 6.8: Survey 2 – Tukey HSD, 'acceptable'

as values not crossing the 0-line on the graph 6.8.

Also in the case of answers marked as 'acceptable', certain similarities to the cognitive model can be observed. For instance, BNs are significantly different from definite singular and two plural forms which was also confirmed in the previous section on the correct answers. The tendencies of the NP types marked as 'acceptable' are very similar to those marked as 'correct', indicating rather positive reading of this grade in the survey. However, in the case of acceptable answers, definite plural was recognised as different from every other form which can confirm its limited use in ambiguous generic contexts.

6.4 Conclusions

The survey conducted with the use of the corpus material was an Acceptability Judgement Task where participants' intuitions about generic expressions were tested. The survey contained 20 test items and 5 filler items, both in a form of generic sentences with gaps. The participants rated each of the NP types that could be inserted into the gaps. The rating scale had three levels: 1 – 'correct', 2 – 'acceptable' and 3 – 'incorrect/awkward'.

The majority of the test items were rated in a way that mirrored the original material, namely the NP types judged most often as 'correct' and/or 'acceptable' were in accordance with the forms used originally in the corpus. In each of the examples more than one form was rated as 1 or 2, indicating a high degree of interchangeability of the forms. However, certain NP types, such as indefinite singular, were not very common in the answers.

The results collected in this survey allowed to create a matrix of generic NP types, similarly as in the previous chapter. There were fewer available contexts in this study, due to the smaller amount of the material. However, certain generic contexts were confirmed by the respondents such as conceptual generics expressed with BNs, reference to prototypes rendered with the singular forms or describing certain gregarious animal species with definite plural generics.

The last part of this chapter consisted in the statistical analysis of the collected results, where a number of statistical tests allowed to verify the distribution of the data. However, the statistical tests cannot be used to illustrate the research question but rather to present main tendencies when it comes to values themselves. In the following chapter the results of the pilot survey, the corpus study and the AJT survey will be discussed together in order to create a cognitive model of generics in Norwegian.

7 Discussion and conclusions

The project described in this dissertation consisted of three tasks: conducting a pilot survey, building and analysing a corpus of generic texts and conducting a second survey, based on the corpus material. The first task was performed in 2017 during a 3-months stay in Norway, whereas the two other tasks were conducted in the years 2018-2019.

The goal of this project was to investigate and analyse cognitive status of Norwegian generics, according to the models proposed by Leslie et al. (2011); Radden and Dirven (2007) and Radden (2009). This way, a unified model for Norwegian was created based on the corpus and the AJT survey. The model will be discussed in detail in section 7.2.

7.1 General results – discussion

Pilot study The first part of this project consisted in designing and conducting a pilot study in a form of a survey. The survey was inspired by the study of Ionin et al. (2011), namely it contained 30 short generic texts with gaps to fill. The survey was published online and the respondents were supposed to choose the NP types that would fit the gaps the best. The goal was to examine as many generic contexts in Norwegian as possible, based on the speakers' intuitions. The total of 630 people took part in the survey, of which 599 were recognised as native speakers of Norwegian.

The results of the first survey have shown a high percentage of BNs in generic contexts with somewhat limited use of definite plural generics. The second most chosen form was indefinite plural, suggesting the similarity to many Germanic languages when it comes to treating indefinite plurals as *default* generics. Apart from those, all other NP types were used frequently by the participants but in different texts.

In order to further determine which NP types were chosen by the respondents in generic contexts, the texts were divided into groups based on the types of generic predications they contained. The model proposed by Leslie et al. (2011) was used in order to do it.

The texts contained 26 countable nouns and 4 mass nouns. When it comes to texts with countable nouns, all types of generic predications were observed with majority characteristic and false generalisations being the most common ones (8 texts in each of the two categories). The other frequent category was quasi-definitional generics, recognised

in 4 of the survey texts with countable nouns. Majority generics occurred in 3 texts, minority characteristic in 2 and striking generics in 1. When it comes to mass nouns, 2 of the texts contained majority predications, 1 majority characteristic and 1 striking generics.

In the survey texts, all five NP types were chosen by the respondents but with different frequency. BNs, indefinite plurals and indefinite singulars were the most frequent. However, in a number of examples delimited generics were preferred. Apart from the texts about people ('child' and 'politician'), also other subjects occurred as definite plural (e.g. 'truck' and 'ceramic stove'). Definite forms occurs more often in predications classified as quasi-definitional or false generalization. Those two categories have a lot in common when it comes to the interpretation. Quasi-definitional generics state a truth that is applicable to the whole kind and do not accept exceptions, whereas false generalizations state false truths that have clear counterparts. Nevertheless, both predication types have a form of a definition, just with different truth-conditions.

When it comes to BNs that were chosen in many of the survey texts, they occurred with all of the predications. According to the scholarly literature on generics in Germanic languages (e.g. Carlson (1977) for English, Pettersson (1976) for Swedish and Halmøy (2016) for Norwegian), indefinite plurals are considered *default*. Nevertheless, already the pilot study has shown that bare nouns occur in many generic contexts and with different types of noun phrases (animate, inanimate, countable and mass). The two other analyses of this project have proved the same, as can be seen below.

Corpus analysis   The second and most important part of the study was building a tailor-made corpus of generic texts. In total, 170 texts (27 761 tokens) were retrieved from the online enecyclopaedia *Store norske leksikon*. The material was divided into 5 thematic categories, namely 1) 'people' 2) 'animals' 3) 'plants' 4) 'tools' and 5) 'other'. The texts were tagged with regard to genericity. The analyses of all text categories have shown that all five NP types were used in generic contexts, with definite plurals being more frequent than in the pilot survey. However, one of the forms, namely indefinite singular, was rather rare throughout the corpus and even absent in two of the text categories ('animals' and 'plants'). Indefinite plurals were again very common in all of the texts, proving the special status of BPs also in Norwegian. What is more, a distinction between prototypes and actual representations in generic texts and sentences was observed. For instance, generic expressions referring to professions were slightly more abstract than those describing groups of people or objects.

Also in the corpus texts, BNs were very frequent and occurred in all of the text categories. One of the reasons for such a high frequency of BNs are encyclopaedic entries which in many cases were also subject in the first sentences in many of the corpus texts. However, in many cases BNs were utilised also in subsequent mentions of a given topic,

indicating that the form has a generic reading also in other contexts. In section 7.2, possible contexts for BNs and other NP types will be provided, based on the corpus and the AJT survey.

AJT survey The third and last part of the project was the Acceptability Judgement Task survey built with the use of the corpus material. 20 generic sentences were chosen from the corpus as test items, together with 5 accompanying filler items. The judgement scale used in the survey had 3 levels: 1 – 'correct', 2 – 'acceptable' and 3 – 'incorrect/awkward'. In total, 100 native speakers of Norwegian took part in the study, most of them aged 18-30.

The majority of the participants opted for *clear-cut* answers (grades 1 and 3), leaving out the 'acceptable' option in many of the survey questions. However, the results align in a very high degree with model sentences from the corpus, meaning that despite the lack of a broader context, the participants were still able to interpret the sentences as generic and grade them accordingly. The main tendencies connected to the use of certain NP types were similar as in the two previous parts of the project (the pilot survey and the corpus, see section 7.2).

The answers from the AJT survey provided further interpretations of the NP types occurring in generic contexts. For instance, according to the respondents kind reference can be rendered with 4 different NP types (except for definite plural), with slight difference in meaning. As has been discussed in chapter 6, reference to whole kinds can be expressed either through denotation of individual entities or through reference to a group of entities. Singular forms base on prototypes, whereas BN renders a generic reference through a concept. When it come to indefinite plural, it refers to a kind by describing characteristics of a given group.

Definite plural generics, called also delimited generics in cognitive theory on the matter, also allow to make generalizations about kinds but in a slightly different manner. In the corpus and in the survey, the form was used in relation to gregarious animals, namely animals that are usually seen in flocks. When it comes to plants, mass-like plants were most often described with the use of this form, both in the corpus and in the survey.

7.2 Cognitive status of Norwegian generics

The core of the analyses in this project are cognitive models for English and Swedish. As has been presented in chapters 5 and 6, the material collected for the project allowed to rework the model of generics and adjust it to the Norwegian data. First, the data was analysed according to the generic interpretations provided by Radden and Dirven (2007); Radden (2009) and Pettersson (1976). Then, a simplified matrix of generic noun was created – one for the corpus data and one for the AJT survey answers. The second part

of the cognitive analysis consisted in combining the two paradigms of generic NP types and reworking the model proposed by Radden (2009), where different forms and their generic meanings are provided.

