Sentence-initial indefinite subjects in English and Norwegian

Hilde Hasselgård

University of Oslo (Norway)

Abstract: The present study uses the English-Norwegian Parallel Corpus to investigate the frequency and use of indefinite noun phrases as subjects in English and Norwegian. Since subjects in both languages tend to appear clause-initially, indefinite subjects represent a deviation from the information principle. The clearest difference between the languages is the greater frequency of indefinite subject NPs in English. The lexicogrammatical features of the indefinite subjects and their immediate contexts are relatively similar in both languages. The indefinite subjects most commonly occur with intransitive verb phrases, and often in clauses with presentative or generic meaning. Translation correspondences of indefinite subjects show that the subject NP is retained in congruent form in the majority of cases, but more changes are made in translations from English into Norwegian than the other way round. This is taken to support the findings of the contrastive analysis and furthermore indicates that the light subject constraint is applied more strictly in Norwegian than in English.

Keywords: indefinite NP, subject, information structure, presentative, English, Norwegian

1.   Introduction

In languages such as English and Norwegian, subjects realized by indefinite noun phrases seem anomalous as regards information structure. Both languages tend to place the subject clause-initially (Hasselgård 2004) and are furthermore assumed to organize sentences in agreement with the information principle, that is, a distribution of given and new information in the clause that corresponds "to a gradual rise in information load" (Biber *et al.* 1999: 896). Thus, the beginning of the sentence – the subject position – is associated with given information. Since indefinite NPs are typically associated with new information, they should be disfavoured as subjects (Biber *et al.* 1999: 269), while definite noun phrase realization should be expected (Prince 1992). Yet, indefinite NP subjects are found in both languages, as shown by (1) and (2) from the English-Norwegian Parallel Corpus (ENPC).[1]

(1)   *En tyv* er ikke voldsom, men stillferdig. (KF1)
        A thief is not violent but quiet. (KF1T)

---

[1] Examples from the ENPC are given with the original first. Norwegian examples are followed by a literal translation if their wording diverges from that of their English translations/sources found in the corpus.

(2)   *A rat* crept out of the hole behind the dresser ... (GN1)
      En rotte kom ut av hullet bak kommoden ... (GN1T)

Previous contrastive studies of English and Norwegian have noted cross-linguistic differences in the use of indefinite subjects, particularly that Norwegian is less tolerant of them than English (Ebeling 2000: 191, Hasselgård 2004: 201, Johansson 2004: 41). However, these studies had other primary concerns (see Section 3). The present study zooms in on sentence-initial indefinite subjects in both languages. It first compares original texts in English and Norwegian to answer the following research questions:

-   How frequent are indefinite subjects in English and Norwegian?

-   What are the lexicogrammatical features of such subjects (e.g. +/- modification) and their verbs (e.g. +/- transitive)?

-   What are the contexts for indefinite subjects – and are they the same in both languages?

The features of the subject NPs include their semantics, particularly referent type and specificity. Contexts include features of the verb phrase as well as other clause elements, particularly objects/complements and locative adverbials. For example, both English and Norwegian have indefinite subjects in contexts of generic reference, as in (1), and in clauses denoting the existence or appearance of a specific subject referent, as in (2).

The second part of the study probes further into cross-linguistic differences in the use of indefinite subject NPs by exploring their translations, the frequency with which changes are made, and the nature of such changes. Given that indefinite subjects are comparatively rare in both languages, the translation principle of normalization (e.g. Baker 1996: 183) might prompt translators to make changes to either the form of the subject (as in (3)) or the placement of the indefinite NP, as in (4).

(3)   *Cultured pearls* are in the vault. (DF1)
      Kunstperlene ligger i velvet. (DF1T)
      "The cultured pearls lie in the vault."

(4)   *En gammel kvinne* tok imot oss, vennlig, men uten smil. (JW1)
      "An old woman received us…"
      We were received by *an old lady*, in a friendly but unsmiling fashion. (JW1T)

Since indefinite subjects are assumed to be less frequent in Norwegian than in English, translations into Norwegian are expected to involve a change of the subject NP more often than translations into English, especially if the NP has specific reference, as is the case in (3) and (4). Information structure and semantics, especially the notions of existence/appearance are also expected to play a role. The overall aim is to learn more about information structuring in both languages, in particular the conditions in which apparent breaches of the given-to-new principle are tolerated, and whether (and how) the languages differ in this respect.

2.   Indefinite noun phrases in English and Norwegian

English and Norwegian indefinite NPs are rather similar in form: there are indefinite articles for the singular only (English *a/an*, Norwegian 'bokmål' *en/ei/et* and Norwegian 'nynorsk' *ein/ei/eit*). Indefinite plurals and uncountable nouns occur with no article or with certain types of quantifying determiners (e.g. cardinal numbers, *some*; see Lyons 1999: 33 ff). The English indefinite articles vary according to the phonological context and the Norwegian ones according to grammatical gender (masculine, feminine, neuter), but they have the same functional properties and relatively similar conditions of use (for some exceptions, see Hasselgård *et al*. 2012: 120). It may be noted that definiteness of nouns is generally marked by suffixes rather than articles in Norwegian. The forms are shown in (5), in which the inflectional suffixes have been underlined.

(5)   en hund – hund<u>en</u> – hund<u>er</u> – hund<u>ene</u>
       "a dog – the dog – dogs – the dogs"

While the indefinite article clearly marks an NP as indefinite in both languages, NPs occurring without a determiner or with determiners other than the articles may be harder to classify. Lyons (1999: 33) argues that e.g. cardinal numbers "do not encode [-Def]", and concludes that a noun phrase may be classified as indefinite simply through "the absence of any definite determiner" (*ibid*.). In cases of doubt, Abbott (2006) suggests locative existentials as a test for definiteness. That is, if it is possible to insert an NP in the formula "There is NP", with no special interpretation of the existential, then the NP is indefinite.[2] In my classification of definite and indefinite noun phrases I have mainly followed Lyons, with the exception of NPs with the determiner *all* or the corresponding Norwegian *alle*. According to Lyons (1999: 44) *all*, being a universal quantifier, is a definite marker when used as a determiner. However, in both English and Norwegian *all(e)* may or may not co-occur with the definite article/suffix, and so I have chosen to classify *all/alle* + indefinite plural as indefinite. NPs with *each/every,* on the other hand, have not been included, as they are generally associated with inclusiveness and hence definiteness (Lyons 1999: 32) and do not co-occur with (other) definiteness markers. Other notable exclusions are NPs introduced by *one of / en/ei/et av* + def. NP, which are 'containing inferables' (Prince 1981) and thus regarded as definite, and the indefinite pronouns *one/en* (as in *One might get £300,000… / En kunne få £300 000 …* (FF1)), which is homonymous with the numeral in both languages. An overview of the types of indefinite NPs included in the material is given in Table 1.

Table 1. Types of indefinite noun phrases in English and Norwegian.

Singular article	*a woman*	*en kvinne*
Singular, no determiner	*water*	*vann*
Singular quantifier	*one/no man*	*en/ingen mann*
Plural, no determiner	*families*	*familier*
Plural quantifier	*several witnesses* *a couple of books* *two republicans*	*flere vitner* *et par bøker* *to republikanere*
Comparative determiner	*other youngsters*	*andre ungdommer*

---

[2] Two of her examples are *There is a book in the shop window* and *There is the book in the shop window*, of which the latter requires the special interpretation of 'list existential' to be grammatical (Abbot 2006).

