

# Norwegian bare singulars revisited

Victoria Rosén and Kaja Borthen

**Abstract.** Borthen (2003) analyzed bare singulars in Norwegian. In this paper some of the claims made there are reexamined by searching in NorGramBank. The study provides new empirical support to the claims put forth in the earlier work, but it also reveals problems with the prior analysis.

## 1 Introduction

A bare singular is a countable, singular and indefinite nominal constituent that does not have a phonetically realized determiner. Whereas some languages do not have indefinite articles at all, other languages do while still allowing for bare singulars in some cases. These languages include Danish (Asudeh and Mikkelsen 2000), Swedish (Delsing 1993), English (Stvan 1998), Dutch, German and French (De Swart and Zwarts 2009), Hungarian (Kiefer 1994), Albanian (Kallulli 1999), and Brazilian Portuguese (Schmitt and Munn 1999), to mention a few.

In spite of crosslinguistic similarities, the distribution pattern of bare singulars in languages that have indefinite articles varies. For instance, Norwegian allows for bare singulars more extensively than English does. Examples of bare singulars in Norwegian are shown in (1)–(4). Bare singular NPs are marked with boldface in examples here and throughout the article.

- (1) *Per er lærer.*  
Per is teacher  
'Per is a teacher.'
- (2) *Han kjører bil.*  
he drives car  
'He drives a car.'
- (3) *Hest er et koselig dyr.*  
horse is a nice animal  
'The horse is a nice animal.'

- (4) *Taxi er dyrt.*  
 taxi is expensive  
 ‘Taking a taxi is expensive.’

Norwegian bare singulars must often be translated into English with an indefinite article, as in (1) and (2). Sometimes however, a definite article is more appropriate, as in the generic statement in (3). In still other cases, neither type of determiner will suffice, and the translation must be rendered periphrastically, as in (4).

English has a quite restrictive use of bare singulars (see e.g. Stvan 1998). According to De Swart and Zwarts (2009), there are five constructions that license bare singulars in English, named ‘bare location’, ‘bare coordination’, ‘bare predication’, ‘bare reduplication’ and ‘bare incorporation’. Their examples (ibid. p. 280) are given in (5)–(9).

- (5) John is in **hospital**. (Bare location)  
 (6) the way to use **knife** and **fork** (Bare coordination)  
 (7) Mary is **chair** of the department. (Bare predication)  
 (8) He found **door** after **door** closed. (Bare reduplication)  
 (9) She is playing **piano** for the choir. (Bare incorporation)

The meanings expressed in (5)–(9) are possible with bare singulars in Norwegian, and more generally, bare singulars in Norwegian and English share important features. Still, the use of bare singulars in each language shows idiosyncratic patterns. As De Swart and Zwarts (2009, p. 7) put it, bare singulars operate “[...] at the border [...] of syntax and lexicon, of rules and lists, of regularities and idioms”. This poses a particularly strong need for thorough empirical investigations, both in order to accurately describe the distribution pattern of bare singulars in one particular language, and in order to detect crosslinguistic similarities and differences.

In her PhD thesis, Borthen (2003) provided an analysis of Norwegian bare singulars, attempting to account for the necessary and sufficient conditions for the use of these phrases. At the time the research was begun, in 1999, it was not straightforward to search for bare singulars in electronic corpora. In the first place, such corpora for Norwegian were not annotated for the distinction between mass and count nouns. In the second place, it was complicated to pick out NPs without determiners from existing corpora. Several searches for consecutive words such as [V + N<sub>sg, indef</sub>], [V + Adj + N<sub>sg, indef</sub>], and [P + N<sub>sg, indef</sub>] were used. However, such search expressions cannot reliably identify bare singulars, and there is also the risk of finding them only in certain syntactic positions. It was furthermore practically impossible to search for bare singulars in specific constructions. Since it was difficult to reliably find bare singulars

through corpus searches, the data used for the thesis comprised approximately 400 manually collected examples from a number of texts plus a large number of examples that were invented by the author.

The availability of treebanks, which are syntactically annotated corpora, marks a radical change in the study of language. NorGramBank (Dyvik et al. 2016) is a treebank of modern Norwegian constructed by automatically parsing a corpus with NorGram, a computational LFG grammar for Norwegian (Dyvik 2000). A small part of the corpus (approx. 315,000 words) was manually disambiguated using computer-generated discriminants, while the rest (approx. 60 million words) was stochastically disambiguated. NorGramBank was developed in the INESS<sup>1</sup> treebanking infrastructure project (Rosén et al. 2012), which also developed the search language INESS Search (Meurer 2012). The detailed syntactic annotation in NorGramBank and the sophisticated search language make it possible to conduct very fine-grained searches for exactly the phenomena the researcher is interested in.

The main goal of this study is to test some aspects of the theoretical analysis of Borthen (2003) on new, authentic data. We claim that the data made available through NorGramBank and the search options provided by INESS Search constitute an excellent basis for improving the theoretical analysis of the phenomenon under investigation.

## 2 Bare singulars in Norwegian

Borthen (2003) makes a number of observations regarding the syntactic properties of Norwegian bare singulars, some of which are listed below (*ibid.* p. 68).

### Syntactic properties of Norwegian bare singulars:

- They can occur in all basic syntactic positions available for nominal phrases in Norwegian, but not “freely”.
- They can be modified and coordinated.
- They are usually not affected by syntactic alternations such as nominalization, passivization, topicalization, raising, question formation, and subject-object alternations for arguments of presentational verbs.
- Adverbs can freely intervene between Norwegian bare singulars and their co-occurring verbal predicates.