The original model can be found in Table 2 in Radden (2009: 224), repeated here as figure 7.1.

	generic type	generic form	ex-/inclusiveness	generic meaning
(a)	representative generic	indefinite singular	exclusive	arbitrary instance representing its type
(b)	proportional generic	indefinite plural	exclusive/ inclusive	salient proportion of the type's reference mass
(c)	kind generic	definite singular	inclusive	prototypical subtype of a well-established type
(d)	delimited generic	definite plural	inclusive	delimited human set within a domain

Figure 7.1: 'Types of generic reference' (Table 2) in Radden (2009: 224).

Both the descriptive and the statistical analyses of all parts of the research, namely the pilot survey, the corpus and the AJT survey, show that the diversity of the forms that appear in generic contexts is by no means random. Based on the corpus material and judgements from native speakers of Norwegian, a simplified matrix of generic NPs was created. One of the categories from Radden's model, namely 'ex-/inclusiveness' was replaced with the contexts in which the generic NPs occurred. The contexts marked in bold appeared both in the corpus and in the judgements in the AJT survey. 'Kinds' marked with an asterisk indicate that the reference was recognised as conceptual, whereas the generic meaning marked with italics represent the element of the original model that did not need to be modified.

Each of the rows in the table 7.1 contains generic types, generic forms and meanings, as well as contexts in which a given NP type was used most often. The proposed classification is based only on the material collected for this project and is by no means exhaustive. A potential expansion of the system could be provided should other aspects of genericity be analysed (e.g. generics in the spoken language, differences between the two written standards, different source material etc.). What is more, certain generic sentences can sometimes have more than one interpretation so the matrix may differ slightly depending on one's language intuition (cf. the interpretation of WEKs and so-called familiar nouns in the language in chapter 5).

As can be observed in the table 7.1, certain elements were modified compared to the original model. First of all, the Norwegian model contains five NP types and not only four, since BNs are widely used in generic contexts in Norwegian. Second of all, generic meanings are slightly different based on the available contexts.

	generic type	generic form	context	generic meaning
(a)	conceptual generic	BN	mass & quasi-mass nouns, **WEKs**, **concepts**, professions, *kinds	concept representing its type
(b)	representative (prototypical) generic	indefinite singular	stand-alone objects, professions, **prototypes**, *kinds	arbitrary instance (prototype) representing its type
(c)	prototypical & kind generic	definite singular	mass nouns, **WEKs**, official functions, ***kinds**, prototypes	*salient proportion of a type's reference mass*
(d)	kind generic	indefinite plural	**kinds, groups of people**	majority representing its type
(e)	delimited generic	definite plural	**ethnic/social groups, gregarious animals,** subkinds, hyperonyms, mass-like plants	delimited set within a given domain

Table 7.1: Generic NPs – a simplified matrix based on the corpus and the AJT survey

The model is based on the empirical material consisting of two parts. However, many of the contexts were the same, both in the corpus and in the AJT survey and those are marked with bold. Certain contexts were lacking in the corpus but were observed in the respondents' judgements. One of such contexts included a conceptual reference to kinds that was judged as correct and/or acceptable in the case of BNs. Prototypical reference to kinds was rendered with the two singular forms – indefinite and definite. In Radden's model, definite singular has exactly the same function, which was marked with italics in the Norwegian model.

The use of BNs in the collected material, apart from its conceptual interpretation, was often connected to mass nouns and those perceived as mass, e.g. 'moss' (grammatically countable in Norwegian). Also well-established kinds and nouns considered familiar in everyday speech appeared as BNs both in the surveys and in the corpus. Concepts such as 'religion' and abstract nouns were also referred to in this way. Interpreting certain professions as concepts is also known in Norwegian and can result in such sentences as the one in (129a), repeated here as (141).

(141)   Psykolog   er en person med utdanning i   psykologi.
        psychologist-Ø is a   person with education in psychology
        A psychologist is a person with education in psychology.

Even though the word 'person' appears in the description of the notion, the interpretation is rather that 'psychologist' is a profession that includes certain tasks, education level and so on – in other words it is a prototypical psychologist that has education in the field of

psychology, not necessarily a particular member of this group.

The use of indefinite singulars in generic contexts was also interesting, especially that this form did not appear in two of the corpus categories, namely 'animals' and 'plants', whereas native speakers judged sentences about animals with indefinite singulars as correct and/or acceptable, although the percentage of such answers was relatively low (see chapter 6).

The indefinite singular form was also used when referring to stand-alone objects and tools, such as 'a knife' or 'a pen', namely objects that are usually perceived as individual tools rather than occurring in sets. This interpretation also brings to mind the notion of a prototype, as in the sentence *A knife is a tool used for cutting*. It is not a certain knife, nor an example of a knife that has the function but rather a prototype of it. This way, by assigning a characteristic feature to an object, a generalisation about the whole kind is made. In the second survey, certain nouns were judged as correct in such contexts, e.g. 'a gene' and 'a ballpoint pen'. Also professions referring to a prototypical person performing a given task were described this way (e.g. in sentences such as *A teacher is someone who works at school*). Reference to kinds rendered with indefinite singular was prototypical (cf. Lyons 1977 and Pettersson 1976).

When it comes to definite singulars, apart from mass nouns and WEKs, also official functions ('the king', 'the County Governor') and certain animal species were described with the use of this form. The animal species that appeared as definite singulars included both *exotic* kinds such as 'the giraffe', as well as species present in the Scandinavian landscape such as 'the eel'. In Radden's model, definite singular is recognised as kind generic, whereas in Norwegian it also serves as a prototypical generic reference.

Indefinite plurals, as the least controversial of all generic NP types, were frequent both in the corpus and in the two surveys. The judgements of native speakers of Norwegian have confirmed the special status of this form, often seen as *default* in many Germanic languages. It is seen as correct almost in any generic context, be it with a kind-predicate or not (see the examples in 142). Studies on generics in other languages have also confirmed this claim (see e.g. the study of Karczewski (2016) where English and Polish generics are compared[1]).

(142)   a. Krokodiller lever i vann og finnes i tropiske områder over hele
            croodile-PL live in water and finds in tropical climate in   whole
            kloden.
            world
            Crocodiles live in water and they can be found in tropical climate in the whole
            world.

        b. Hos haier   som lever av fisk og lignende er det utviklet   skjære-
            with shark-PL that live on fish and similar   is it   developed flattened

---

[1]In Polish, a language without articles, plural form was seen as a corresponding one to English bare plurals.

eller gripetenner.

or needle-like.teeth

Sharks that live on fish and similar food have developed flattened or needle-like teeth.

The sentence in (142a) shows what can be considered a *classic* generic reference denoting a whole kind. The use of the predicate 'to live in water' makes the reference possible only when it concerns the whole kind. It is not possible that only certain crocodiles would live in water as it is a feature characteristic for the whole kind. Indefinite plural was judged correct 93 times – most often among all of the NP types.

The sentence in (142b) on the other hand, shows the use of regular predicate with indefinite plural in a generic sentence. The respondents judged the form as correct 88 times, making it most widely used form in this example.

Both sentences in (142) show that the respondents chose the same form as in the model sentences from the corpus. What is more, the predicates used in the two examples confirm the claims of Pettersson (1976), Carlson (1977), Chierchia (1998) and others that bare plurals can be considered a *default* generic form in English and other Germanic languages.

The last of the mentioned NP types, namely the definite plural, is not very common in generic contexts e.g. in English, but it appears quite often in Norwegian. In the collected material, definite plural generics were most often used in texts and sentences about people (e.g. 'the Sami'), as well as animal species considered gregarious or usually seen in flocks (e.g. 'the bumblebees'). In certain cases, subkinds were accessed through this form (e.g. 'the tipis' – a particular type of a tent), whereas in other definite plurals denoted hyperonyms (e.g. 'the drilling machines' – a general category). The use of definite plural in generic contexts in Norwegian is indeed delimited but still allows for a wider range of contexts compared to English.

The types of generic references and their meanings presented in table 7.1 were chosen among those that appeared most often in the corpus and in the AJT survey. Certain contexts, especially in the corpus, depended on the writing style. Encyclopaedic texts belong to a genre where the information needs to be concise and understandable. What is more, in many texts the first mention of a text's topic was both an encyclopaedic entry and a subject of the first sentence in a text. In such cases, the noun phrase was often a BN. However, the use of BNs was analysed in detail in order to account for other possible contexts and not only those of the first mentions.

