

The way of the world: the colligational framework “the N1 of the N2” and its Norwegian correspondences

Hilde Hasselgård, University of Oslo

Abstract

The English-Norwegian Parallel Corpus is used to study the colligational framework “the N1 of the N2” (e.g. *the way of the world*). The aim is to discover recurrent patterns in the lexical and semantic make-up of such sequences in English and their correspondences in Norwegian. The correspondences are described in terms of both structure and meaning. In spite of similarities in the structural potential of English and Norwegian noun phrases, the majority “the N1 of the N2” sequences have divergent correspondences in Norwegian. This points to differences between the languages in their preferred lexicogrammatical patterns. The study confirms the feasibility of carrying out cross-linguistic studies on the basis of patterns of function words.

1. Introduction

This paper explores the colligational framework “the N1 of the N2” and its Norwegian correspondences.¹ In so doing, it represents an attempt at extending the domain of contrastive studies to patterns based on function words. The term “colligational framework” is modelled on Renouf & Sinclair’s “collocational framework”, defined as a discontinuous sequence of (grammatical) words (Renouf & Sinclair 1991: 128). Since it is not grammatically complete (or “self-standing”), its “wellformedness is dependent on what intervenes” (ibid.). Unlike Renouf & Sinclair’s study, the present one specifies word class membership for the two empty slots, hence the modified term. Renouf & Sinclair demonstrate that such frameworks are highly selective of their collocates. For example, focusing on the intermediate word in the framework “a _ of”, they find that the quantifying nouns *couple*, *series*, *pair* and *lot* occur in this framework more frequently than they occur outside it (ibid.: 131).

Hunston (2008) argues that “small words” play an important role in the identification of grammar patterns, which in turn can form semantic sequences, i.e. “series of meaning elements that can be demonstrated to

¹ I am grateful to the reviewers of this paper for constructive criticism, and in particular for suggesting the term “colligational framework”.

Hasselgård, Hilde. 2016. “*The way of the world: the colligational framework “the N1 of the N2” and its Norwegian correspondences.*” *Nordic Journal of English Studies* 15(3):55-79.

occur regularly” (2008: 271). That is, the core of semantic sequences need not be lexical words, since grammatical words have also been shown to have collocates; see e.g. Groom (2007), who studies the phraseological profiles of a number of grammatical words. Groom finds, for instance, that in a random sample of 100 instances of *of*, a number of semantic sequences can be recognized, such as PROPERTY + *of* + PHENOMENON, illustrated by *the essential values of academic life* (2007: 84). As also noted by Hunston, prepositions can “serve to classify semantically the lexical words with which they occur” (2008: 202). Such semantic classifications are of great interest in a contrastive study.

In the British National Corpus (BNC) *the end of the* and *the rest of the* are both among the ten most frequent 4-grams.² This suggests that “the N1 of the N2” is a common and productive colligational framework worth investigating contrastively. In addition, complex noun phrases are intriguing from an English-Norwegian contrastive perspective: as Elsness (2014: 91) observes, “there are some notable differences in the structure of the noun phrase between the two languages”, particularly in the use of modifiers. In spite of this, contrastive studies of complex noun phrases in English and Norwegian are scant, Elsness (2014), focusing on clausal modifiers, being an exception.

This study explores the colligational framework “the N1 of the N2” with the aim of discovering recurrent patterns as regards the selection of lexical items in the frame as well as its Norwegian correspondences. The investigation of English will focus on the patterns that emerge with regard to the types of nouns in the open slots and the relation between the N1 and N2. The Norwegian correspondences of the pattern will be studied in order to discover the extent to which Norwegian uses similar or different patterns, and whether different semantic types of nouns, or combinations of nouns in the “the N1 of the N2” framework, trigger different structural types in Norwegian. Because of the exploratory nature of the study, and in order to limit the material, the investigation is limited to definite noun phrases, thus ignoring e.g. the related framework “a(n) _ of” studied by Renouf & Sinclair (1991). It is hoped not only that this limited investigation will point to systematic similarities and differences between English and Norwegian complex noun phrases, but

² According to a search using “Phrases in English”: <http://phrasesinenglish.org>, accessed 29 July 2015.