That Norwegian bare singulars cannot occur “freely” in nominal positions means, for instance, that they often occur as direct objects, but not of just any verb and not in just any context. Similarly, they sometimes occur as subjects, but only rarely. This means that the generation of bare singulars is not as productive as that of singulars with overt determiners. On the other hand, we cannot account for Norwegian bare

---

<sup>1</sup> <http://clarino.uib.no/iness>

singulars merely by assuming that they are part of fixed multiword expressions. Such an explanation is unlikely since bare singulars can be modified and coordinated, adverbs can freely intervene between bare singulars and their selecting predicates, and their acceptability is usually not affected by syntactic alternations. As we will see later in this paper, additional evidence for the view that the generation of Norwegian bare singulars is productive comes from the high number of unique combinations of verbs and bare singulars.

Four construction types were posited in Borthen (2003) in order to predict the productive use of Norwegian bare singulars. Each construction type is illustrated with some examples below (ibid. p. 117, 165, 171, 194, 212, 215).

The ‘conventional situation type’ construction

- (10) *Hun er elev.*  
 she is pupil  
 ‘She is a pupil.’
- (11) *Hun går på skole.*  
 she goes to school  
 ‘She goes to school.’

The ‘profiled *have*-predicate’ construction

- (12) *Hun hadde rød ytterfrakk.*  
 she had red coat  
 ‘She had a red coat.’
- (13) *Han mangler sovepose og regnfrakk.*  
 he lacks sleeping bag and rain coat  
 ‘He lacks a sleeping bag and a rain coat.’
- (14) *Vi trenger nytt telt.*  
 we need new tent  
 ‘We need a new tent.’
- (15) *et bord med hvit duk*  
 a table with white cloth  
 ‘a table with a white cloth’
- (16) *Hva skulle vi gjort uten do?*  
 what should we done without toilet  
 ‘What should we have done without a toilet?’

The ‘taxonomic’ construction

- (17) *Det hjelpemiddelet som er mest brukt er datamaskin.*  
 the tool that is most used is computer  
 ‘The type of tool that is used the most is the computer.’
- (18) *Buss er et naturvennlig kjøretøy.*  
 bus is a nature friendly vehicle  
 ‘A bus is a non-polluting vehicle.’

The ‘covert infinitival clause’ construction

- (19) *Sykkel er kult.*  
 bike is cool  
 ‘To ride a bike is cool.’
- (20) *Jeg vil anbefale telt.*  
 I will recommend tent  
 ‘I would recommend (having/using) a tent.’

According to Borthen (2003, p. 153–154), the ‘conventional situation type’ construction licenses bare singulars as long as the bare singular and its selecting predicate denote a conventional situation type. A conventional situation type is a property, state, or activity type that occurs frequently or standardly in a given contextual frame and has particular relevance in this frame as a recurring situation type (Borthen 2003, p. 160). This predicts that bare singulars such as *er elev* ‘is a pupil’ in (10) and *går på skole* ‘goes to school’ in (11) will be acceptable as long as the verb phrases they are part of are intended to describe a conventional situation type. This construction is more general than the ones that license bare singulars in English, and it subsumes the constructions called ‘bare location’ (5), ‘bare predication’ (7), and ‘bare incorporation’ (9) in English. The constructions that license bare singulars may lead to the development of multiword expressions with bare singulars over time, and the ‘conventional situation type’ construction is particularly likely to do so because of phonological and semantic characteristics of the construction (Borthen 2003, p. 153–154).

The ‘profiled *have*-predicate’ construction licenses bare singular arguments on certain interpretations of *have*-predicates, according to Borthen (2003). A *have*-predicate is a predicate that expresses a *have*-relation (an asymmetrical coexistence relation) directly or that can be decomposed into a structure that includes one. For instance, to lack something means to not have something, and to need something means to have a desire or urge to have something. Thus, the verbs *mangle* ‘order’ and *treng* ‘need’ in (13) and (14) are *have*-predicates. Similarly, the preposition *med* ‘with’ in (15) can

be seen as denoting a *have*-relation directly whereas the preposition *uten* ‘without’ in (16) denotes a negated *have*-relation. Bare singulars are licensed as arguments of *have*-predicates as long as the context is such that the focus is on the state in which the denotation of the bare singular simply coexists with some other entity mentioned in the sentence (Borthen 2003, p. 187–188). Due to this construction, verbs such as *ha* ‘have’, *ønske seg* ‘wish for’, *mangle* ‘lack’, *få tak i* ‘get hold of’, *ta med* ‘bring’, *hente (seg)* ‘fetch (for oneself)’, and *ta med (seg)* ‘bring (for oneself)’ allow for bare singular objects on certain interpretations in Norwegian.

The third bare singular licensing construction was originally called the ‘comparison of types’ construction in Borthen (2003). It has been renamed here as the ‘taxonomic’ construction, due to the fact that the denotation of the bare singular is presented as having a specific position in a taxonomic hierarchy. Illustrated in (17) and (18), this construction always involves the copular verb *være* ‘be’ and one preverbal and one postverbal nominal phrase where the bare singular is presented as a hyponym of the denotation of the other nominal phrase. This construction often licenses bare singular subjects.

Finally, according to Borthen (2003), Norwegian has a construction which allows for a “covert infinitival clause interpretation” of indefinite noun phrases in subject or object position. This construction is, however, not a construction that licenses bare singulars directly; it licenses them only if the underlying predication (in the covert infinitival clause) is one which would naturally be expressed by a phrase containing a bare singular object (*ibid.* p. 222). Thus, bare singulars that occur in this kind of example, as in (19) and (20), could be considered to be licensed by the ‘conventional situation type’ construction or the ‘profiled *have*-predicate’ construction.

As for why bare singulars are licensed by the four constructions listed above, Borthen (2003) argues that this has to do with their semantics. Some semantic characteristics of Norwegian bare singulars are listed below (*ibid.* p. 50–51).

Semantic properties of Norwegian bare singulars:

- They can never take wide scope.
- They can never be referential.
- They can never be partitive.
- They can be generic, but not with a (quasi-)universal interpretation.
- They are poorer antecedent candidates of token pronouns than corresponding expressions with indefinite articles, but they can be antecedents of some identity-of-sense anaphors.
- Their descriptive content cannot be too general.