Genericity as a part of general knowledge The idea that genericity is a part of general knowledge is not new (see e.g. Leslie 2007 and Karczewski 2016 among others). What is more, some studies have shown that the ability to form and interpret generic expressions

might be innate or it is developed at a very young age.[2] The empirical material presented in this dissertation also suggests that despite the normative language rules, the cognition itself is a key factor in understanding and expressing generics. General knowledge and interpretation of certain phenomena have an influence on the way genericity is rendered in a language. For instance, understanding whether one speaks about a prototypical object or a specific one is not only conditioned by the sentence's structure but also by the context and one's cognitive abilities. The so-called *generic knowledge* (Karczewski 2016: 19) can be observed when the corpus material (chapter 5) is compared with the results of the AJT survey (chapter 6). In many of the cases, the participants intuitively graded the *default* forms (namely the ones that occurred in the corpus) as correct and/or acceptable. Also the recognition of the special status of the noun 'stave church' was observed, indicating that the notion of well-established kinds also might be a part of one's generic knowledge.

7.3   Implications for further research

Some of the aspects of genericity that were out of the scope of this project include the study of local varieties of the Norwegian language and contrastive study of generics in the two written standards of Norwegian. The status of the language is rather complicated: as has been said, there are two written standards of Norwegian and no spoken standard. What is more, there are certain varieties of *bokmål* and *nynorsk* – some are more conservative, other more liberal. Furthermore, the number of local varieties and dialects in Norway is difficult to estimate, given such a big diversity and a lack of standardised language.

The choice of *bokmål* for this project proved effective, however a comparison of the two written standards could provide a broader perspective on genericity in Norwegian. Even though a great majority of the written resources are available in *bokmål*, a set of data in *nynorsk* could be collected, for instance with the use popular science magazines and websites such as forskning.no, where the texts are published in two language variants. Also some articles from *Store norske leksikon* are available in *nynorsk*, even though they are few compared to the ones written in *bokmål*.

When it comes to analysing generics in local varieties of Norwegian, the case is a lot more complex. Dialects are mostly spoken, however it is possible to find texts written in non-standard Norwegian, e.g. on a forum or in the comment section on YouTube. There are certain tools that allow for automatic extraction of such texts (such an application can be constructed e.g. in R or Python). Nevertheless, filtering out the needed information cannot be automatised in such a high degree since the written versions of spoken dialects are very diverse and they differ from the standardised *bokmål* and *nynorsk*. Another option for analysing the dialects of the language is conducting a study similar to the one

---

[2]The study of Leslie and Gelman (2012) proves for instance that children recognise generic statements before they are able to interpret quantificational ones.

of Oosterhof (2008), where a survey was constructed and speakers of certain local varieties of the language were consulted. Such a project could provide an insight into generics in Norwegian dialects but it would certainly require a lot of time and resources.

As has been demonstrated throughout this dissertation, the scholarly literature on genericity concerns mostly English and other widely spoken languages. Testing the existing models of generics in other languages, especially those less studied, would certainly contribute to the development of the models, as well as provide more information on genericity in general.

Bibliography

Allan, Keith (2014). *Lingustic meaning.* London and New York: Routledge.

Allan, Keith (ed.) (2016). *The Routledge Handbook of Linguistics.* London and New York: Routledge.

Becker, Misha (2014). *The Acquisition of Syntactic Structure. Animacy and Thematic Alignment.* Cambridge: Cambridge University Press.

Behrens, Leila (2000). *Typological Parameters of Genericity. (Arbeitspapier 37.).* Köln: Institut für Sprachwissenschaft der Universität zu Köln.

Behrens, Leila (2005). Genericity from a cross-linguistic perspective. *Linguistics*, 43(2), pp. 275–344.

Borthen, Kaja (2003). *Norwegian Bare Singulars.* Ph.D. thesis, Norwegian University of Science and Technology, Trondheim.

Borthen, Kaja (2007). The correspondence between cognitive status and the form of kind-referring NPs. In: Jeanette K. Gundel et al. (eds.), *The Grammar-pragmatics Interface. Essays in Honor of Jeanette K. Gundel*, chap. 7. Amsterdam/Philadephia: John Benjamins Publishing Company, pp. 143–169.

Carlson, Greg (2010). Generics and Concepts. In: Francis Jeffry Pelletier (ed.), *Kinds, Things, and Stuff. Mass Terms and Generics*. New Directions in Cognitive Science, Oxford: Oxford University Press, pp. 16–35.

Carlson, Greg N. (1977). A Unified Analysis of the English Bare Plural. *Linguistics and Philosophy*, 1, pp. 413–457.

Carlson, Greg N. (1979). Generics and Atemporal 'When'. *Linguistics and Philosophy*, 3, pp. 49–98.

Carlson, Greg N. (1982). Generic terms and generic sentences. *Journal of Philosophical Logic*, 11, pp. 145–181.

Carlson, Greg N. (1983). Logical form: types of evidence. *Linguistics and Philosophy*, 6, pp. 295–317.

Carlson, Gregory N. and Pelletier, Francis Jeffry (eds.) (1995). *The Generic Book.* Chicago and London: The University of Chicago Press.

Carlsson, Yvonne (2012). *Genericitet i text.* Ph.D. thesis, Stockholms universitet, Stockholm.

Chierchia, Gunnaro (1998). Reference to kinds across languages. *Natural Language Semantics*, 6, pp. 339–405.

Collins, John (2018). Genericity sans Gen. *Mind and Language*, 33, pp. 34–64.

Comrie, Bernard (1989). *Language Universals and Linguistic Typology.* Chicago: The University of Chicago Press, 2 edn.

Creswell, John W. and Clark, Vicki L. Plano (2018). *Designing and Conducting Mixed Methods Research. [Ebook].* Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE, 3. edn.

Dahl, Östen (1975). On generics. In: E. Keenan (ed.), *Formal semantics of natural languages.* Cambridge: Cambridge University Press, pp. 99–111.

Dahl, Östen (1985). Remarques sur le generique. *Langages*, 79, pp. 55–60.

Dahl, Östen (2015). *Grammaticalization in the North. Noun phrase morphosyntax in Scandinavian vernaculars, Studies in Diversity Linguistics*, vol. 6. Berlin: Language Science Press.

Dahl, Östen and Koptjevskaja-Tamm, Maria (2001). Kinship in Grammar. In: Irene Baron et al. (eds.), *Dimensions of Possession, Typological Studies in Language*, vol. 47. Amsterdam/Philadephia: John Benjamins Publishing Company, pp. 201–226.

Delsing, Lars-Olof (1993). *The Internal Structure of Noun Phrases in the Scandinavian Languages A Comparative Study.* Ph.D. thesis, Lunds universitet, Institutionen för nordiska språk, Lund.

Devitt, Michael (2010). Linguistic Intuitions Revisited. *The British Journal for the Philosophy of Science*, 61(4), pp. 833–865.

Dyvik, Helge (1979). Omkring fremveksten av artiklene i norsk. Språklig markering av referensielle forutsetninger. *Maal og Minne*, 1-2, pp. 40–78.

Faarlund, Jan Terje (1988). *Morfologi. Bøyningssystemet i nynorsk og bokmål.* Oslo: Det Norske Samlaget.

Faarlund, Jan Terje et al. (1997). *Norsk referansegrammatikk.* Oslo: Universitetsforlaget.

Farkas, Donka F. and de Swart, Henriëtte (2009). Stability and variation in article choice: generic and non-generic contexts. In: Anastasia Giannakidou and Monika Rathert (eds.), *Quantification, Definiteness & Nominalization.* Oxford Studies in Theoretical Linguistics, Oxford: Oxford University Press, pp. 188–213.

Frege, Gottlob ([1892] 2010). On sense and reference. In: Darragh Byrne and Max Kolbel (eds.), *Arguing About Language.* London and New York: Routledge, pp. 35–56.

Galmiche, Michel (1985). Phrases, syntagmes et articles génériques. *Langages*, 79, pp. 2–39.

Gelman, Susan A. (2003). *The essential child.* New York: Oxford University Press.

Gelman, Susan A. and Bloom, Paul (2007). Developmental changes in understanding of generics. *Cognition*, 105, pp. 166–183.

Gelman, Susan A. and Tardif, Twila (1998). A cross-linguistic comparison of generic noun phrases in English and Mandarin. *Cognition*, 66, pp. 215–248.

Golden, Anne et al. (2008). *Norsk som fremmedspråk. Grammatikk.* Oslo: Universitetsforlaget, 3. edn.

Gundel, Jeanette et al. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, pp. 274–307.

Gunnerud, Sverre Martin (2017). Grunnskole-tall: 87,7 prosent bokmål, 12,2 prosent nynorsk. URL `https://www.riksmalsforbundet.no/grunnskolen-877-bokmal-122-nynorsk/`.

Hagen, Jon Erik (1998). *Norsk grammatikk for andrespråkslærere.* Oslo: Ad notam Gylendal.