An indefinite noun phrase can have specific, non-specific or generic reference in both English and Norwegian, as explained by Biber *et al.* (1999: 260) and exemplified in (6)–(8):

(6) SPECIFIC: *Mosquitoes and fireflies* had come in. (BO1)
Moskitoer og ildfluer var kommet inn. (BO1T)

(7) NON-SPECIFIC: *En drosjesjåfør* ville kjørt helt bort til Hansdals inngangsdør… (EG1)
"A taxi driver would (have) driven right over to Hansdal's front door…"
A taxi driver would have driven the old lady right to the door. (EG1T)

(8) GENERIC: *A child* needs security. (ABR1)
Et barn trenger trygghet. (ABR1T)

According to Hawkins,

> The speaker performs the following speech acts when using an indefinite article to achieve (specific) indefinite reference: He (a) introduces a referent (or referents) to the hearer; and (b) refers to a proper subset, i.e. not-all, of the potential referents of the referring expression. (1978: 186–187)

While non-specific reference involves reference to "any member of a total class" (*ibid*.: 215), a generic referent is considered "a typical representative of its class" (*ibid*.). These principles are assumed to apply to both English and Norwegian.

3.   Literature review

As noted above, the use of an indefinite subject in sentence-initial position typically involves a deviation from normal or unmarked information structure. Quirk *et al.*, for example, observe that "the organization of sentences … generally presumes that a sentence begins with a reference to 'given' information and proceeds to 'new' information" (1985: 1402). Although there is no syntactic rule against using an indefinite subject as NP, Quirk *et al.* argue that "… a certain awkwardness is sensed where the recipient is expected to interpret a theme as entirely new and unconnected with anything previously introduced" (*ibid*.). Quirk *et al.*'s views on word order and information structure are clearly influenced by the Prague School and its concept of Functional Sentence Perspective (FSP), in which information structure, or the distribution of communicative dynamism (CD), is seen as a crucial structuring principle for word order. According to Firbas (1992: 118), "the principle of FSP arranges the sentence elements in a Th[eme] – Tr[ansition] – Rh[eme] sequence. If asserting itself to the full extent, it … induces the sentence to display what has been termed 'the basic distribution of CD'", which implies a gradual rise (*ibid*.: 10).

Firbas (1992) devotes a section to *context-independent* subjects, i.e. subjects that do not link up with the preceding text or the situational context, and that are typically realized by indefinite noun phrases. Such subjects are likely to occur with (mostly intransitive) verbs that denote "appearance or existence on the scene" (1992: 60). In these sentences, the subject carries the highest degree of communicative dynamism (CD). However, as Dušková (1999: 249) argues, placing new information at the beginning of a sentence runs counter to the basic distribution of CD. Thus context-independent, rhematic subjects in subject position constitute a "deviation from basic distribution of CD" (*ibid*.: 254), and moreover, may represent an unresolved conflict between the basic distribution of CD and the grammatical word order principle (*ibid*.: 260). In a later paper, Dušková (2015) shows that sentence-initial rhematic

subjects in English, though rare, tend to occur in the pattern S-V-A and thus belong to the presentation scale (2015: 21–23).

Similar views on sentence openings and information structure are expressed by other functionally oriented linguists. For example Chafe (1994: 83–92) argues that subjects are governed by the *light subject constraint*: i.e., in conversation subjects normally refer to given or accessible information, and the new information (if any) comes later in the clause.[3] Although Halliday (1994) makes a distinction between theme (what comes first) and given information (what the speaker expects the hearer already knows), he concedes that the theme is typically selected from given information (1994: 299). A clause that starts with new information will typically have a marked information focus, provided it also contains some recoverable information (*ibid*.: 301).

Biber *et al*. (1999) investigate the distribution of definite and indefinite noun phrases (as marked by articles) across syntactic functions. They find that "indefinite subjects are much less common [than definite ones] but by no means rare" (1999: 269). This applies in all four registers studied (fiction, news, academic prose and conversation). The object function is much more characteristic of indefinite NPs, and to some extent prepositional complement (*ibid*.). This distribution is explained by the information principle: "When new information is introduced in subject position, it is marked as thematically important" (*ibid*.). Similarly, in a study of (English) subjects and information status, Prince (1992: 316) reports that only 10% of indefinite NPs in a single text had subject function as opposed to 38% of definite NPs. She suggests that definiteness is "a grammaticization of Hearer-status" (*ibid*.: 317), i.e. a marker that an entity is assumed to be identifiable by the hearer/reader. De Hoop & Krämer (2006) offer an alternative, processing-based explanation for the scarcity of indefinite subjects, arguing that they involve a "conflict of interpretation" because subjects in general (in the standard subject position) "favour a referential reading" while indefinites "favour a nonreferential reading" (2006: 119).

The Norwegian reference grammar (Faarlund *et al*. 1997) states that the subject usually expresses given information in Norwegian too: it does not normally convey new information (1997: 691). There are, however, contexts that make a "new" subject more likely, for example if it refers to an element within a recoverable group (i.e. what Prince 1981 refers to as "containing inferable") or if it is singled out for contrastive focus (Faarlund *et al*. 1997: 692). Furthermore, both generic reference and the presence of certain (plural) quantifiers also make an indefinite subject acceptable (*ibid*.: 690). In a paper on the typology of subjects, Faarlund questions the acceptability of the Norwegian sentence *Ein mann arbeider på vegen* ('a man works on the road') because the subject, though meeting the subjecthood requirement of agency, does not convey given information (1988: 199). A presentative with *det* ('there') is suggested as a fully acceptable alternative, namely *Det arbeider ein mann på vegen* ('there works a man on the road') (*ibid*: 200).

Indefinite subjects have not, to my knowledge, been the main topic of any cross-linguistic study of English and Norwegian, but they have come up in previous studies of presentative constructions (Ebeling 2000), thematic choice (Hasselgård 2004, 2005) and the extent to which syntactic subjects are preserved in translation (Johansson 2004).

A construction with an indefinite subject, an intransitive verb, and most typically a locative adverbial may have a presentative function even in the absence of the explicit presentative marker *there/det*, hence the term "bare presentative" (Ebeling 2000:157 ff). Ebeling discusses both subject-initial bare presentatives (as in example (3) above) and those

---

[3] The light subject constraint (Chafe 1994: 85) is also referred to as the "light starting point constraint". Chafe (1986) acknowledges that writing may not adhere as strictly to it as conversation does.

with locative inversion (*ibid*.: 160, 173).[4] An important finding is that English bare presentatives have a greater variety of lexical verbs than full presentatives with *there* (*ibid*.: 161). The difference in verb choice between the two construction types is somewhat smaller in Norwegian, whose full *det*-presentative construction is more flexible than the English one as regards its verbal collocates (*ibid*.: 176). Ebeling finds fewer examples of subject-initial bare presentatives in Norwegian than in English (*ibid*.: 179). He therefore suggests that "Norwegian is more susceptible to FSP than English" (*ibid*.). Locative-initial bare presentatives, on the other hand, are more frequent in Norwegian (*ibid*.: 186). These findings reveal "a greater tendency in Norwegian to avoid indefinite NPs in subject position than in English" (*ibid*.: 187).