In order to account for these properties and the construction types that license bare singulars, Borthen (2003) assumes that bare singulars are *type emphasizing*. All countable nouns have a dual aspect to them; on the one hand they denote a property, a

type of thing, and on the other hand they may be used to refer to tokens in the world. Whereas indefinites with the indefinite article indicate relative emphasis on the token involved in the given situation, bare singulars emphasize the type of thing introduced and are only licensed in specific constructions that go naturally along with such interpretations. The syntactic constructions that allow for bare singulars are thus motivated but not fully predicted by the semantics of bare singulars.

As mentioned earlier, bare singulars in languages that have the indefinite article share many properties. Still, there are crosslinguistic differences. Borthen (2003) explains this by proposing that type emphasis is a scalar notion. That is, bare singulars in various languages may point to different positions on a scale of type/token emphasis. Bare singulars across languages are similar because they are all type emphasizing (compared to corresponding phrases with the indefinite article). As such, they are restricted semantically as well as destined to appear in constructions that go particularly well together with type emphasis. At the same time, bare singulars are different across languages since they may differ with respect to where on the scale of type emphasis they are positioned. This, in turn, affects the set of constructions that license them (Borthen 2003, p. 226–227).

### 3 Problems with Borthen (2003)

Many of the claims put forth in Borthen (2003) are based on invented examples, introspection and impressions. For instance, the following statement is made: “*Ønske seg* ‘want’ and *dele ut* ‘hand out’ belong to a semantically related group of verbs that co-occur particularly easily, and thus relatively frequently, with bare singulars in Norwegian” (Borthen 2003, p. 164). The group of verbs referred to in this quote are the ones labeled *have*-predicates above. The claim that bare singulars licensed by these verbs are particularly frequent would be more convincing if it were supported by authentic examples and, for the frequency claim, some statistics.

Another claim in Borthen (2003) is that bare singulars “[...] tend to be unacceptable if they have too little descriptive content”; the invented examples in (21) and (22) are meant to illustrate this (*ibid.* p. 50).

(21) *Det ligger kniv på bordet.*  
 it lies knife on the table  
 ‘There is a knife on the table.’

(22) *??Det ligger ting/dings/greie på bordet.*  
 it lies thing/gizmo/thingamajig on the table  
 ‘There is a thing/gizmo/thingamajig on the table.’

Whereas the example in (21) is perfectly fine, the examples in (22) are intuitively unnatural. This claim was based on the intuitions of the author, and would be more convincing with empirical evidence to back it up.

The original study also claimed that “[...] the extensive use of bare singulars in idioms and as part of multi-word lexical entries is striking. In fact, they seem to be more frequent than bare singulars licensed by the general constructions proposed in this thesis” (ibid. p. 342). The claim was also made that “Norwegian bare singulars are usually not affected by syntactic alternations like nominalization, passivization, topicalization, [...]” (ibid. p. 68), from which it follows that bare singulars should show up in these sentence structures also in authentic language use. Again, the question must be asked what evidence there is for these claims.

## 4 Searching for evidence in NorGramBank

We conducted searches in NorGramBank for evidence that can answer research questions such as the following, posed in Borthen (2003).

1. Is it true that bare singular nouns with only very general descriptive content such as *ting* ‘thing’, *dings* ‘gizmo’ and *greie* ‘thingamajig’ are particularly unlikely?
2. What verbs tend to take bare singular arguments? Is it true that *have*-predicates are particularly frequent?
3. What are the most common verb–noun combinations? Is it true that there are more instances of idiomatic expressions with bare singulars than productive uses?
4. Can bare singulars occur in all kinds of non-canonical sentence structures, for example in topicalizations, left-dislocations and clefts?

In 4.2–4.5 these questions will be examined based on searches in NorGramBank.

### 4.1 Bare singulars in NorGramBank

In order to find bare singulars in NorGramBank, we must know what characterizes them. In addition, we must know how bare singulars are represented in NorGramBank to know what features of the treebank annotation to search for.

A bare singular noun phrase is headed by a count noun in the singular form and does not have an article or a determiner. In English and many other languages, this is a sufficient characterization. In Norwegian, however, we must add that the phrase must be indefinite. The reason for this is that Norwegian nouns are inflected for definiteness; in (23) the noun *lærer* ‘teacher’ is inflected for definiteness by adding the singular definite suffix *-en* to the stem.

- (23) *Per er læreren.*  
 Per is the teacher  
 ‘Per is the teacher.’

The lack of a determiner is therefore not a sufficient criterion for the phrase being indefinite; the noun must also be in the indefinite form.

The syntactic annotation in NorGramBank is in the Lexical-Functional Grammar (LFG) formalism (Bresnan 2001; Dalrymple 2001). Each sentence has a constituent structure (c-structure) and a functional structure (f-structure). The c-structure is a context-free phrase structure tree showing the relations of dominance and linear precedence. The f-structure is an attribute–value matrix which provides information about syntactic functions, such as subject and object, and grammatical features, such as number, gender and tense. The properties that characterize bare singulars are represented in the f-structure. We can examine some f-structures to see how this is done. Figure 1 shows the f-structure for the noun *lærer* ‘teacher’.