Halmøy, Madeleine (2016). *The Norwegian Nominal System – A Neo-Saussurean Perspective*, *Trends in Linguistics. Studies and Monographs*, vol. 294. Berlin, Boston: De Gruyter Mouton.

Hansen, Erik (1994). Generisk substantiv. In: *Språkbruk, grammatik och språkförändring: en festskrift till Ulf Teleman 13.1.1994.* Lund: Lunds universitet, Institutionen för nordiska språk, pp. 137–144.

Hansen, Erik and Heltoft, Lars (2011). *Grammatik over det Danske Sprog*, vol. 1-3. Odense: Syddansk Universitetsforlag.

Hawkins, John A. (2015). *Definitness and Indefinitness. A Study in Reference and Grammaticality Prediction.* London and New York: Routledge.

Hesse-Biber, Sharlene Nagy (2010). *Mixed Methods Research. Merging Theory with Practice.* New York and London: The Guilford Press.

Ionin, Tania et al. (2011). An experimental investigation of the expression of genericity in English, Spanish and Brazilian Portuguese. *Lingua*, 121, pp. 963–985.

Jacoby, David M. P. et al. (2014). Shark personalities? Repeatability of social network traits in a widely distributed predatory fish. *Behavioral Ecology and Sociobiology*, 68(12), pp. 1995–2003.

Julien, Marit (2005). *Nominal Phrases from a Scandinavian Perspective.* Amsterdam/Philadephia: John Benjamins Publishing Company.

Kabatek, Johannes (ed.) (2013). *New Perspectives on Bare Noun Phrases in Romance and Beyond, Studies in Language Companion Series (SLCS)*, vol. 141. John Benjamins Publishing Company.

Karczewski, Daniel (2016). *Generyczność w języku i w myśleniu. Studium kognitywne.* Kraków: Universitas.

Kreidler, Charles W. (1998). *Introducing English Semantics.* New York: Routledge.

Krifka, Manfred et al. (1995). Genericity: An Introduction. In: Gregory N. Carlson and Francis Jeffry Pelletier (eds.), *The Generic Book*. Chicago and London: The University of Chicago Press, pp. 1–124.

Kulbrandstad, Lars Anders (1998). *Språkets mønstre. Grammatiske begreper og metoder.* Oslo: Universitetsforlaget.

Kurek, Anna (2017). Hvordan skildres generisitet i norsk faglitteratur? *Folia Scandinavica Posnaniensia*, 22, pp. 4–17.

Lazaridou-Chatzigoga, Dimitra et al. (2015). Genericity is easy? Formal and experimental perspectives. *Ratio*, XXVIII, pp. 470–494.

Leslie, Sarah-Jane (2007). Generics and the Structure of the Mind. *Philosophical Perspectives, Philosophy of Mind*, 21, pp. 375–403.

Leslie, Sarah-Jane (2015). Generics Oversimplified. *Noûs*, 49(1), pp. 28–54.

Leslie, Sarah-Jane and Gelman, Susan A. (2012). Quantified statements are recalled as generics: Evidence from preschool children and adults. *Cognitive Psychology*, 64, pp. 186–214.

Leslie, Sarah-Jane et al. (2011). Do all ducks lay eggs? The generic overgeneralization effect. *Journal of Memory and Language*, 65, pp. 15–31.

Liebesman, David (2011). Simple Generics. *Noûs*, 45(3), pp. 409–442.

Lønning, Jan Tore (1987). Mass terms and quantification. *Linguistics and Philosophy*, 10, pp. 1–52.

Lyons, John (1977). *Semantics.* Cambridge: Cambridge University Press.

MacDonald, Kirsti (1997). *Spørsmål om grammatikk. Når norsk er andrespråk.* Oslo: Cappelen.

MacDonald, Kirsti (2009). *Norsk grammatikk. Norsk som andrespråk. Teoribok.* Oslo: Cappelen Damm.

Mari, Alda et al. (eds.) (2013a). *Genericity.* Oxford: Oxford University Press.

Mari, Alda et al. (2013b). Introduction. In: Alda Mari et al. (eds.), *Genericity.* Oxford: Oxford University Press, pp. 1–92.

McEnery, Tony and Hardie, Andrew (2012). *Corpus Linguistics. Method, Theory and Practice.* Cambridge Textbooks in Linguistics, Cambridge: Cambridge University Press.

Medin, Douglas L. and Ortony, Andrew (1989). Psychological essentialism. In: Stella Vosniadou and Andrew Ortony (eds.), *Similarity and analogical reasoning.* New York: Cambridge University Press, pp. 179–195.

Mertens, Willem et al. (2017). *Quantitative Data Analysis. A Companion for Accounting and Information Systems Research.* Switzerland: Springer.

Molnár, Krisztina (2014). *Substantivdetermination im Deutschen und im Ungarischen. Eine sprachtypologisch-kontrastive Analyse.* Berlin, Boston: De Gruyter.

Nickel, Bernhard (2017). Generics. In: Bob Hale et al. (eds.), *A Companion to the Philosophy of Language.* Oxford: Wiley Blackwell, pp. 437–462.

Oosterhof, Albert (2008). *The Semantics of Generics in Dutch and Related Languages.* Amsterdam/Philadephia: John Benjamins Publishing Company.

Pelletier, Francis Jeffry (ed.) (2010). *Kinds, Things, and Stuff. Mass Terms and Generics.* New Directions in Cognitive Science, Oxford: Oxford University Press.

Pettersson, Thore (1976). Bestämda och obestämda former. In: Eva Gårding (ed.), *Kontrastiv fonetik och syntax med svenska i centrum.* Liber Läromedel, pp. 119–142.

Prasada, Sandeep and Dillingham, Elaine M. (2009). Representation of principled connections: A window onto the formal aspect of common sense perception. *Cognition*, 33(3), pp. 401–448.

Quine, Willard Van Orman (2013). *Word and Object. New Edition*. Cambridge: The MIT Press.

Radden, Günter (2009). Generic reference in English. A metonymic and conceptual blending analysis. In: Klaus-Uwe Panther et al. (eds.), *Metonymy and Metaphor in Grammar*, *Cognitive Human Processing*, vol. 25. Amsterdam/Philadephia: John Benjamins Publishing Company, pp. 199–228.

Radden, Günter and Dirven, René (2007). *Cognitive English Grammar*, *Cognitive Linguistics in Practice*, vol. 2. Amsterdam/Philadephia: John Benjamins Publishing Company.

Rasinger, Sebastian M. (2013). *Quantitative Reserach in Linguistics. An Introduction*. London - New Delhi - New York - Sydney: Bloomsbury, 2 edn.

Riazi, A. Mehdi (2016). *The Routledge Encyclopedia of Research Methods in Applied Linguistics. Quantitative, qualitative, and mixed-methods research*. London and New York: Routledge.

Riemer, Nick (2016). Internalist semantics: meaning, conceptualization and expression. In: Nick Riemer (ed.), *The Routledge Handbook of Semantics*. London and New York: Routledge, pp. 30–47.

Rosén, Victoria and Borthen, Kaja (2017). Norwegian bare singulars revisited. In: Victoria Rosén and Koenraad De Smedt (eds.), *The very model of a modern linguist*, *BeLLS*, vol. 8. University of Bergen, pp. 220–240.

Rosenbach, Anette (2008). Animacy and grammatical variation – Findings from English genitive variation. *Lingua*, 118, pp. 151–171.

Saeed, John I. (2009). *Semantics*. Malden, Mass: Wiley Blackwell.

Schütze, Carson T. (2016). *The empirical base of linguistics. Grammaticality judgements and linguistics methodology*, *Classics in Linguistics*, vol. 2. Berlin: Language Science Press.

Skrzypek, Dominika and Kurek, Anna (2018). Generics in Mainland Scandinavian languages. *European Journal of Scandinavian Studies*, 48(2), pp. 253–266.

Sterken, Rachel Katharine (2016). Generics, Covert Structure and Logical Form. *Mind and Language*, 31, pp. 503–529.

Strandskogen, Åse Berit and Strandskogen, Rolf (1995). *Norwegian. An Essential Grammar*. London and New York: Routledge.

Teleman, Ulf et al. (1999). *Svenska Akademiens grammatik*, vol. 1-4. Stockholm: Norstedts i distrobution.

Yamamoto, Matsumi (1999). *Animacy and Reference. A cognitive aproach to corpus linguistics*. Amsterdam/Philadephia: John Benjamins Publishing Company.