Johansson (2004) studies the change or preservation of the subject in translation from English into Norwegian. He finds that the vast majority of subjects (around 90%) are preserved in translation (2004: 33, 49) probably because "the subject is essential in the building of both sentences and texts" (*ibid*.: 32). However, "changes may be induced by formal differences between the languages, or they may be due to differences in stylistic norms" (*ibid*.). Johansson notes that English indefinite subjects may be rendered by a construction with the anticipatory subject *det* ('it/there') in the Norwegian translation. This is related to "a lower tolerance for indefinite subjects in Norwegian than in English" (*ibid*.: 41).

English and Norwegian clause themes have been studied contrastively by Hasselgård (e.g. 2004, 2005) and Rørvik (2013).[5] (See also Rørvik and Monsen, this volume.) Although Norwegian is a V2 language while English is an SV language, there are great cross-linguistic similarities. However, an important difference is that Norwegian has more fronted non-subjects, particularly adverbials, as well as more anticipatory subjects realized by *det* ('it/there') (Hasselgård 2004: 192). The latter finding was "associated with a greater tendency in Norwegian than in English to avoid indefinite subject NPs and NPs conveying new information in thematic position" (*ibid*.: 208). It was found that translated texts in both languages show a great degree of similarity with their source texts, indicating that the differences between English and Norwegian originals pertain more to preferences than to syntactic constraints (*ibid*.: 190). The study indicated that "word order is freer Norwegian than in English, [but] *information structure* seems to be more flexible in English. I.e. English is more tolerant of new information in the Theme, while Norwegian has a stronger preference for 'light' Themes" (*ibid*.: 208). In a follow-up study it was argued that Norwegian has a greater preference than English for syntactically and informationally light themes (Hasselgård 2005: 46) due to differences in the functional load of themes between the languages: Norwegian themes are less often associated with contrast and more often with cohesion, and may therefore be less prominent (*ibid*.).

4.   Material and method

The material for this study consists of declarative main clauses with indefinite subjects from the fiction part of the English-Norwegian Parallel Corpus (ENPC), i.e. 30 text extracts of 10,000–15,000 words in each language plus their translations (Johansson *et al*. 1999/2002, Section 1.2).[6] As the corpus is not parsed, subjects were retrieved (from the original texts) using a combination of lexical searches, positional criteria and PoS tags. Indefinite singular

---

[4] An example of locative inversion is *Behind the policemen was a middle-aged woman …* (DF1) > *Bak politimennene var en middelaldrende dame …* (Ebeling 2000: 159).
[5] In these studies, theme is defined as the first clause element with an experiential/referential function plus any preceding elements (Halliday and Matthiessen 2014: 105).
[6] The corpus was accessed from the PerlTCE interface (http://www.tekstlab.uio.no/cgi-bin/omc/PerlTCE.cgi).

subjects were retrieved by searching for indefinite articles and the numerals *one/én/ett* in first and second position in an s-unit.[7] Indefinite plurals and uncountable nouns were retrieved by searching for nouns in first position and in second position without a preceding definite determiner. Most of the resulting concordances had fairly (or very) poor precision, but were cleaned up manually to remove irrelevant hits, where the initial NP was either not indefinite or not a subject. The resulting concordances with their translations were transferred to a database and annotated for lexicogrammatical features and type of translation correspondence, as specified below.

The search method has (at least) two shortcomings. Most importantly, it misses indefinite subjects that are not sentence-initial. These include notional subjects in *there/det* presentatives, subjects occurring after initial adverbials of more than one word (most noticeable for English), subjects in post-verbal position (most noticeable for Norwegian, which is a V2 language) and subjects in conjoined clauses. There is no obvious way in which these subjects could have been retrieved reliably except manually. Since non-initial indefinite subjects are not quite so blatantly in breach of the information principle, this shortcoming, though not trivial, was deemed acceptable for the purposes of this study. The second shortcoming is that the PoS tagging, particularly for Norwegian, is not 100% reliable. Sifting through the concordance lines, I found examples of words wrongly tagged as nouns, which makes it likely that some nouns are wrongly tagged as something else. Short of reading through the whole corpus, there is little to be done about this. Note that Ebeling (2000: 158) used slightly different, and more comprehensive, search procedures to identify instances of bare presentatives, but still acknowledged less than full recall.

5.   Contrastive analysis

This section compares indefinite subjects in English and Norwegian original texts. In addition to frequency information, consideration is given to syntactic and semantic features of the subject NP, the verb phrase and the clause pattern in which the subject appears in order to investigate both the properties and the contexts of indefinite subjects.

5.1   Frequency and form of indefinite subjects in original texts

As expected, indefinite subjects are more frequent in English than in Norwegian original texts, as shown in Table 2. The raw frequencies are comparable since the amount of text in the two languages is similar. The last row of the table gives the frequency of indefinite subjects per 1000 s-units. This was considered the best available measure considering the opportunity of occurrence for a sentence-initial subject, even though not all s-units are complete sentences with a grammatical clause structure. Like the raw frequencies, the relative frequency of indefinite subjects per 1000 s-units is greater in English.[8]

---

[7] An s-unit is roughly equivalent to an orthographic sentence (Johansson *et al*. 1999/2002, section 2.3.4), i.e. a stretch of language starting with a capital letter and ending with a final punctuation mark.
[8] The (raw) frequency difference relative to the number of s-units is significant (LL=49.33, p<0.0001).

Table 2. Frequency of indefinite subjects in English and Norwegian (original texts).

Subject form	English	Norwegian
Sg N, indefinite article	273	268
Sg N, no determiner	112	55
Sg N, quantifier	34	13
Pl N, no determiner	267	161
Pl N, quantifier	38	49
Comparative determiner	8	2
TOTAL	**732**	**548**
Freq. per 1000 s-units	24.9	16.8

Singular noun phrases with indefinite articles are the most common realization of indefinite subject in both languages. Both singular and plural nouns with no determiner are much more frequent in English than in Norwegian indefinite subjects, and thus account for most of the frequency difference between the languages. Indefinite subjects are used in all the original corpus texts except one in Norwegian, with both the highest and the lowest number per text being higher in English (49 vs. 38 examples and 9 vs. 6). Thus, the material supports the hypothesis, derived from previous contrastive studies, that indefinite subjects are (even) less common in Norwegian than in English.

5.2 Lexicogrammatical features of the subject NP

In the analysis of indefinite subject noun phrases, the following features were noted:

- NP complexity, i.e. the absence/presence of premodifier(s) and postmodifier(s).

- NP semantics, i.e. whether the referent of the NP is human, (non-human) animate, concrete, or abstract.

- NP specificity, i.e. whether the reference of the NP is specific, non-specific, or generic.

An overview of noun phrase complexity is shown in Table 3. Simple (unmodified) noun phrases are most common, with a slightly higher proportion in English than in Norwegian. Premodification and postmodification are equally common in English, but in Norwegian premodified NPs are almost twice as common as postmodified ones.

Table 3. Complexity of NPs functioning as indefinite subjects.

	English		Norwegian	
	N	%	N	%
Simple NP	416	56.8	280	51.1
Premodifier + head	134	18.3	150	27.4
Head + postmodifier	137	18.7	82	15.0
Premodifier + Head + postmodifier	45	6.1	36	6.6
	732	99.9	548	100.1

Table 4 shows categories of referents of the indefinite subject NPs. Concrete (non-animate) referents are most common in both languages, followed by human referents. Abstract and (non-human) animate referents are comparatively rare. The proportions of each type of referent are fairly similar between the languages.