<b>PRED</b>	<b>'lærer'</b>								
<b>NTYPE</b>	<table style="border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;"> <b>NSEM</b> </td> <td style="padding: 2px 5px;"> <b>7</b> </td> <td style="padding: 2px 5px;"> <b>COMMON</b> </td> <td style="padding: 2px 5px;"> <b>count</b> </td> </tr> <tr> <td style="border-right: 1px solid black; padding: 2px 5px;"> <b>NSYN</b> </td> <td colspan="3" style="padding: 2px 5px;"> <b>common</b> </td> </tr> </table>	<b>NSEM</b>	<b>7</b>	<b>COMMON</b>	<b>count</b>	<b>NSYN</b>	<b>common</b>		
<b>NSEM</b>	<b>7</b>	<b>COMMON</b>	<b>count</b>						
<b>NSYN</b>	<b>common</b>								
<b>GEND</b>	<b>4</b>	<b>MASC +, FEM -, NEUT -</b>							
<b>PERS</b>	<b>3</b>	<b>DEF-MORPH -, NUM sg</b>							

Figure 1: F-structure for the noun *lærer*

The f-structure consists of unordered pairs of attributes and values. Some attributes have simple values; an example is the attribute NUM (number) which has the value *sg*. Some attributes, such as NTYPE, have other f-structures as their values. The value of NTYPE is a new f-structure (labeled with the index ‘6’) which has an attribute NSEM, which in turn has an f-structure (labeled with the index ‘7’) as its value. The innermost f-structure has the attribute COMMON with the value *count*. In such cases we speak of a *path* of attributes that leads to a value; here the path is NTYPE NSEM COMMON. The attribute PRED (for *predicate*) has a special type of value called a *semantic form*. This is usually the citation form of the word in single quotes, sometimes followed by a list of arguments, such as for the subcategorized arguments of verbs. The PRED value in the f-structure in Figure 1 is ‘*lærer*’.

For a noun phrase to be a bare singular, its f-structure must have the values *sg* and *count*. In addition, the noun phrase must not have a determiner and it must not be definite. In Figure 1 there is an attribute DEF-MORPH with the value – (minus). This means that the noun is in the indefinite form, but not necessarily that the noun phrase is indefinite, since an indefinite noun may occur together with a definite determiner in a definite noun phrase. We therefore need to know that the noun phrase is not definite and that the noun phrase does not have a determiner.

In Figure 2 the f-structure for the noun phrase *denne læreren* ‘this teacher’ is shown. This noun phrase has an attribute DEF with the value +; this is the value that must not

be present in order for the noun phrase to be indefinite. The determiner *denne* ‘this’ is represented by the f-structure with the attribute SPEC and its value; the attribute SPEC does not occur in the f-structure of a bare singular.

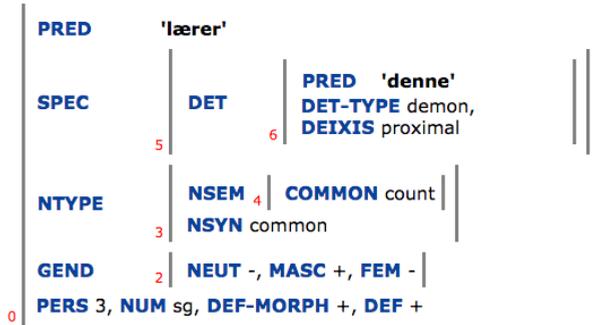


Figure 2: F-structure for the noun phrase *denne læreren*

Searching in NorGramBank is done with INESS Search. The search expression in (24) is designed to pick out bare singular nouns.

- (24) #x\_>PRED #p &  
 #x\_>NUM 'sg' &  
 #x\_>(NTYPE NSEM COMMON) 'count' &  
 !(#x\_>SPEC) &  
 !(#x\_>DEF '\+') &  
 !(#x\_>PRED 'pro') &  
 !(#x\_>(OBL PSEM) 'part')

This expression searches for properties in the f-structure. It consists of seven conjuncts, each of which constrains the search to certain properties which the analysis either must have or must not have. The first conjunct says that there is an f-structure #x\_ that has an attribute PRED with the value #p.<sup>2</sup> The second conjunct states that #x\_ has an attribute NUM with the value 'sg' (atomic f-structure values must be enclosed in single quotes in INESS Search expressions). The third conjunct says that #x\_ has a path of attributes NTYPE NSEM COMMON with the value 'count'. The exclamation point in the final four conjuncts is the negation operator; these conjuncts state which properties the f-structure must not have. It must not have a SPEC, it must not have a definite marking, and it must not have the value 'pro' for its PRED. The final conjunct states that the f-structure must not have a path of attributes OBL PSEM with the value 'part' for *partitive*; this ensures that the lexical item that is the value of #x\_ is not a

<sup>2</sup> All node variables in INESS Search expressions are marked with either #, in which case the node variable is taken to be existentially quantified, or with %, in which case it is universally quantified.

quantifier in a partitive phrase. Together these properties target bare singulars as they are represented in NorGramBank.

INESS Search can present search results as frequency tables, making it easy to examine the results. The search expression indicates which elements are to be displayed in the table by the use of an underscore. If a variable has an underscore, its values are not shown in the table; if a variable does not have an underscore, its values are shown together with their frequencies.

Figure 3 shows the c- and f-structures for (25), one of the sentences found by the search expression in (24). INESS Search highlights the parts of the structure that were searched for in the results. The f-structure of the NP headed by the noun *natt* ‘night’ is marked by a red box and labeled in the top left corner by the variable `#x_`. The values required by the search expression are also marked by red boxes.

- (25) *Det var natt.*  
 it was night  
 ‘It was nighttime.’