Summary in Polish. Streszczenie w języku polskim

Generyczność w języku norweskim. Analiza kognitywna

Tematem rozprawy jest generyczność w języku norweskim w ujęciu kognitywnym. Zjawisko to występuje w każdym zbadanym dotąd języku, jednak w żadnym z nich nie zidentyfikowano dotąd narzędzia, za pomocą którego można jednoznacznie wyrazić generyczność (Behrens 2005). W zależności od rodziny języków, odniesienia generyczne są wyrażane na różne sposoby. W językach germańskich, do których należy język norweski, paleta możliwości jest szczególnie szeroka. W wyrażeniach generycznych stosowane mogą być zarówno formy określone grup imiennych jak i nieokreślone, a w niektórych przypadkach również rzeczowniki nagie.

Celem niniejszego badania jest zweryfikowanie w jaki sposób generyczność może być wyrażana w języku norweskim, a także stworzenie modelu opartego na teoriach kognitywnych, który usystematyzowałby użycie poszczególnych form generycznych fraz imiennych. Głównym założeniem pracy jest wypełnienie luki w badaniach nad generycznością w mniej zbadanych językach, a także przeprowadzenie badania empirycznego języka norweskiego.

Pytania badawcze niniejszej rozprawy to:

- Jakie formy fraz imiennych występują w kontekstach generycznych w języku norweskim?
- Czy i w jaki sposób kontekst wpływa na interpretację generyczności?
- Czy i w jaki sposób zdolności kognitywne natywnych użytkowników języka wpływają na wyrażenie odniesień generycznych?

Przeprowadzenie niniejszego badania oraz odpowiedzi na powyższe pytania pozwoliły na zbudowanie kognitywnego modelu generyczności, który obrazuje występowanie zjawiska w języku norweskim. Podstawę teoretyczną pracy stanowią teorie kognitywne oraz modele wykorzystywane do analizy generyczności (m.in. Radden 2009; Pettersson 1976 i Leslie et al. (2011)).

Język norweski jest dość nietypowy, ponieważ nie tylko nie ma żadnego standardu

wymowy, ale także posiada dwa warianty języka pisanego – *bokmål* i *nynorsk*. *Bokmål* jest o wiele bardziej powszechny, zarówno w literaturze, mediach, życiu codziennym jak i edukacji (Gunnerud 2017). Co więcej, dostępność tekstów napisanych w tym wariancie języka norweskiego jest większa, co również miało wpływ na wybór wariantu *bokmål* do niniejszego badania.

Jak wspomniano powyżej, generyczność jest zjawiskiem, które obecne jest we wszystkich zbadanych dotąd językach, jednak nie stwierdzono istnienia elementu, który stosowany byłby jedynie do wyrażenia tego typu odniesienia. Jest to rodzaj odniesienia do całego gatunku lub całej grupy osób/przedmiotów, o których jest mowa w danym zdaniu czy tekście (Lyons 1977).

Według literatury fachowej (zob. np. Carlson and Pelletier 1995; Mari et al. 2013a i Carlson 2010), generyczność może być rozumiana dwojako, a mianowicie jako *klasyczne* odniesienie gatunkowe (*kind reference*), odnoszące się do całych gatunków oraz jako zdanie habitualne (*habitual sentence*), opisujące powtarzające się czynności. Co więcej, generyczność może być wyrażona w języku na trzech poziomach: 1) grupy imiennej, 2) zdania oraz 3) tekstu (Behrens 2005).

Generyczność na poziomie grupy imiennej zakłada, że sama grupa imienna jest interpretowana jako generyczna, bez względu na użyte w zdaniu predykaty. Najczęściej jest to określenie gatunkowe, a więc odnoszące się do wszystkich lub większości przedstawicieli danego gatunku, jak np. podmioty w zdaniach *Psy szczekają* czy *Włosi lubią makaron*. W językach germańskich, a więc także i w języku norweskim, funkcję tę pełni najczęściej forma nieokreślona liczby mnogiej, uznawana za *domyślną* formę generyczną grupy imiennej.

Drugim poziomem, na którym może wystąpić odniesienie generyczne jest poziom zdania. Wówczas to predykat zastosowany w danym zdaniu decyduje o jego interpretacji. Najczęściej są to predykaty gatunkowe lub takie, które nie występują w odniesieniach specyficznych. Na przykład predykat 'być ssakiem' może odnosić się jedynie do całego gatunku. Nie możemy powiedzieć, że tylko niektóre psy lub koty są ssakami, podczas gdy inne nie. W ten sposób każde zdanie, w którym znajdzie się tzw. predykat gatunkowy (*kind predicate*) jest generyczne. Zdania generyczne mogą również zawierać generyczne grupy imienne.

Trzecim poziomem generyczności wyszczególnionym przez Behrens są tzw. teksty generyczne. Teksty te mogą zawierać zarówno generyczne grupy imienne jak i zdania, jednak ich ogólny wydźwięk uzależniony jest od kontekstu. Teksty generyczne to najczęściej teksty naukowe i popularnonaukowe opisujące gatunki, przedmioty lub zjawiska naturalne. Tego typu teksty opisują cechy charakterystyczne dla danego gatunku, przedmiotu czy zjawiska, a więc są generyczne. Jednakże nie wszystkie zdania czy grupy imienne obecne w takich tekstach zostaną zinterpretowane jako generyczne bez szerszego kontekstu. W niniejszej dysertacji zostaną zbadane wszystkie trzy poziomy generyczności

– badanie pilotażowe dotyczy głównie poziomu 1 i 2, badanie korpusowe skupia się na poziomie 3, a ankieta AJT na poziomie 2.

Rozprawa jest podzielona na siedem rozdziałów. Pierwszy rozdział stanowi ogólne wprowadzenie do badań oraz zarysowanie celów i pytań badawczych projektu. Rozdziały 2-3 stanowią tło teoretyczne dla części empirycznej. Rozdziały 4-6 zawierają opisy przeprowadzonych badań: ankiety pilotażowej, badania korpusowego oraz ankiety Acceptability Judgement Task (AJT). Ostatni rozdział z kolei zawiera szczegółowe omówienie przedstawionych wcześniej wyników, konkluzje oraz perspektywy dalszych badań nad generycznością.

W rozdziale pierwszym przedstawione zostały cele i pytania badawcze, a także wybrana metodologia. Metodologia badań zastosowana w projekcie to tzw. *mixed methods research*, w skrócie MMR. Polega ona na połączeniu metod kwalitatywnych (w tym przypadku modeli kognitywnych) z metodami kwantytawnymi (w tym przypadku są to testy statystyczne).

Rozdział drugi zawiera przegląd najważniejszych publikacji dotyczących generyczności. Omówione prace dotyczą głównie języka angielskiego oraz innych, dobrze zbadanych języków. Zdecydowana większość teorii dotyczących generyczności opiera się na przykładach z języka angielskiego. Zarówno język angielski jak i norweski należą do rodziny języków germańskich, co umożliwiło zastosowanie omówionych w rozdziale modeli do przeprowadzenia niniejszych badań.

W rozdziale trzecim omówiono dostępną literaturę dotyczącą genryczności w językach skandynawskich – duńskim, szwedzkim i norweskim. Jest to ogólny zarys problematyki, oparty głównie na opracowaniach dostępnych w gramatykach wspomnianych języków. Jedynie dla języka szwedzkiego istnieje publikacja dotycząca empirycznych badań nad generycznością przeprowadzonych przez Carlsson (2012). Nieliczne opracowania dotyczące języka norweskiego stanowią jedynie część większych badań dotyczących kategorii określoności lub rzeczowników (m.in. Borthen 2003 i Halmøy 2016), a nie samej generyczności.

Rozdział czwarty opisuje przeprowadzone badanie pilotażowe, składające się z ankiety. Forma ankiety, a mianowicie 30 krótkich tekstów generycznych z lukami do wypełnienia, została zaczerpnięta z badania Ionin et al. (2011). Teksty zostały napisane w odmianie *bokmål* i przeczytane przez natywnego użytkownika języka norweskiego. Ankieta, wraz z pytaniami kontrolnymi mającymi na celu rozpoznanie natywnych użykowników języka, została stworzona za pomocą Google Forms i opublikowana w internecie. Podczas dwóch tygodni, gdy ankieta była dostępna online, 630 respondentów udzieliło odpowiedzi na pytania. 599 z nich zostało zidentyfikowanych jako natywni użytkownicy języka norweskiego.

Badanie pilotażowe wykazało, że spośród pięciu form rzeczownika w języku norweskim (forma naga, forma określona i nieokreślona liczby pojedynczej oraz forma określona i nieokreślona liczby mnogiej), wszystkie mogą mieć znaczenie generyczne, jednak

nie w każdym możliwym kontekście. Co więcej, formy mogą do pewnego stopnia być stosowane zamiennie, jednak i w tym przypadku nie może być mowy o zupełnej dowolności.