Table 4. Referents of indefinite subject NPs.

Subject NP referents	English		Norwegian	
	N	%	N	%
Human	238	32.5	212	38.7
Animate	51	7.0	26	4.7
Concrete	356	48.6	248	45.3
Abstract	87	11.9	62	11.3
	732	100	548	100

The analysis of noun phrase specificity also gave a similar distribution between the languages, as Table 5 shows. Specific reference is most common, followed by generic and non-specific. Example (9) thus shows a prototypical representative of an indefinite subject: a singular, unmodified noun with an indefinite article, a concrete referent and specific reference.

(9)   *En stemme* kom fra treet: (THA1)
       A voice came from the tree: (THA1T)

Table 5. Specificity of indefinite subject noun phrases.

NP specificity	English		Norwegian	
	N	%	N	%
Specific	418	57.1	317	57.8
Non-specific	137	18.7	102	18.6
Generic	177	24.2	129	23.5
	732	100	548	99.9

The proportion of generic reference in indefinite subjects seems high compared to Biber *et al.*'s (1999: 266) report that less than 2.5% of definite NPs in fiction had generic reference. (Figures for generic indefinite NPs are not given separately.) This indicates that indefinite subjects may be particularly prone to generic interpretation. Note that generic reference is possible with both singular and plural nouns in both languages, as shown in (10) and (11); see also Biber *et al.* (1999: 265).

(10)   *Scholars and artists* have no morals whatever about grants of money. (RDA1)
       Forskere og kunstnere har ingen som helst moralske skrupler overfor stipendier. (RDA1T)

(11)   Men *en soldat* så da ikke slik ut. (KAL1)
       But then *a soldier* did not look like that. (KAL1T)

With respect to the distinction between given and new information, Chafe (1994) argues that generic reference has a special status:

> …sharing knowledge of generic referents is different from sharing knowledge of particular referents. Knowing a category, like the category that allows something to be called an elephant,

entails knowing something about a *typical* instance of that category, whereas the sharedness involved in identifiability depends on knowing a *particular* instance. (Chafe 1994: 103)

An indefinite NP with generic reference may thus not be considered to convey new information to the same extent as one that introduces a new specific referent. If this is the case, a generic sentence-initial subject will seem less objectionable as it does not so clearly violate the information principle.

5.3   Lexicogrammatical features of the verb phrase

The following features were noted in the analysis of verb phrases occurring with sentence-initial indefinite subjects:

-   **Voice**: active/passive[9]

-   **Transitivity**: intransitive, monotransitive, ditransitive, complex transitive, copular

-   **Semantics** (cf. Halliday 1994): material, mental, verbal, existence/ appearance, attributive, identifying, possessive

The classification of verb semantics is based on Halliday's process types (1994: 173), but modified to include a category of verbs of existence and appearance, mostly inspired by the presence of this category in FSP analyses of the presentation scale (Firbas 1992: 59, 67). The category overlaps with Halliday's existential, relational and material processes, thus reducing or replacing these categories in comparison with a purely systemic-functional analysis. Further, behavioural processes have been combined with material (as in Matthiessen 1995).

The distribution of active and passive voice is similar in both languages, with roughly 8.5% of clauses with indefinite subjects in the passive. This is slightly above the ratio of passives in fiction found by Biber *et al*. (1999: 476), who do not report exact figures, but present a diagram where the percentage of passives seems to be around five.

The two languages also have similar distributions of transitivity types. As shown in Table 6, approximately half of the clauses are intransitive in both languages, with mono-transitive verbs accounting for just over a quarter. Copular patterns take up about a fifth, while ditranstive and complex transitive verbs are infrequent in both languages. Note that transitivity type has been assigned from the verb in context, not from the lexeme, so that any clause not containing a grammatical object or predicative (including passives) have been classified as intransitive.

Table 6. Verb transitivity in clauses with indefinite subjects (original texts).

Verb transitivity (in context)	English		Norwegian	
	N	%	N	%
Intransitive	361	49.3	281	51.3
Monotransitive	193	26.4	154	28.1
Copular	156	21.3	98	17.9
Ditransitive	12	3.0	9	2.7
Complex transitive	10		6	
Total	732	100	548	100

---

[9] All verb phrases not marked for passive voice have been classified as active.

Intransitives are more common than expected from the general distribution of transitivity types reported by Oostdijk & de Haan (1993: 48). In their study of the distribution of clause patterns in main clauses, intransitives are only slightly more frequent than monotransitive and copular patterns, at 30%, 25% and 24%, respectively. It is thus a fair assumption that the intransitive clause pattern is a favourable context for indefinite subjects.

As regards the verb semantics, the most frequent process type is material, followed by attributive for English and existence/appearance for Norwegian (see Table 7).[10] The results are difficult to compare to other studies because of the modifications of the classification system. However, the proportional distribution of process types in clauses with indefinite subjects appears to be fairly similar to the general distribution of process types reported in Matthiessen (1999).

Table 7. Semantic types of verbs in clause with indefinite subjects (original texts)

Verb semantics	English		Norwegian	
	N	%	N	%
Material	378	51.6	261	47.6
Attributive	131	17.9	78	14.2
Ex/app	103	14.1	103	18.8
Mental	43	5.9	25	4.6
Verbal	30	4.1	40	7.3
Identifying	25	3.4	22	4.0
Possessive	22	3.0	19	3.5
Total	732	100	548	100

Based on the most frequent verb phrase properties, a prototypical sentence with an indefinite subject would thus be as in (12) in both English and Norwegian: it has active voice, an intransitive verb and refers to a material process.

(12) Et kaldt solgløtt *gnistret* i rutene på Deichmanske Bibliotek. (BV1)
A bleak ray of sunshine *sparkled* in the windows of the Deichman Library. (BV1T)

5.4 Lexicogrammatical features of the clause

Besides the subject and the verb, I considered the presence of other constituents in clauses with indefinite subjects. In clauses with a direct object or predicative (combined here under the label 'complement', as in Halliday 1994: 80) the realization of the post-verbal element was analysed, for example to determine the extent to which the clause conveys given/identifiable information in a later position than the indefinite subject. As shown in Table 8, definite NPs (including pronouns) are the most common realization of complements in English, but not in Norwegian, where indefinite complements are preferred. Especially complements realized by indefinite NPs and adjectives are differently distributed across the languages. The more frequent adjective realization in English can be related to the higher frequency of attributive processes shown in Table 7.

---

[10] Material verbs denote processes of doing and causing (Halliday 1994: 109); attributive processes ascribe an attribute to the subject referent and are a subtype of relational processes (*ibid.*: 173).

Table 8. Form of complement in transitive and copular clause patterns

Complement	English			Norwegian		
	N	%		N	%	
Indef NP	107	28.2		96	35.4	
Def NP	79	20.8	124	52	19.2	82
Pronoun	45	11.8	32.6%	30	11.1	30.3%
Adjective	88	23.2		41	15.1	
Clause	41	10.8		41	15.1	
PP + adverb	20	5.3		11	4.1	
Total	380	100.1		271	100	

The higher frequency of indefinite complements in Norwegian is an interesting finding. But even for English, it appears that the proportion of indefinite NP complements in the present material is higher than in Biber *et al.*'s figures for fiction (1999: 269), which only includes object NPs with definite and indefinite articles. In terms of Functional Sentence Perspective, with communicative dynamism being a relative concept, a clause with indefinite clause complementation after an indefinite subject is less at odds with the basic distribution of CD than one with an indefinite subject and a definite complement (Firbas 1992: 8). It appears that an indefinite complement makes the indefinite subject more palatable since the information structure of the clause will appear as all new, as in (13). By contrast, the information in (14) proceeds from new to given: the door is inferable from previous mention of a building, while the brunette is mentioned for the first time.