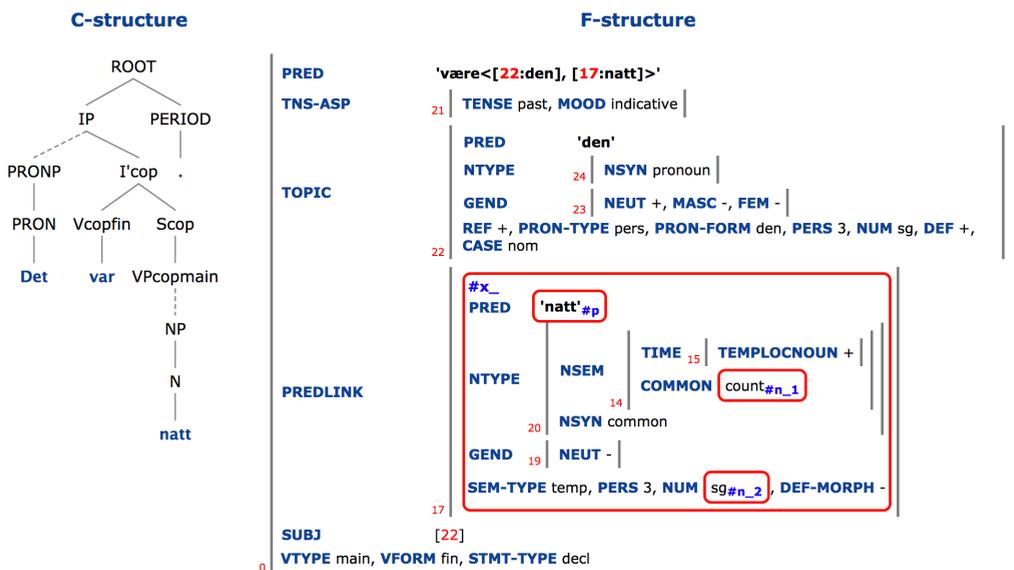


Figure 3: C- and f-structures for the sentence *Det var natt.*

## 4.2 Can nouns with little descriptive content occur as bare singulars?

In order to investigate whether it is true that very general nouns tend not to appear as bare singulars, we searched in the entire automatically disambiguated treebank for Norwegian Bokmål, since it is important to search through large amounts of data to detect a potentially rare phenomenon. The only restriction was that we did not search

among fragment analyses, since these will often assign incorrect bare singular analyses to nouns. The search expression in (24) was amended by adding the desired PRED form to the first conjunct: #x\_ >PRED #p:'dings'. This search resulted in eight hits. One of these was in a sentence that had received an incorrect analysis of the phrase *en såpass diger dings* 'such a large thing', which is not a bare singular. One occurrence was in a headline, where noun phrases are often abbreviated with bare singulars that are not acceptable in a normal context. Three occurrences involve the use of *dings* as a euphemism for the male sex organ and two cases involved compounds that end with the element *dings*; these do not involve the very general descriptive content that is normal with this word. Finally, the example in (26) shows a kind of play on words, where *dings* is contrasted with the nonsense word *dangs*.

- (26) *Testen og eksamen ble laget av en mann, sensor kjenner*  
 the test and exam were made by a man, examiner knows  
*igjen kvinneskrift, NTH er dings, kjemi er dangs osv.*  
 again female writing NTH is gizmo, chemistry is dangs etc.  
 'The test and exam were made by a man, the examiner recognizes the  
 handwriting style of a woman, NTH is gizmo, chemistry is 'dangs', etc.'

The context of this sentence makes it clear that what is being listed here are rationalizations for why there is a lack of gender equality in education. The last two clauses contrast *dings* with *dangs* in a sarcastic manner; the idea is to say that NTH (the Norwegian Institute of Technology) and chemistry are just *this and that* in a pejorative sense. This is not a normal use of the word *dings*. In conclusion, none of the occurrences of *dings* found by this search constitute legitimate examples of bare singulars.

We performed similar searches for the words *ting* 'thing' and *greie* 'thingamajig'. These words get many more hits than *dings* – 1804 for *ting* and 437 for *greie* – simply because they are more frequent words. It has not been feasible to examine all of the hits as we did above, but browsing through them we were not able to find any genuine occurrences of bare singulars. This result must be said to support the claim that bare singulars tend to be unacceptable if their descriptive content is only very general (Borthen 2003, p. 50).

### 4.3 Verbs that take bare singulars as objects

A second claim in Borthen (2003) is that *have*-predicates are particularly frequent with bare singular objects. In order to find out which verbs typically take bare singular objects, the search expression in (24) must be augmented to include specifications about the verb, as shown in (27). The first conjunct says that there is an f-structure #v\_ (the f-structure of the verb) that has an attribute PRED with the value #p, and the second conjunct states that this verb must have the bare singular as its object (OBJ). The third conjunct specifies that the f-structure must have an attribute VFORM; this ensures that the bare singular is the object of a verb, and not of a preposition, for example. The

other conjuncts are the same as in (24) except that *#x\_ >PRED #p* is omitted, since the purpose here is simply to list the most frequent verbs that co-occur with bare singulars, and not the nouns themselves. In order to get the most accurate results possible, we ran the search only on the manually disambiguated part of the corpus, excluding fragment analyses.

(27) *#v\_ >PRED #p &*  
*#v\_ >OBJ #x\_ &*  
*#v\_ >VFORM &*  
*#x\_ >NUM 'sg' &*  
*#x\_ >(NTYPE NSEM COMMON) 'count' &*  
*!(#x\_ >SPEC) &*  
*!(#x\_ >DEF '\+') &*  
*!(#x\_ >PRED 'pro') &*  
*!(#x\_ >(OBL PSEM) 'part')*

The most frequent verbs according to this search are shown in Table 1, with the *have*-predicates in boldface. Note that it is the PRED value of the verb that is listed in the table. Most words have their citation form as their PRED value, but some words have special values; the verb *være* 'to be' has 'exist' as its PRED value in presentational constructions, while the PRED values *få#øye\*på* and *legge#merke\*til* are the predicate names of verbal idioms (see also 4.4 below).

Count	#p: value
178	<b>ha</b> 'have'
86	<b>få</b> 'get'
35	<b>gi</b> 'give'
34	<b>ta</b> 'take'
29	exist 'be-existential'
26	<i>få#øye*på</i> 'catch sight of'
17	<i>legge#merke*til</i> 'notice'
12	<b>bruke</b> 'use'
12	si 'say'
11	holde 'hold'

Table 1: The ten most frequent verbs with bare singular objects

The data in Table 1 lend some support to the claim that *have*-predicates are particularly likely to take bare singular objects, as the four most frequent verbs on the list are *have*-predicates.<sup>3</sup> A search for verbs with bare singular predicative NPs (i.e.