Wyniki badania pilotażowego zostały przeanalizowane na podstawie modelu zaproponowanego przez Leslie et al. (2011), który zakłada występowanie sześciu różnych typów predykatów generycznych: 1) quasi-definicji, 2) cechy większościowej, 3) cechy mniejszościowej, 4) predykatu większości, 5) predykatu *szokującego* oraz 6) fałszywego uogólnienia (Leslie et al. 2011: 19).

W przeprowadzonym badaniu pilotażowym zaobserwowanego występowanie każdego z predykatów generycznych, jednak nie we wszystkich przypadkach respondenci wybierali te same formy grupy imiennej. Również testy statystczne potwierdziły występowanie znaczących różnych pomiędzy testowanymi grupami, tj. pięcioma formami grupy imiennej. Oznacza to, że kontekst generyczny dyktuje wybór danej formy.

Rozdział piąty jest najobszerniejszym rozdziałem empirycznym i dotyczy przeprowadzonego badania korpusowego. Korpus do analizy został zbudowany za pomocą programu R i oparty na tekstach pobranych ze *Store norske leksikon*, norweskiej encyklopedii online. Korpus liczył w sumie 170 tekstów (27 761 tokenów), podzielonych na pięć kategorii tematycznych: 1) ludzie, 2) zwierzęta, 3) rośliny, 4) narzędzia, 5) inne. Kategoria 'ludzie' zawierała 20 tekstów, kategorie 'zwierzęta', 'rośliny' oraz 'narzędzia' po 25 tekstów, natomiast kategoria 'inne' składała się z 75 tekstów. Każdy z tekstów miał długość przynajmniej jednego paragrafu i był napisany w wariancie języka *bokmål*.

Dane pozyskane w badaniu korpusowym również zostały przeanalizowane na dwa sposoby, zgodnie z metodą MMR. Pierwszy etap analizy zakładał zastosowanie modeli kognitywnych dla języka angielskiego (Radden 2009) i szwedzkiego Pettersson (1976). Na podstawie tychże modeli pogrupowano dane według typów grup imiennych w każdym z tekstów, a także według kontekstów generycznych, w których występowały. Analiza ta pozwoliła na stwierdzenie, że forma naga przyjmuje często interpretację konceptualną, podczas gdy dwie formy liczby pojedynczej wyrażają generyczność za pomocą odniesień do prototypów. Z kolei forma nieokreślona liczby mnogiej, podobnie jak w języku angielskim, niemal zawsze odnosi się do całych gatunków. Tak zwana ograniczona generyczność (ang. *delimited generics*) ma w norweskim szersze zastosowanie niż w języku angielskim i może być stosowana w większej ilości kontekstów, nie tylko tych dotyczących opisów ludzi. Zebrane dane pozwoliły na stworzenie matrycy form generycznych dla języka norweskiego.

Drugim elementem analizy korpusowej było przeprowadzenie testów statystycznych. Testy te wykazały główne tendencje w występowaniu poszczególnych form grup imiennych, jak i różnice pomiędzy nimi. Główne tendencje zbadane statystycznie pokryły się w dużej mierze z analizą opisową opartą na modelach kognitywnych.

W rozdziale szóstym opisana została ostatnia część materiału empirycznego, a mianowicie ankieta AJT. Ankieta zawierała 20 zdań generycznych oraz 5 zdań niezawierają-

cych odniesienia generycznego. Wszystkie zdania zostały zaczerpnięte z korpusu. Ankieta została zbudowana za pomocą aplikacji Shiny oraz programu R, a następnie umieszczona na serwerze. Badanie zostało przeprowadzone na Uniwersytecie w Tromsø oraz w bibliotece miejskiej w Tromsø, gdzie zrekrutowano w sumie 100 osób.

Respondenci mieli za zadanie dokonanie oceny każdej z form rzeczownika w wybranych do badania zdaniach i z wykorzystaniem skali. Skala ocen poszczególnych form obejmowała 3 stopnie: 1 – forma poprawna, 2 – forma akceptowalna i 3 – forma niepoprawna/ nienaturalna. Zbrane w ten sposób odpowiedzi zostały przeanalizowane w opraciu o modele kognitywne, podobnie jak w rozdziale piątym, oraz za pomocą testów statystycznych.

Badanie wykazało, że w większości przypadków formy grup imiennych w oryginalnych zdaniach korpusowych pokrywały się z odpowiedziami respondentów oznaczonymi jako poprawne (ocena 1). Co więcej, ocena 2 dostarczyła dodatkowych informacji na temat kontekstów generycznych, w których może występować więcej niż jedna forma grupy imiennej. Z kolei ocena 3 pozwoliła na wykluczenie form grup imiennych w danych kontekstach generycznych.

Testy statystyczne zostały przeprowadzone dla odpowiedzi oznaczonych jako 1 i 2, a więc jako formy poprawne i akceptowalne. W ten sposób potwierdzone zostały główne tendencje, widoczne również dzięki analizie kognitywnej. Również dla danych pozyskanych w ankiecie AJT utworzono matrycę form grup imiennych w kontekstach generycznych, która uzupełnia matrycę z rozdziału piątego.

Rozdział siódmy zawiera podsumowanie przeprowadzonych badań oraz syntezę wszystkich danych. W rozdziale uwzględniono wszystkie trzy elementy projektu, a na podstawie dwóch z nich, tj. analizy korpusowej oraz ankiety AJT, zaproponowano model kognitywny form generycznych dla języka norweskiego. Model ten pokrywa się w pewnym stopniu z modelami Raddena i Petterssona, ale także ukazuje konteksty generyczne nieobecne w tychże modelach. Ostatnia część rozdziału siódmego dotyczy perspektywy dalszych badań nad generycznością z wykorzystaniem teorii kognitywnych.

Appendices

A   Pilot study – survey 1 texts

1. Det å trene og utvikle muskulaturen har i det siste blitt veldig populært verden rundt, også i Norge. I dag er det ikke bare menn som ønsker å bli sterkere og se strammere ut. De fleste vet at ... kan gi gode resultater både hos kvinner og menn.
   (a) styrketrening
   (b) en styrketrening
   (c) styrketreningen
   (d) styrketreninger
   (e) styrketreningene

2. At et sunt kosthold er viktig når man trener, vet alle. Det som skaper uenighet blant personlige trenere og helseentusiaster er hva et sunt kosthold egentlig er. Noen påstår for eksempel at ... bør spises daglig for å forebygge mangel på viktige vitaminer og næringsstoffer.
   (a) eple
   (b) et eple
   (c) eplet
   (d) epler
   (e) eplene

3. Blant dagens mange dietter og kosthold finner man forskjellige tilnærminger til næring. Alle har sine egne meninger om hva som bør spises og hva som bør unngås. Det er for eksempel enighet om at:
   (a) appelsin er en god kilde til c-vitaminer;
   (b) en appelsin er en god kilde til c-vitaminer;
   (c) appelsinen er en god kilde til c-vitaminer;
   (d) appelsiner er en god kilde til c-vitaminer;
   (e) appelsinene er en god kilde til c-vitaminer.

4. Har du lyst til å dekorere huset ditt, er det viktig med en detaljert plan og en god oversikt over budsjettet du har til rådighet. Det kan være vanskelig å bestemme om hvordan leiligheten eller huset skal se ut etter oppussing. Ikke la deg påvirke av design-blader eller selgere i butikker som håper å tjene på deg, for ... gir nok de beste rådene.
   (a) venn

(b) en venn

(c) vennen

(d) venner

(e) vennene

5. Verdenspolitikken har i det siste vært i krise. I mange land, blant annet Norge, minsker samfunnets tillit til politikere og offentlige organisasjoner. Flere og flere mennesker innrømmer at de ikke tror på det de store politiske partiene påstår. Bortsett fra det mener mange at ... har relativt stor makt i det moderne samfunnet.

(a) politiker

(b) en politiker

(c) politikeren

(d) politikere

(e) politikerne

6. Antall bilder og videoer med katter som man finner på nettet er overraskende stort. Katteelskere betrakter sine firbeinte venner som konger. Likevel synes fortsatt mange at det faktisk er ... som er det beste selskap.

(a) hund

(b) en hund

(c) hunden

(d) hunder

(e) hundene

7. Stadig flere sliter med depresjon og angsttilstander som et resultat av stress. For mange kan plikter på jobb og i hverdagslivet være en hovedgrunn til økt stressnivå. Blant vanlige behandlinger og terapier, finner man også en del psykologer som anbefaler alternative behandlingsmetoder. For noen mennesker kan for eksempel ... være til hjelp.