(13)   Bønder satte opp *uthus og innhegninger for hester og kveg* og lot en plog lage furer i jorden. (SH1)
Farmers put up *barns and corrals for horses and cattle* and ploughed long furrows in the earth. (SH1T)

(14)   A fresh-faced brunette woman in her thirties, wearing a flowery apron, opened *the back door*. (MM1)
En dame i trettiårene med brunt hår og et spill levende ansikt åpnet *døren*. (MM1T)
"A woman in the thirties with brown hair and a most lively face opened the door"

The pattern of indefinite subject + indefinite complement also occurs in descriptive passages presenting a series of observations, as in (15), which occurs in a series of independent observations. In contrast to the brunette introduced in (14), the freemasons in (15) are not maintained as a topic of the ensuing discourse.

(15)   A party of freemasons scrutinised a globe. (BC1)
En gruppe frimurere studerte en globus. (BC1T)

Finally, the pattern of new information in both subject and complement position includes a great number of generic sentences such as the one in (16). As noted above, generic noun phrases have a special status in information structure since they do not introduce actual discourse referents (Chafe 1994: 103). As example (16) illustrates, such sentences can be definition-like. This particular example also demonstrates the almost non-referential nature of generic reference, since the word *menasjeri/menagerie* also occurs in the previous sentence of the text, where it is used by an adult, while the current sentence explains the word for the benefit of a child.

(16) Et menasjeri var en samling av forskjellige dyr, ... (JG1)
      A menagerie was a collection of animals. (JG1T)

The second feature that was noted was the absence/presence of place adverbials. This adverbial type was singled out due to the close association between indefinite subjects and bare presentatives (Ebeling 2000) and between location and the presentation scale (Firbas 1992; Dušková 2015). Figure 1 shows the percentage of intransitive, monotransitive and copular clause patterns that also include a place adverbial. There were no place adverbials in ditransitive clauses and only one in a Norwegian complex transitive clause.
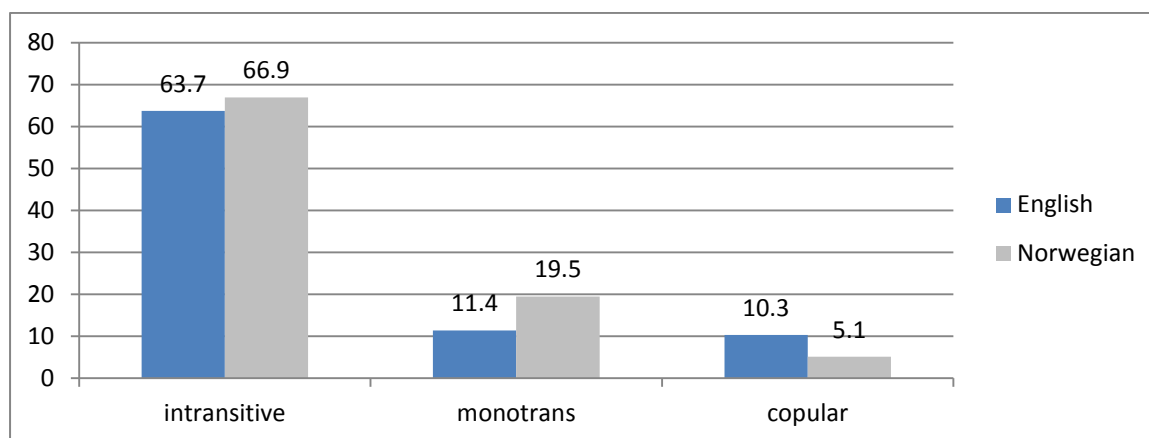

Figure 1. Place adverbials in clauses with indefinite subject (percentages)

Intransitive clauses are not only the most frequent environment for indefinite subjects but also clearly those most likely to occur with a place adverbial. However, it appears that intransitive clauses are on the whole most likely to occur with (all kinds of) adverbials: according to Oostdijk & de Haan (1993: 59) about 75% of intransitive non-embedded clauses contain one or more adverbials, as against about 50% of intensive clauses and 58% of monotransitive clauses. Considering that not all adverbials are place adverbials, the percentage of intransitive clauses containing a place adverbial shown in Figure 1 is still strikingly high.[11]

The co-occurrence of indefinite subject and place adverbial is frequent enough to constitute a pattern that tends to have some kind of presentative function, irrespective of process type. Such bare presentatives (Ebeling 2000: 157) typically involve intransitive verbs and subjects with specific reference. The pragmatic function is similar to that of full presentatives with *there/det*. Ebeling found bare presentatives to be more frequent in Norwegian than in English, a finding which is corroborated by the present study.[12] It may be noted that bare presentatives, as discussed by Ebeling (2000), may have an initial adjunct and a clause-final subject. This word order pattern has not been included in the present study; see Section 4. A typical example of the construction is given in (17), which contains a verb of existence/appearance.

---

[11] According to Hasselgård (2010: 34) and an unpublished study of Norwegian by the same author, close to 40% of adjunct adverbials are spatial in both languages.
[12] In Norwegian, bare presentatives account for 29.6% of the sentences with indefinite subjects (162 out of 548), and in English for 22% (161 out of 732).

(17) *En skikkelse* kommer til syne rundt hjørnet,… (LSC1)
      A figure comes into sight around the corner,…

In FSP (see e.g. Firbas 1992, Dušková 1999), sentences such as (17) represent the presentation scale. The subject has the dynamic role of Phenomenon with an Ex/App verb (and optionally an adverbial). Most such sentences could have been reformulated as full presentatives with *det/there* as an anticipatory subject, but as Ebeling (2000: 141) shows, English *there*-clauses are more restrictive than Norwegian *det*-clauses as regards the type of verb that occurs in them. The Norwegian original of (18) seems to have some sort of presentative function in spite of the fact that it does not contain a verb of existence/appearance. This type of passive construction is also noted by Dušková (1999: 255), who classifies it as presentational because the subject is taken to be rhematic. Example (19) is also presentational: although the verb denotes a material process, a full presentative would have been (marginally) acceptable in Norwegian (*Det plystret en fugl i hagen*), but not in English (**There chirruped a bird in the garden*). This suggests that the constructional meaning of indefinite subject + V + locative adverbial is more clearly presentative − and more readily paraphraseable by a full presentative − in Norwegian than in English.

(18) *En flokk griser* jages mot meg av folk med kjepper og stokker, … (SL1)
      "A herd (of) pigs is chased towards me by people with sticks and canes"
      A herd of pigs rushes towards me driven by men with whips and sticks, …

(19) *A bird* chirruped in the garden. (MM1)
      En fugl plystret i hagen. (MM1T)

5.5   Summary of the contrastive analysis

The contrastive analysis of indefinite subject NPs in comparable original data has shown that the clearest cross-linguistic difference concerns frequency. While indefinite subjects are more frequent in English, their lexicogrammatical features are relatively similar in both languages. However, the zero article is more common in English than in Norwegian, and premodified NPs are proportionally more common in Norwegian. The lexicogrammatical features of the verb phrases in clauses with indefinite subjects are also similar. Most VPs are intransitive with a material process verb in the active voice. However, verbs of existence/appearance are more common in Norwegian while attributive verbs are more common in English. The lexicogrammatical features of the clause pattern differ somewhat between the languages. Place adverbials are frequent in both, but NP complements are more likely to be indefinite in Norwegian than in English. It was found that generic reference and presentative function provide favourable conditions for indefinite subject NPs in both languages.