<sup>3</sup> However, the hits include a considerable number of sentences with bare nouns that may be argued to have a mass interpretation rather than a singular count interpretation. The reason for this is most likely

bare singulars with the syntactic function PREDLINK instead of OBJ) results in 255 matches, which all involve the verb *være* ‘be’. This is less than the 345 matches with *have*-predicates in Table 1. These numbers suggest that *have*-predicates constitute the class of verbs that co-occur with bare singulars most often, even if the copular verb is included among potential verbs.

The verbs in Table 1 are frequent also with corresponding objects with the indefinite article. In fact, the first five verbs listed in Table 1 are top five also if one runs a search for the most frequent verbs that take indefinite singular objects with a specifier. In other words, bare singulars are particularly likely to occur as arguments of the most frequent verbs. What is crucial is the relative frequency of verbs with bare singular objects compared to the relative frequency of verbs with other types of objects. For instance, the present search resulted in 178 hits for the verb *ha* ‘have’ with bare singular objects, while the corresponding search for *ha* and singular indefinite objects with a determiner resulted in 181 hits. This means that for the verb *ha*, bare singulars constitute almost fifty percent of all singular indefinite nominal objects of the verb, given that all hits are correct. This can be contrasted to other verbs, which have a much lower percentage of bare singular objects. To conclude, Table 1 lends some support to the claim in Borthen (2003) that *have*-predicates are particularly likely to take bare singular objects, but the question requires a more thorough empirical investigation and statistical analysis to be answered firmly.

Another more crucial insight of the search for verbs that take bare singular objects is that the border between the ‘conventional situation type’ construction and the ‘profiled *have*-predicate’ construction is not as clear cut as it appears to be in Borthen (2003). Some examples of sentences found with bare singulars and the verb *ha* ‘have’ are provided in (28)–(31), with both the verb and the bare singular in boldface.

- (28) *Skal du ikke ha fest da?*  
shall you not have party then  
‘Aren’t you going to have a party then?’
- (29) *Den fjerde mai hadde jeg bursdag.*  
the fourth May had I birthday  
‘The fourth of May was my birthday.’
- (30) *Den natta hadde ikke Kato mareritt.*  
that the night had not Kato nightmare  
‘That night Kato didn’t have a nightmare.’

---

that the mass–count distinction was not encoded in NorKompLeks (Nordgård 1998), the lexical resource that is the basis for the NorGram lexicon. Mass readings have been added by the annotators in the INESS project as they have been encountered during disambiguation, but there are certainly many mass nouns which have not received the proper encoding.

- (31) *Jeg har dessverre type, men det er ikke sikkert det varer  
I have unfortunately boyfriend but it is not certain it lasts  
lenge.  
long  
'I unfortunately have a boyfriend, but it's not for sure it will last long.'*

Bare singulars that occur as complements of *have*-predicates, such as those in (28)–(31), may well be part of verb phrases that denote conventional situation types such as having a party, having a birthday, having a nightmare, and having a boyfriend. This suggests that the ‘*have*-predicate’ construction and the ‘conventional situation type’ do not exist side-by-side as two distinct ways of generating bare singulars, as proposed in Borthen (2003).

#### 4.4 Bare singulars in idiomatic expressions

According to Sag et al. (2002, p. 2), multiword expressions are “idiosyncratic interpretations that cross word boundaries (or spaces)”. Sag et al. distinguish between two main types of expressions: lexicalized phrases and institutionalized phrases. Lexicalized multiword expressions are idiosyncratic with respect to their syntax and/or semantics, and they sometimes contain words which do not occur in isolation. Some are fixed expressions with rigid word order, while others are syntactically flexible. Institutionalized multiword expressions have normal syntactic and semantic properties, but the words that make them up co-occur with markedly high frequency. When Borthen (2003) made the claim that most bare singulars occur in multiword expressions, it was with the first multiword category in mind, i.e. expressions where the meaning and possibly also the syntax of the expression cannot be deduced from the meaning and the syntax of the individual words and the way they are put together.

Verbal idioms that involve a verb plus an object are a common type of multiword expression; well-known examples mentioned in Sag et al. (2002, p. 5) are *kick the bucket*, *shoot the breeze*, and *spill the beans*. These all involve definite objects, but bare singulars also occur, for example *give way*, *catch fire*, and *play possum*. In order to gather evidence to investigate the claim that most bare singulars occur in lexicalized multiword expressions, we searched for combinations of verbs and bare singular objects. The search expression is the same as in (27), but with the addition of *#x\_ >PRED #q*, since we want both the verb and the head noun to appear in the frequency list. As in 4.3 we ran the search only on the manually disambiguated part of the corpus, excluding fragment analyses. Table 2 lists the most frequent verb–noun combinations.

Six of the ten most frequent hits are analyzed as verbal idioms by NorGram, as can be seen by the predicate names that incorporate the lexical items that the multiwords consist of. The other four verb–noun combinations are also multiword expressions, although NorGram analyzes them compositionally and not as verbal idioms. But in addition to the highly frequent combinations shown in the table, the search results

Count	#p: value of V	#q: value of N	translation V+N
27	ha 'have'	rett&right 'right'	'be right'
26	få#øye*på 'get eye on'	øye 'eye'	'catch sight of'
17	legge#merke*til 'lay mark to'	merke 'mark'	'notice'
9	få#tak*i 'get hold of'	tak 'hold'	'obtain'
6	få 'get'	melding 'message'	'get word'
6	ha 'have'	råd*til 'affordance to'	'be able to afford'
6	ta#slutt 'take end'	slutt 'end'	'end'
6	holde#øye*med 'keep eye on'	øye 'eye'	'keep an eye on'
6	ha 'have'	tid*til 'time to'	'have time for'
6	sette#pris*på 'set price on'	pris 'price'	'appreciate'

Table 2: The ten most frequent combinations of verbs and their bare singular objects

also include many verb–noun combinations that occur only once and thus are unlikely to constitute multiword expressions. In (32)–(35) are some examples of unique combinations of verbs and bare singular nouns (with both the verb and the noun in boldface).