(a) religion

(b) en religion

(c) religionen

(d) religioner

(e) religionene

8. Hver dag bruker man opp til fire timer på internett. De fleste bruker nettet både på datamaskin, nettbrett og telefon. Dekningen her i landet er såpass bra at det ikke er noe problem å surfe på nettet utenfor hjemmet eller arbeidsplassen. Ifølge de nyeste statistikkene er det nettopp ... de fleste bruker til å sjekke e-post og sosiale medier.

(a) smarttelefon

(b) en smarttelefon

(c) smarttelefonen

    (d) smarttelefoner

    (e) smarttelefonene

9. Antall bilulykker i Europa er på et relativt høyt nivå. Tusenvis av lastebiler som kjører gjennom kontinentet bidrar til stor trafikk på motorveier og i byer. For å kjøre trygt og unngå ulykker er det viktig å huske at ... veier flere tonn og trenger derfor lang tid til å bremse.

    (a) lastebil

    (b) en lastebil

    (c) lastebilen

    (d) lastebiler

    (e) lastebilene

10. Hvert år bygges det nye sykkelveier i mange norske byer og tettsteder. De som bytter sine biler mot sykler påstår at det ikke bare er bra for helsen å sykle, men at det også kan fikse budsjettet og hjelpe å spare litt ekstra hver måned. Det er jo klart at ... er billige og økovennlige transportmidler.

    (a) sykkel

    (b) en sykkel

    (c) sykkelen

    (d) sykler

    (e) syklene

11. Bør man spise meieriprodukter for å få i seg nok kalsium? Og kan kalsium fra meieriprodukter styrke bein i kroppen? De nyeste undersøkelsene viser at i land der det spises mest meieriprodukter forekommer benskjørhet mye oftere enn vanlig. Er det faktisk slik at ... forårsaker dette?

    (a) melk

    (b) en melk

    (c) melken

12. Autistiske barn har det vanskelig med å tilegne seg grunnleggende kunnskaper som lesing og skriving. Gjennom årene har man jobbet med nye opplæringsteknikker og verktøy som kan brukes på skolen. Det har blitt bl.a. påvist at ... hjelper barn med autisme til å fokusere bedre på skolen.

    (a) musikk

    (b) en musikk

    (c) musikken

13. I dagens Norge er tilpasset opplæring og inkluderende skoler de viktigste målene man arbeider mot. Lærere og foreldre kommer stadig med nye idéer om hva som er best for de yngste. Hva om elever selv kunne bestemme hva de vil lære? Det er viktig å huske at ... også utgjør en viktig del av samfunnet

    (a) barn [sg.]

(b) et barn

(c) barnet

(d) barn [pl.]

(e) barna

14. Det er blitt lettere å være entreprenør i Norge. Blant de nye prosjektene som støttes er det flere små bryggerier rundt omkring i landet. Nybegynnere i bransjen ser det som en sjanse til å komme inn på ølmarkedet. For disse entreprenørene betyr ... de selger en liten, men fast inntekt.

(a) øl [sg.]

(b) et øl

(c) ølet

(d) øl [pl.]

(e) ølene

15. Det er typisk norsk å dra på hyttetur. Uansett om du reiser alene eller med venner må du ikke glemme det som er viktigst på slike turer, nemlig maten! Taco og varme retter kan være fristende, men det er også greit med noe mer tradisjonelt. Derfor bør man alltid ha ... på hytta si.

(a) brødmat

(b) en brødmat

(c) brødmaten

16. I en digitalisert verden der nesten alt foregår på nettet, står papirlitteraturen overfor den største krisen noensinne. Det er flere som velger å laste ned en app istedenfor å lese en bok i fritiden. Mange mener også at vanlige papirbøker er i ferd med å forsvinne helt fra markedet. I stedet kan imidlertid ... være en god litterær opplevelse

(a) lydbok

(b) en lydbok

(c) lydboken

(d) lydbøker

(e) lydbøkene

17. Hvordan bør man kle på seg på en jobbtur? Om du jobber i et stort firma der det er en kleskode, kan det være en krevende sak. Det er ikke uvanlig å ta på seg en dress eller en kjole på en slik tur. Problemet kan være tilbehør man må ha med seg. Hvordan reiser man med dokumenter og en pc når man også skal være elegant? ... er ikke lenger forbudt i den profesjonelle verden.

(a) Ryggsekk

(b) En ryggsekk

(c) Ryggsekken

(d) Ryggsekker

(e) Ryggsekkene

18. I høstmånedene er det mange som velger å kose seg hjemme om kvelden med et glass vin og en god bok. Lurer du på hvordan du kan skape den koselige høststemningen i ditt eget hjem? Det er relativt enkelt og krever ikke veldig mye tid. Først og fremst er det fint med ... hjemme.

   (a) stearinlys[sg.]

   (b) et stearinlys

   (c) stearinlyset

   (d) stearinlys [pl.]

   (e) stearinlysene

19. Når vinteren er godt i gang, er det greit å handle vinterklær på salg. Flere butikker tilbyr da salg opp til 50% fra de vanlige prisene. Lurer du på hva som er best å handle om vinteren? De fleste velger ... av god kvalitet

   (a) dunjakke

   (b) en dunjakke

   (c) dunjakken

   (d) dunjakker

   (e) dunjakkene

20. Sommerferie er noe de fleste ser fram til hele året. Når sommeren endelig kommer, er det bare kos og gøy man er opptatt av. Da er det enkelt å glemme riktig beskyttelse mot solen. Solkrem har jo de fleste, men man burde også beskytte hodet. Da kan det være både praktisk og stilig med ...

   (a) stråhatt

   (b) en stråhatt

   (c) stråhatten

   (d) stråhatter

   (e) stråhattene

21. Hvordan velger man den rette leiligheten? Det kan være fristende å gå på alle mulige visninger i området, men ekspertene mener at man heller burde velge to-tre leiligheter man liker godt. Til syvende og sist bestemmes gjerne valget av ... man får.

   (a) god magefølelse

   (b) en god magefølelse

   (c) den gode magefølelsen

   (d) gøde mageføleler

   (e) de gode magefølelsene

22. Det er mange faktorer som må tas hensyn til når man skal pusse opp et kontor. Gode og komfortable møbler må man selvfølgelig ha, men tilbehøret er også viktig. Flere undersøkelser har for eksempel vist at ... hjelper til å fokusere bedre.

    (a) grønn lampe

    (b) en grønn lampe

    (c) den grønne lampen

    (d) grønne lamper

    (e) de grønne lampene

23. Det å planlegge et bryllup kan ta mye tid og penger. Særlig for kvinner er det vanskelig siden de må velge en spesiell kjole til seremonien. Bryllupsmoten endrer seg like raskt som den hverdagslige moten. I dag er det vel ikke bare ... som velges til bryllup.

    (a) hvit kjole

    (b) en hvit kjole

    (c) den hvite kjolen

    (d) hvite kjoler

    (e) de hvite kjolene

24. I dagens verden er det viktig å være online døgnet rundt. Dette gjør at telefonprodusentene kommer med stadig nye funksjoner i produktene sine. Ifølge brukerne er det i dag bare ... som teller.

    (a) stor trykkskjerm

    (b) en stor trykkskjerm

    (c) den store trykkskjermen

    (d) store trykkskjermer

    (e) de store trykkskjermene

25. Interiørdesign endrer seg stadig og kan noen ganger bli litt overraskende. Moderne designere bruker ofte gamle kopper og keramikk til å pynte kjøkkenet og stuen. Man kan for eksempel bruke ... til å pynte kjøkkenet

    (a) vanlig ølflaske

    (b) en vanlig ølflaske

    (c) den vanlige ølflasken

    (d) vanlige ølflasker

    (e) de vanlige ølflaskene

26. Et jobbintervju er kanskje det viktigste steget når man leter etter en jobb. Det første inntrykket arbeidsgiveren får av deg er viktig og kan være avgjørende. Pass derfor på at du kler deg riktig. Har du ingen idé om hva du kan ta på deg, kan ... passe best.

    (a) hvit skjorte

    (b) en hvit skjorte

    (c) den hvite skjorten

    (d) hvite skjorter

    (e) de hvite skjortene

27. Kjøkkenutstyr er noe alle må velge en gang i livet. Kommer du til å møblere ditt eget kjøkken snart, er det lurt å sammenligne forskjellige typer komfyr før du bestemmer deg. Blant alle de tilgjengelige komfyrene er det ... som anbefales oftest
    (a) keramisk komfyr
    (b) en keramisk komfyr
    (c) den keramiske komfyren
    (d) keramiske komfyrer
    (e) de keramiske komfyrene

28. Det blir stadig vanligere å endre litt på utseendet når man blir eldre. I dag tilbys plastiske operasjoner og ulike kurer som skal utsette aldringsprosessen. Mange velger friske farger når de begynner å få
    (a) grått hår;
    (b) et grått hår;
    (c) det grå håret.