6.   The translation of indefinite subject NPs

This section discusses the translation of sentences with indefinite subjects from English into Norwegian and vice versa. Because indefinite subjects are comparatively rare in both languages, it is hypothesized that translators will make a number of changes to avoid the markedness of a rare construction and atypical information structure. And because indefinite subjects are less frequent in Norwegian than in English, Norwegian translators are expected to make more changes than English translators to avoid indefinite subject NPs in clause-initial position, due to the translation principle of normalization (Baker 1996).

6.1 The classification of translation correspondences

The starting point of the classification scheme is that of Johansson (2007: 23) into *congruent, divergent* and *zero* correspondences. Congruent correspondences are those that have the same types of linguistic form in both languages, and divergent correspondences are those where the realization types differ. With zero correspondence, either the source or the translation lacks the item under investigation. In the current material, zero correspondences were only found when whole s-units were omitted in the translation.

A more fine-grained classification system than Johansson's (2007) was needed for the present analysis, and so the categories of congruent and divergent were subdivided. Table 9 shows the correspondence types according to degrees of congruence. Congruent correspondences are those in which the subject NPs of the original and the translation correspond practically word for word with each other. Semi-congruent correspondences have the same phrase type, but with different internal structures. Divergent correspondences may involve correspondences between an indefinite NP and a definite one or a pronoun, or more substantial reorganizations of the clause, including the use of a full presentative (with the anticipatory subject *there/det*), a rearrangement of the clause content that involves a change of subject NP, and even more radical changes to clause/sentence structure, subsumed under the label of "rephrasing".

Table 9. Congruent and divergent correspondences.

		Original	**Translation**	**Gloss of Norwegian example**
Congruent: indefinite NP		*En tyv* er ikke voldsom, men stillferdig. (KF1)	*A thief* is not violent but quiet.	A thief is…
Semi-congruent: restructured indefinite NP		*Faces showing shock* looked out at our passing. (DF1)	*Tydelig sjokkerte ansikter* kikket ut på oss da vi passerte.	Clearly shocked faces…
Divergent correspondences	Definite NP	*Animals* don't eat me, and I don't eat them. (PDJ3)	*Dyrene* spiser ikke meg, og jeg spiser ikke dem.	The animals eat not me…
	Pronoun	Men *et slikt bytte* er like umulig: (JG1)	But *this* is equally unacceptable.	But a such reward is equally impossible
	Full presentative	*Puddles* had formed everywhere, …(SG1)	*Det var* sølepytter overalt, …	There were puddles everywhere
	Subject change	*En så viktig beslutning* må ikke avledes. (FC1)	*One* mustn't be side-tracked from *an important decision like that*.	A so important decision must not be sidetracked
	Rephrasing	*A raised voice* is remarked on, … (RR1)	Hvis noen hever stemmen, blir det bemerket, …	If anyone raises the voice, is it remarked

6.2 Changes made in translation

Table 10 shows the extent to which translations are congruent or divergent, and the types of changes made in case of divergence. The two directions of translation differ markedly from each other: fewer changes are made in translations from Norwegian into English than in the other translation direction. That is, congruent and semi-congruent correspondences are much more common in Eng→Nor than in Nor→Eng.

Table 10. Translation correspondences of indefinite subject NPs.

		Norwegian→English		English→ Norwegian	
		N	%	N	%
Congruent		407	76.4	402	55.3
Semi-congruent		55	10.3	73	10.0
Divergent	Definite NP	17	3.2	104	14.3
	Pronoun	8	1.5	14	1.9
	Full presentative	2	0.4	33	4.5
	Subject change	31	5.8	72	9.9
	Rephrased	13	2.4	29	4.0
	Total	533	100.1	727	100

The divergence that involves turning an indefinite subject into a definite one is clearly more frequent in Eng→Nor. The example of this type of change in Table 9 has generic reference. It appears that Norwegian definite plurals can have generic reference more readily than English ones, according to Johansson and Lysvåg (1987: 43), who continue: "If this tendency is carried over into English the result would be a noun phrase with more or less clear specific reference" (*ibid.*). However, the change from indefinite to definite form is also found when the subject has specific reference, as shown in (20).

(20) *Walls* had been pulled down to make this a room that accommodated nearly all the ground floor. (DL1)
Veggene var revet ned for å lage dette til et rom som tok opp nesten hele første etasje. (DL1T)
"The walls were torn down…"

The large room described in (20) has in fact been mentioned in a preceding sentence; hence the *walls* are inferable. This fact will have justified the definite form in the translation. But although many examples are of this type, others are harder to explain, such as (21), in which no experts have been mentioned or can be inferred from the context. It is tempting to assume that the change from indefinite to definite subject has been made chiefly to create an apparently smoother information structure.

(21) *Experts* restored the canvas by repairing the boot. (JH1)
Ekspertene restaurerte lerretet ved å reparere støvelen. (JH1T)
"The experts…"

The divergent correspondence that involves turning a bare presentative into a full presentative is less frequent than the indefinite-to-definite change, but still noticeably more common in translations from English into Norwegian. The example given in Table 9 is typical in that the verb is *be*, corresponding to the Norwegian *være*. However, the same type of change is found with other verbs, as shown in (22).

(22) *A queue* had formed in the area newly designated for waiting in... (RR1)
Det hadde dannet seg kø i det nye feltet som var avsatt for ventende kunder… (RR1T)
"There had formed itself queue…"

As noted above, the Norwegian presentative construction is more flexible than the English one as regards the range of lexical verbs that occur in it (Ebeling 2000: 260–261). The translator of (22) thus exploits a possibility that exists in Norwegian, but not in English, of relegating the subject to non-initial position while retaining a full, lexical verb.

Subject change is almost twice as common in translations from English into Norwegian as vice versa. Examples (23)–(25) show three recurrent types of change within this category: (i) another participant from the clause is promoted to subject; (ii) the impersonal pronoun *man* ('one') is introduced; (iii) a pronominal subject is supplied from the context.

(23)　A loud sharp barking suddenly disturbed *the silence*. (MM1)
　　　Stillheten ble plutselig forstyrret av skarp bjeffing. (MM1T)
　　　"The silence was suddenly disturbed by sharp barking"

(24)　A lot of inquiries can be done by phone… (SG1)
　　　Man kan gjøre en god del undersøkelser per telefon... (SG1T)
　　　"One can do a good deal (of) inquiries by phone, …"

(25)　Three men from town and another farmer named Hawkins helped. (JSM1)
　　　De fikk hjelp av tre karer fra byen og en gårdbruker som het Hawkins. (JSM1T)
　　　"They got help from three men from town…"

Subject change is the most frequent type of divergence in translations from Norwegian into English, and the examples are mainly of the same type as those in the other direction of translation. Thus, in (26) a pronominal subject is supplied from the context, and in (27) another participant is promoted to subject in the translation.