- (32) *Jeg har kjøpt interrailbillett.*  
 I have bought interrail ticket  
 'I have bought an interrail ticket.'
- (33) *Han har lys stemme og snakker fort.*  
 he has light voice and talks fast  
 'He has a high-pitched voice and talks fast.'
- (34) *Å pusse tennene til barna eller smøre matpakke, ga en enorm glede.*  
 to brush the teeth to the children or butter sandwich package, gave a enormous joy  
 'Brushing the kids' teeth or making their lunch provided enormous happiness.'
- (35) *Der er det kø ved disken.*  
 there is it line at the counter  
 'There is a line at the counter.'

To test whether *most* bare singulars occur in lexicalized multiword expressions is difficult without a manual check of all of the hits. This is beyond the scope of the

present paper. Here we can only report on our impression from browsing through the search results, and they suggest that bare singulars – indeed – occur in multiword expressions most of the time, as claimed in Borthen (2003).

#### 4.5 Bare singulars in non-canonical sentence structures

In Borthen (2003) it is claimed that the acceptability of bare singulars is mostly unaffected by syntactic alternations. From this it follows that it should be possible to find instances of bare singulars in sentences with non-canonical sentence structure. For instance, it is expected that bare singulars can occur as topicalized and left-dislocated objects and as the postcopular element of cleft sentences. These claims can be tested through searches in NorGramBank. Since we assumed that bare singulars in these constructions would be rare, we searched in the entire corpus, with the only restriction being that we did not search among fragment analyses.

Topicalized bare singulars can be searched for by adding the constraints in (36) to the search expression in (27).

- (36) #w\_ >\* #v\_ &  
#w\_ >TOPIC #x\_

The first conjunct in (36) says that there is an f-structure #v\_ (the f-structure of the verb) that is a sub-f-structure of another f-structure #w\_. The second conjunct says that this f-structure (#w\_) has an attribute TOPIC with the value #x\_ (the f-structure of the bare singular). The sentences in (37)–(39) provide examples of topicalized bare singulars identified through this search.

- (37) *Men hønsehjerne kan de ha selv!*  
but hen brain can they have self  
'Let them be birdbrains!'
- (38) *Men personlig rådgiver kan jeg velge selv.*  
but personal adviser can I choose self  
'But I can choose a personal adviser myself.'
- (39) *For jordkjeller hadde farmor også hatt.*  
for earth cellar had grandma also had  
'For Grandma had also had a cellar.'

These data show that bare singulars, just as other objects, can be moved from their base position and placed in topic position.

Left dislocation differs from topicalization in that the left-dislocated constituent co-occurs with a coreferential pronoun in subject or object position in the sentence. When the left-dislocated constituent is a bare singular, the coreference relation does not exist at the token level, but at the type level. That is, the left-dislocated constituent and

the coreferential pronoun refer to the same type of thing. Left-dislocated bare singulars can be identified by simply adding #x\_ >ADJUNCT-TYPE 'left-disloc' to the search expression in (24). This constraint states that the bare singular must have the attribute-value pair ADJUNCT-TYPE 'left-disloc' in its f-structure. The sentences in (40)–(42) are examples of left dislocation found by this search expression.

- (40) *Skole – det er OK for noe, som å bli lærer eller forsker.*  
 school that is OK for something like to become teacher or researcher  
 'School – it's OK for something, like becoming a teacher or a researcher.'
- (41) *Men statsministerbolig, det kan minne om stormannsgalskap?*  
 but prime minister residence that can remind of megalomania  
 'But the residence of the prime minister, isn't that reminiscent of megalomania?'
- (42) *Kniv i ranselen, det kunne være livsfarlig.*  
 knife in the satchel that could be deadly  
 'A knife in the satchel, that could be deadly.'

As for clefted bare singulars, these can be identified by adding #y\_ >FOCUS #x\_ to the search expression for bare singulars in (24). This uniquely identifies the element in the postcopular position of cleft sentences. Some of the resulting sentences are presented in (43)–(45) below.

- (43) *Det var først og fremst boktyv mor var.*  
 it was first and foremost book thief mother was  
 'It was above all a book thief my mother was.'
- (44) *Det var hund hun ønsket seg.*  
 it was dog she wished herself  
 'It was a dog she was wishing for.'
- (45) *En stund var han overbevist om at det var maler han ville bli.*  
 a while was he convinced about that it was painter he wanted become  
 'For a while he was convinced that it was a painter he wanted to become.'

In sum, we have found new evidence in favor of the claim that Norwegian bare singulars can take part in syntactic alternations and appear in various types of non-canonical sentence types, here exemplified by topicalization, left-dislocation and clefting. Despite the fact that we have presented only a handful of examples, the new data are more convincing than what was provided in Borthen (2003), since authentic examples (that the reader can consult) must be said to constitute more convincing evidence than invented examples whose acceptability is merely judged by the researcher.

The complete set of output sentences that resulted from the searches presented in this section do contain some undesired hits (wrong analyses), as can be expected when complicated constructions in natural language are automatically parsed and stochastically disambiguated. However, since NorGramBank allows for the manual creation of subcorpora, the output may be manually cleaned if desirable. That way one may compare the relative frequency of bare singulars in the investigated constructions with other nominals, which will add yet another level of insight.

## 5 Conclusion

In this paper we have illustrated how bare singulars can be searched for in NorGramBank. The study has provided some (though not full) support for the following four claims about Norwegian bare singulars put forth in Borthen (2003):

1. Nouns with very general descriptive content tend not to appear as bare singulars;
2. *Have*-predicates are the most frequent bare singular-selecting verbs;
3. Most bare singulars are part of multiword expressions;
4. Bare singulars can occur in non-canonical sentence structures; for instance they can be topicalized, left-dislocated and clefted.