29. Norsk arkitektur og interiørdesign har gått gjennom flere endringer i de siste årene. Det som var typisk norsk for tjue år siden ikke nødvendigvis er det i dag. Men noen ting blir alltid knyttet til den norske kulturen som for eksempel ... i skogen
    (a) rødt hus
    (b) et rødt hus
    (c) det røde huset
    (d) røde hus
    (e) de røde husene

30. Ansiktsuttrykk og symboler kan være forskjellige i de enkelte landene. Men noen tegn er universelle verden rundt. For eksempel vises overgivelse i den militære verden alltid med ...
    (a) hvitt flagg
    (b) et hvitt flagg
    (c) det hvite flagget
    (d) hvite flagg
    (e) de hvite flaggene

B Corpus nouns

ID	Norwegian	English
CATEGORY: ANIMALS		
1	måke	seagull
2	krabbe	crab
3	manet	jellyfish
4	maur	ant
5	øyenstikker	dragonfly
6	humle	bumblebee
7	mygg	mosquito
8	gresshoppe	grasshopper
9	ål	eel
10	skate	batoidea/ray
11	krokodille	crocodile
12	hai	shark
13	frosk	frog
14	høne	hen
15	dronte	dodo bird
16	flaggermus	bat
17	oter	otter
18	vaskebjørn	raccoon
19	hvalross	walrus
20	lama	llama
21	hund	dog
22	esel	donkey
23	sjiraff	giraffe
24	afrikamus	nesomyidae
25	neshorn	rhino

ID	Norwegian	English
CATEGORY: PLANTS		
26	solsikke	sunflower
27	nellik	carnation
28	påskelilje	wild daffodil
29	stemorsblom	viola tricolor (Johnny Jump up)
30	nøkleblom	primula
31	valmue	papaver
32	åkertistel	Cirsium arvense
33	løvetann	common dandelion
34	palme	palm
35	skjermbladtre	umbrella pine
36	mammuttre	giant sequoia
37	edelgran	silver fir
38	alm	wych elm/ Scotch elm
39	sørgepil	weepin willow
40	apebrødtre	baobab
41	lind	small-leaved lime
42	bøk	common beech
43	bregne	fern
44	torvmose	peat moss
45	rødalge	red algae
46	fluesopp	amanita
47	valnøtt	walnut
48	eik	oak
49	hengebjørk	silver birch
50	lerk	larches
CATEGORY: TOOLS		
51	tang	pliers
52	kanyle	cannula
53	intravenøs kanyle	intravenous cannulation
54	sentralt vanekateter	central venous catheter
55	loddebolt	soldering iron
56	lysrør	fluorescent lamp
57	sveising	welding
58	rulle	mangle
59	kniv	knife
60	høvel	hand plane

ID	Norwegian	English
61	kloakk	sewage system
62	fødselstang	obstetrical forceps
63	spiker	nail
64	våpen	weapon
65	skrue	screw
66	skruetrekker	screwdriver
67	skive	washer
68	mutter	nut
69	støvsuger	vacuum cleaner
70	glasskjærer	glass cutter
71	skrunøkkel	spanner
72	spøytepistol	spray gun
73	kipperøks	axe
74	vaterpass	spirit level
75	sag	saw

CATEGORY: OTHER

ID	Norwegian	English
76	aksel	axis
77	hodeplagg	headgear
78	gitar	guitar
79	saksofon	saxophone
80	fyllepenn	fountain pen
81	satelitt	satelite
82	fjernkontroll	remote
83	kalkulator	calculator
84	jekk	jack
85	boremaskin	drill
86	bildekk	car tyre
87	forbrenningsmotor	internal combustion engine
88	stridsvogn	tank
89	boreskip	drillship
90	tanskip	tankship
91	ferge	ferry
92	gaffeltruck	forklift
93	dal	valley
94	barkan	barchan
95	dyne	dune
96	permafrost	permafrost
97	savanne	savannah

ID	Norwegian	English
98	ørken	desert
99	fjord	fjord
100	meander	meander
101	akvarellmaling	watercolor
102	spiseredskap	cutlery
103	gummi	rubber
104	håv	hand net
105	fiskekrok	hook
106	ukulele	ukulele
107	fjernstyring	remote controll
108	lærer	teacher
109	dirigent	bandmaster
110	psykolog	psychologist
111	taksidermist	taxidermist
112	bakkepersonell	aircrew
113	baker	baker
114	butikkslakter	butcher
115	isolatør	insulation installer
116	sølvsmed	silversmith
117	maler	painter
118	maskinist	train driver
119	gevær	gun
120	bazooka	bazooka
121	skriver	printer
122	fritidshus	vacation house
123	stavkirke	stave church
124	badstue	sauna
125	akvedukt	aqeuduct
126	basilika	basilica
127	bu	stall
128	hytte	cottage
129	korskirke	cross-church
130	mausoleum	mausoleum
131	tipi	tipi
132	kulepenn	ballpoint pen
133	blekk	ink
134	akevitt	akvavit
135	fennikel	fennel

ID	Norwegian	English
136	mynt	coin
137	krone	crown
138	taco	taco
139	smalahove	smalahove (traditional Norwegian dish)
140	gen	gene
141	vitamin	vitamin
142	språkvitenskap	linguistics
143	mølle	mill
144	telefon	telephone
145	magnet	magnet
146	papir	paper
147	cellulose	cellulose
148	fremmedord	foreign word
149	demokrati	democracy
150	grunnskole	primary school

CATEGORY: PEOPLE

ID	Norwegian	English
151	kuvlunge	Kuvlung
152	øyskjegge	Øyeskjegg
153	bagler	Bagler
154	birkebeiner	the Birkebein Party
155	konge	king
156	rabbi	rabbi
157	diakon	deacon
158	prest	priest
159	fylkesmann	county governor
160	ombud	public advocate
161	tsar	tsar
162	siktet	accused
163	vitne	witness
164	berserk	berserk
165	adel	man of nobility
166	bonde	farmer
167	biskop	bishop
168	husmann	crofter
169	embetsmann	official
170	sysselmann	sysselmann/ governor

C   Survey 2 sentences

1. Måkene er hvite, grå og svarte (svartbrune, rosenmåken svakt rosa) med lange vinger og til dels svevende flukt.
2. Skater er god matfisk.
3. Dannelsen av meandere er blitt forklart ved at elvevannet kommer i stående svingninger på tvers av strømretningen.
4. I Norge har kniver fremstilt ved småindustri på Toten vært av anerkjent kvalitet, likeså kniver fra Geilo.
5. Fra steinalderen kjennes fiskekroker i stort antall av flint, skifer, men særlig av ben i høyst ulike former.
6. Krokodiller lever i vann og finnes i tropiske områder over hele kloden.
7. Hos haier som lever av fisk og lignende er det utviklet skjære- eller gripetenner.
8. Stavkirke er en høyt utviklet kirketype, oppført i reisverk av tre, kjent fra norsk kristen middelalder.
9. Når kulepennen føres over papiret, ruller kulen mot en blekkpasta (tørrblekk) fra en beholder i skaftet.
10. Gener er oppskrifter for egenskaper hos levende organismer.
11. Birkebeinere var en politisk gruppering i Norge i høymiddelalderen.
12. Diakonen er leder av menighetens diakonitjeneste.
13. Flaggermusene har et enormt næringsbehov.
14. Oteren er utbredt over store deler av Europa, sør for tundragrensen, samt i Nord-Afrika.
15. I naturen finnes sjiraffen i dag bare i Afrika.
16. Torvmosene har en eiendommelig bygning.
17. Humlene danner samfunn og betegnes sammen med honningbiene som sosiale bier.
18. Man skjelner mellom skruer beregnet for stål og metalldeler, og for arbeid i tre.
19. Gitaren kom opprinnelig fra Midtøsten.
20. Ålen finnes i hav og ferskvann fra Nord-Afrika og middelhavslandene i sør til Kvitsjøen i nord.

FILLER ITEMS:

1. De fleste orkestre og kor har en fast sjefdirigent.
2. For husmennene som ikke bodde langs strandkanten, ble jorda en viktig del av næringsveien.
3. Moten med avrundede knivspisser bredte seg fra det toneangivende franske hoff og har holdt seg siden.
4. Sysselmannen på Svalbard er også politimester, notarius publicus og hjelpedommer ved underretten, og har dessuten andre offentlige funksjoner.
5. Det var imidlertid engelsk-amerikaneren Alexander Graham Bell som 1876 konstruerte en brukbar telefon, og som helt til 2002 ble regnet som telefonens oppfinner.