(26)　En arm lå over hoften *hans*. (KAL1)
　　　"An arm lay across the hip his"
　　　He felt an arm resting on his hip, ... (KAL1T)

(27)　En ny tanke slo ned i *henne*… (EG1)
　　　"A new thought struck down in her"
　　　She was suddenly struck by another thought… (EG1T)

6.3　Syntactic restructuring in translation

Beyond the changes in translation mentioned in the previous section, we find instances of syntactic restructuring in translation which removes the indefinite NP from sentence-initial position. That is, some translation correspondences involve a simple reordering of constituents. An example is given in (28), where the adjunct is moved from end to initial position in the translation. This type of change is not very frequent, but found more often in translation from English into Norwegian than in the opposite direction.[13] Interestingly, syntactic reordering in Nor→Eng translation seems to be related to another aspect of adverbial placement, i.e. the lower tolerance for long adverbials in medial position in English than in Norwegian (Johansson and Lysvåg 1987: 264; Hasselgård 2010: 107). Example (29) illustrates the movement of a long medial adverbial to initial position, thus delaying the indefinite subject.

---

[13] There are 11 instances in Nor→Eng translations and 23 in Eng→Nor translations, accounting for 2.1% and 3.2% of the material, respectively.

(28) *A bowl of hot, buttered, boiled potatoes* stood in the middle of the table. (MM1)
Midt på bordet sto *et fat med kokte poteter*. (MM1T)
"In the middle of the table stood a bowl of boiled potatoes"

(29) *En sosial handling* var *i lys av den oppståtte mistanken* blitt asosial, forbrytersk. (KA1)
"A social action had in light of the arisen suspicion become anti-social, criminal"
In light of the suspicion that had arisen, a social action had become anti-social, criminal. (KA1T)

Furthermore, a more thorough restructuring may take place, typically involving subject change (including the use of presentative *there/det*) as discussed in Section 6.2. These types of changes are also much more frequent in English-Norwegian translations. In the Norwegian translations we also find constructions with *det* that are not clearly presentative, but for example impersonal passives or clefts, as illustrated by (30) and (31).

(30) *A house in the city* could be bought for that much. (JH1)
For de pengene kunne det kjøpes et hus midt i byen. (JH1T)
"For that money could there be-bought a house in the middle of the city"

(31) *An intruder* had done this. (RR1)
Det var en tyv som hadde gjort dette. (RR1T)
"It was a thief who had done this"

Syntactic restructuring involving a change of voice is not very frequent in either direction of translation: it is found in only about 5% of the clauses. However, in translations into English, changes from active to passive and from passive to active are about equally common, while in translations into Norwegian, passive to active changes are twice as common as active to passive. This lends some support to Johansson's observation that "English appears to have a greater preference for passive perspectives" (2004: 49), although it should be recalled that the proportions of active and passive clauses were similar in both English and Norwegian originals (Section 5.3).

6.4   Factors affecting congruence in translations

As shown in Table 10, the large majority of indefinite subjects remain unchanged in translation, particularly going from Norwegian into English. Table 11 reports findings from a cross-tabulation of full congruence with the lexicogrammatical features of the subject NP, the verb phrase and the clause pattern. The aim of this exercise is to discover contexts that are particularly favourable to the preservation or change of an indefinite subject NP. A lexicogrammatical feature has been considered to promote either congruence or divergence if the percentage of congruent correspondences differs by at least 5 percentage points from the mean for each direction of translation. The features not mentioned in the table do not seem to pull the degree of congruence up or down from the mean percentage.

Table 11. Full congruence and lexicogrammatical features of the subject/context.

	English→Norwegian	Norwegian→English
Mean rate of congruence	54.9%	74.3%
NP complexity	Complex NP promotes congruence	Simple NP promotes congruence
NP specificity	Generic NP promotes congruence	
NP semantics	Animate and human reference promotes congruence.	Non-human animate promotes congruence.
	Abstract reference reduces congruence.	
VP transitivity	Transitive verb promotes congruence; intransitive reduces it.	Copular pattern reduces congruence.
Complement form	Clausal and indefinite NP complements promote congruence.	
		Adjectival complement reduces it.
Voice	Passive voice reduces congruence	

It is interesting that complex subject NPs should increase the chance of a congruent translation from English into Norwegian. Possibly, the presence of modifiers makes the NP referent more readily identifiable and therefore more acceptable as a subject. In the case of (32), the next sentence also makes anaphoric reference to the clause-final *pool*.

(32)   A statue that probably represented the pursuit of Daphne by Apollo was reflected in the dark waters of a shallow pool. (RR1)
En statue som så ut til å forestille Apollons forfølgelse av Dafne, speilte seg i det mørke vannet i et grunt basseng. (RR1T)
"A statue that seemed to represent Apollo's pursuit of Daphne…"

Abstract reference of the indefinite subject NP decreases congruence in both directions of translation. This may have to do with the hierarchy of subject selection proposed e.g. by Givón (1993: 93) by which the preferred subject roles are agent > dative > patient > others. That is, abstract referents may be less likely than (human) animate and concrete ones to have the role of agent. The hierarchy of subject selection may also impact on the relatively low degree of congruence with passive clauses noted in Table 11.

Subjects with a transitive verb followed by a complement promote congruence in translation from English to Norwegian, particularly if the complement is realized by a clause or an indefinite NP. Such sentences will adhere to either the information principle (new information last) or the end weight principle (long constituents last) (Biber *et al*. 1999: 896, 898), which is likely to make the indefinite subject more acceptable, as noted in Section 5.4.

6.5   Translation correspondences: main findings and further research

Translation correspondences of indefinite subjects show that although the subject NP is retained in congruent form in the majority of cases, there is a marked difference between the two directions of translation: more changes are made in translations from English into Norwegian than the other way round. This can be related to the lower overall frequency of indefinite subjects in Norwegian original texts, which in turn is taken to reflect a lower tolerance of such subjects. The most frequent change made in translation from English to Norwegian is replacing the indefinite NP with a definite one. This type of change occurs in the other translation direction too, but much less commonly. The most frequent change in Nor→Eng, and the second most frequent in Eng→Nor, is a restructuring which involves subject change. It may be noted that the changes made tend to bring the translated sentence

into better agreement with the information principle. Certain contexts that were found to be welcoming to indefinite subjects in the contrastive study, particularly generic reference, tend to encourage congruence in translation. Similarly, the presence of a post-verbal complement, particularly one that is either indefinite or long, is conducive to a congruent translation.

This study has only considered sentences with indefinite subjects in original texts and their translations. A future complementary study might look into indefinite subjects in translated texts and identify their sources. For example, translations may contain some sentence-initial subjects that do not appear in the original, as in (33), where the Norwegian original has a sentence-initial direct object (and the subject in post-verbal position due to the V2 constraint).

(33)   Kaméen tar sikkert en gullsmed med glede. (KF1)
         "The cameo takes surely a jeweler with pleasure"
         A jeweler will be glad to take the cameo. (KF1T)

The use of indefinite subjects is likely to be sensitive to genre. Thus, another avenue of further research would be to conduct a similar investigation using for example academic texts. Unfortunately, the ENPC does not contain sufficient material in any non-fiction genre to facilitate such an investigation.