More importantly, the present study has revealed some fundamental problems with the original analysis of bare singulars, due to the availability of huge amounts of authentic data. One such observation is related to the fact that the machine annotations sometimes fail to pick out bare singulars uniquely. This reveals an interesting fact: in addition to phrases that clearly have a singular count interpretation and phrases that clearly have a mass interpretation, there are many indefinites that are hard to categorize as either one. Similar observations have been made by Halmøy (2016). This questions the very premise that bare singulars constitute an interesting category on their own – which in turn means that there are probably no constructions (or grammar rules or principles) that license bare singulars specifically in Norwegian.

A second observation that points in the same direction is the fact that many examples with bare singulars fit at least two out of the four ‘bare singular’-licensing constructions proposed in Borthen (2003). This is evidence against the assumption that the four alleged constructions proposed in Borthen (2003) exist side-by-side.

The data presented in this paper may be taken to indicate that bare singulars are not licensed through a set of constructions that are part of the grammar, but rather constitute a phenomenon on a par with the choice between an indefinite or a definite article. If so, the idea that determiners are obligatory for all nominal arguments in Germanic and Romance languages (see e.g. Longobardi, 1994) is threatened. Another possibility is that the secrets of bare singulars lie in the understanding of multiword expressions, ranging from fixed, non-compositional lexicalized phrases to fully compositional institutionalized ones whose status as multiword expressions relies solely on their frequency. Further studies of bare singulars – most likely based on large-scale searchable corpora – will show which of these, or other, approaches will be most successful.

## Acknowledgments

We would like to thank our reviewers for their helpful comments and suggestions. We would also like to acknowledge the outstanding work that Helge Dyvik has done with NorGram and NorGramBank. Without his efforts through many years, studies such as the one reported on in this paper would not be possible.

## References

- Asudeh, Ash and Line Hove Mikkelsen (2000). “Incorporation in Danish: Implications for Interfaces”. In: *Grammatical Interfaces in HPSG*. Ed. by Ronnie Cann, Claire Grover, and Philip Miller. CSLI Publications, pp. 1–15.
- Borthen, Kaja (2003). “Norwegian bare singulars”. PhD thesis. Trondheim: Norwegian University of Science and Technology.
- Bresnan, Joan (2001). *Lexical-Functional Syntax*. Malden, MA: Blackwell.
- Dalrymple, Mary (2001). *Lexical Functional Grammar*. Vol. 34. Syntax and Semantics. San Diego, CA: Academic Press.
- De Swart, Henriette and Joost Zwarts (2009). “Less form – more meaning: Why bare singular nouns are special”. In: *Lingua* 119, pp. 280–295.
- Delsing, Lars-Olof (1993). “The internal structure of noun phrases in the Scandinavian languages. A comparative study”. PhD thesis. University of Lund.
- Dyvik, Helge (2000). “Nødvendige noder i norsk: Grunntrekk i en leksikalsk-funksjonell beskrivelse av norsk syntaks [Necessary nodes in Norwegian: Basic properties of a lexical-functional description of Norwegian syntax]”. In: *Menneske, språk og felleskap*. Ed. by Øivin Andersen, Kjersti Fløttum, and Torodd Kinn. Oslo: Novus forlag, pp. 25–45.
- Dyvik, Helge, Paul Meurer, Victoria Rosén, Koenraad De Smedt, Petter Haugereid, Gyri Smørdal Losnegaard, Gunn Inger Lyse, and Martha Thunes (2016). “NorGramBank: A ‘Deep’ Treebank for Norwegian”. In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*. Ed. by Nicoletta Calzolari,

- Khalid Choukri, Thierry Declerck, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Asunción Moreno, Jan Odijk, and Stelios Piperidis. ELRA. Portorož, Slovenia, pp. 3555–3562.
- Halmøy, Madeleine (2016). *The Norwegian Nominal System*. Berlin/Boston: Walter de Gruyter.
- Kallulli, Dalina (1999). “The Comparative Syntax of Albanian: On the Contribution of Syntactic Types to Propositional Interpretation.” PhD thesis. University of Durham.
- Kiefer, Ferenc (1994). “Noun incorporation in Hungarian”. In: *Acta Linguistica Hungarica* 40.1–2, pp. 149–177.
- Meurer, Paul (2012). “INESS-Search: A search system for LFG (and other) treebanks”. In: ed. by Miriam Butt and Tracy Holloway King. LFG Online Proceedings. Stanford, CA: CSLI Publications, pp. 404–421.
- Nordgård, Torbjørn (1998). “Norwegian Computational Lexicon (NorKompLeks)”. In: *Proceedings of the 11th Nordic Conference on Computational Linguistics (NoDaLiDa), Copenhagen*.
- Rosén, Victoria, Koenraad De Smedt, Paul Meurer, and Helge Dyvik (2012). “An Open Infrastructure for Advanced Treebanking”. In: *META-RESEARCH Workshop on Advanced Treebanking at LREC2012*. Ed. by Jan Hajič, Koenraad De Smedt, Marko Tadić, and António Branco. Istanbul, Turkey, pp. 22–29.
- Sag, Ivan, Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger (2002). “Multiword Expressions: A Pain in the Neck for NLP”. In: *Lecture Notes in Computer Science. Proceedings of the Third International Conference on Computational Linguistics and Intelligent Text Processing*. Vol. 2276. Springer, pp. 1–15.
- Schmitt, Cristina and Alan Munn (1999). “Against the nominal mapping parameter: Bare nouns in Brazilian Portuguese”. In: *Proceedings NELS 29*, pp. 339–353.
- Stvan, Laurel Smith (1998). “The Semantics and Pragmatics of Bare Singular Noun Phrases”. PhD thesis. Northwestern University.