7.   Concluding remarks

The present study set out to investigate sentence-initial indefinite NP subjects in English and Norwegian. Such subjects appear anomalous in the light of word order principles such as the basic distribution of communicative dynamism (Firbas 1992; Dušková 2015) and the light subject constraint (Chafe 1994) because of their association with new information. Most likely for this reason, indefinite subjects have been observed to be rare (Prince 1992, Biber *et al*. 1999). The indefinite subjects were first analysed cross-linguistically on the basis of original texts with respect to their frequencies, their lexicogrammatical features and the contexts in which they occur, particularly regarding features of the verb phrase and the presence of other constituents in the clause. The main hypothesis, based on previous contrastive studies of related issues, was that English would be more tolerant of sentence-initial indefinite subjects than Norwegian. This turned out to be the case. However, the lexicogrammatical features of both subject NPs and their verbs were relatively similar across the languages. This indicates that the differences between the languages are not systemic. That is, the language difference is stylistic rather than structural, although it can be argued, in the words of Johansson (2004: 49), that there is "no clear borderline between structural differences and stylistic preferences."

The study of translations gave additional evidence of the lower tolerance of indefinite subjects in Norwegian in that such subjects are changed more often in translation from English to Norwegian than vice versa. However, the high degree of congruence in both directions give further support to the idea that the differences observed are due to preferences rather than grammaticality. There are probably few cases – given the appropriate lexical resources in the target language – in which a congruent translation would be ungrammatical. When translators make structural changes in spite of the availability of a congruent correspondence, this must reflect language-specific preferences as to which syntactic patterns are perceived as natural and idiomatic.

The results of both parts of the investigation supported the hypotheses but also brought out more detailed information about the use of indefinite subjects. While Johansson (2004) found that the form of subjects in general is preserved in translation from English into Norwegian in about 90% of the cases, the present study showed that sentence-initial indefinite subjects are more prone to change, and moreover, that changes are made much more often in translation from English into Norwegian than from Norwegian into English (Table 10). The difference between translation directions is consistent with Ebeling's (2000) findings regarding the translation of bare presentatives: he argues that since English tolerates indefinite NPs in subject position in bare presentatives to a greater extent than Norwegian, the Norwegian S-V (+Locative) presentatives are expected to be translated by similar English patterns, which is also the case (*ibid.*: 191). Besides the bare presentatives, however, which are an important context for indefinite subjects (*ibid.* and Dušková 2015), generic sentences were found to be favourable to indefinite subjects in both languages. Similarly to sentences that also had indefinite post-verbal complements, these were considered less at odds with the information principle.

The study shows that the light subject constraint (Chafe 1994) is even more apparent in Norwegian than in English. In a cross-linguistic perspective it can be argued that the stricter application of the constraint in Norwegian works well for English in the sense that the Norwegian preferences will generally produce acceptable English sentences. In contrast, the greater tolerance of indefinite subjects in English works less well for Norwegian. Hence translators from English into Norwegian feel a need to change a number of sentences with indefinite subjects, whereas translators from Norwegian into English may find that such sentences in the source text can be rendered congruently because they already lie well within what is considered natural English usage.

References

Baker, M. 1996. Corpus-based translation studies: the challenges that lie ahead. In *Terminology, LSP, and Translation: Studies in Language Engineering in Honour of Juan C. Sager,* H. Somers (ed.), 175–186. Amsterdam: John Benjamins.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Chafe, W. 1986. Writing in the perspective of speaking. In *Studying Writing: Linguistic Approaches,* C.R. Cooper and S. Greenbaum (eds), 12–39. Beverly Hills: Sage Publications.

Chafe, W. 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.

Ebeling, J. 2000. *Presentative Constructions in English and Norwegian. A Corpus-based Contrastive Study*. Doctoral thesis. Oslo: Faculty of Arts, University of Oslo.

de Hoop, H and Krämer, I. 2006. Children's optimal interpretations of indefinite subjects and objects. *Language Acquisition*, 13:2, 103–123.

Dušková, L. 1999. Basic distribution of Communicative Dynamism vs. Nonlinear indication of Functional Sentence Perspective. In *Prague Linguistic Circle Papers*, Vol 3, E. Hajičová, O. Leška, P. Sgall and Z. Skoumalová (eds), 249–261. Amsterdam: John Benjamins.

Dušková, L. 2015. *From Syntax to Text. The Janus Face of Functional Sentence Perspective*. Prague: Karolinum Press.

Faarlund, J.T. 1988. A typology of subjects. In *Studies in Syntactic Typology,* M. Hammond, E.A. Moravcsik and J. Wirth (eds), 193–208. Amsterdam: John Benjamins.

Faarlund, J.T., Lie, S. and Vannebo, K.I. 1997. *Norsk Referansegrammatikk.* Oslo: Universitetsforlaget.

Firbas, J. 1992. *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press.

Givón, T. 1993. *English Grammar: A Function-based Introduction.* Amsterdam/Philadelphia: John Benjamins.

Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*. London: Arnold.

Hasselgård, H. 2004. Thematic choice in English and Norwegian. *Functions of Language* 11:2, 187–212.

Hasselgård, H. 2005. Theme in Norwegian. In *Semiotics from the North: Nordic Approaches to Systemic Functional Linguistics*, K.L. Berge, and E. Maagerø (eds), 35–48. Oslo: Novus.

Hasselgård, H. 2010. *Adjunct Adverbials in English*. Cambridge: Cambridge University Press.

Hasselgård, H., Lysvåg, P. and Johansson, S. 2012. *English Grammar: Theory and Use.* Second edition. Oslo: Universitetsforlaget.

Hawkins, J.A. 1978. *Definiteness and Indefiniteness. A Study in Reference and Grammaticality Prediction*. London: Croom Helm.

Johansson, S. 2004. Why change the subject? On changes in subject selection in translation from English into Norwegian. *Target* 16(1), 29–52.

Johansson, S. 2007. *Seeing through Multilingual Corpora. On the Use of Corpora in Contrastive Studies*. Amsterdam/Philadelphia: John Benjamins.

Johansson, S., Ebeling, J. and Oksefjell, S. 1999/2002. English-Norwegian Parallel Corpus: Manual. http://www.hf.uio.no/ilos/english/services/omc/enpc/ENPCmanual.pdf.

Johansson, S. and Lysvåg, P. 1987. *Understanding English Grammar. Part II: A Closer View*. Oslo: Universitetsforlaget.

Lyons, C. 1999. *Definiteness*. Cambridge: Cambridge University Press.

Matthiessen, C.M.I.M. 1995. *Lexicogrammatical Cartography: English Systems*. Tokyo: International Language Science Publishers.

Matthiessen, C.M.I.M. 1999. The system of transitivity: An exploratory study of text-based profiles. *Functions of Language* 6(1): 1–51.

Oostdijk, N. and de Haan, P. 1993. Clause patterns in Modern British English: A corpus-based (quantitative) study. *ICAME Journal* 18: 41–79.

Prince, E.F. 1981. Toward a taxonomy of given–new information. In *Radical Pragmatics,* P. Cole (ed.), 223–255. New York: Academic Press.

Prince, E.F. 1992. The ZPG letter: Subjects, definiteness, and information-status. In *Discourse Description: Diverse Linguistic Analyses of a Fund-Raising Text,* W.C. Mann, and S.A. Thompson (eds.), 295–326. Amsterdam: John Benjamins.

Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Rørvik, S. 2013. *Texture in Learner Language*. PhD thesis, University of Oslo.

Author's address

Hilde Hasselgård
Department of Literature, Area Studies and European Languages
University of Oslo
PO box 1003
0315 Oslo
Norway

hilde.hasselgard@ilos.uio.no