also that such an investigation will demonstrate the usefulness of starting an investigation from grammatical words and exploring colligational frameworks in contrast.³

2. Complex noun phrases in English and Norwegian

English and Norwegian noun phrases have much the same syntactic structure, with one major exception: the definite form of Norwegian nouns is marked morphologically by a suffix. The definite suffix usually replaces the definite article, but if the noun is premodified, the noun phrase typically (but not obligatorily) has double definiteness, with both article and definite suffix. Examples are shown in (1), in which the definite suffixes have been underlined.⁴

- (1) On *the way* to *the hotel* where he was staying, *the emaciated man* walked along the harbour of *the little town*. (FC1T)
 På *veien* til *hotellet*, der han bodde, gikk *den radmagre mannen* langs *havnen* i *den lille byen*. (FC1)
 Lit: “On way_{def} to hotel_{def}, where he lived, walked the emaciated man_{def} along harbour_{def} in the little town_{def}.”

As the example also shows, nouns in both languages can be premodified by adjectives and postmodified by prepositional phrases. Nominal premodification is not common in Norwegian unless the premodifying noun has genitive form; noun + noun combinations in English may correspond to compounds in Norwegian; see (2).

- (2) It was *Harry’s idea* that they should give a *dinner party* ... (FW1)
 Det var *Harrys idé* at de skulle ha *middagsselskap* ... (FW1T)
 Lit: “It was Harry’s idea that they should have dinnerparty”

³ Miccoli (2010) studied phrasal verbs in English and Norwegian on the basis of lexically specified particles; this is the only previous study to my knowledge that explores collocations on the basis of function words. Brems (2015) studied quantifying binomials with *of*, but with a lexically specified N1.

⁴ All examples are from the English-Norwegian Parallel Corpus. The English version consistently appears first. If it is a translation, its identification tag ends in ‘T’. The Norwegian version is immediately followed by a literal translation marked as “Lit.”. The most relevant parts of the examples have been italicised.

It may be noted that genitive expressions are also similar across the languages: like English, Norwegian has an *s*-genitive, illustrated in (2), as well as an equivalent of the *of*-genitive, using the preposition *til* (e.g. *the name of the journalist – navnet til journalisten*). For further details of complex noun phrase structure in English and Norwegian, see e.g. Biber et al (1999: 574 ff) and Faarlund et al. (1997: 233 ff). The structural similarities are great enough for the “the N of the N” construction to have congruent correspondences in Norwegian, allowing for the replacement of the definite article by the definite suffix, as seen in (3).

- (3) I live in that caravan *on the edge of the cliff*. (PDJ3)
 Jeg bor i campingvognen der borte på *kanten av skrenten*. (PDJ3T)
 Lit: “I live in caravan_{def} there over on edge_{def} of cliff_{def}.”

However, the similarities in syntactic structure, and thereby in constructional potential, are “no guarantee that there is identity of use” (Johansson 2012: 47). Rather, as is often found in contrastive studies, it is likely that the two languages differ in their “preferred ways of expressing similar meanings” (ibid.: 64).

3. The colligational framework “the N1 of the N2”

Sequences that instantiate the framework “the N1 of the N2” are complex noun phrases, according to most accounts consisting of a head noun (N1) plus a postmodifier (*of* N2). According to Biber et al. (1999: 606), prepositional phrases are the most frequent type of postmodifier of nouns, and of these, 60-65% are introduced by *of* (ibid.: 635). Furthermore, lexical bundles involving *of*-phrases are highly frequent in academic prose: a list of four-word bundles of this type contains numerous instances such as *the end of the*, *the beginning of the*, *the base of the* (ibid.: 1014 f). Some such bundles are also noted for conversation, particularly in temporal and spatial references such as *the end/middle of the week/road* (ibid.: 1012).

However, *of* is not a “typical” preposition; in particular it generally does not introduce PPs with an adjunct function (Sinclair 1991; Owen 2007). According to Sinclair, “it may ultimately be considered distracting to regard *of* as a preposition at all” (1991: 83). On the other hand, *of* is highly multifunctional in terms of the meaning relationships it

expresses between nouns preceding and following it. In addition, an “N of N” sequence is not necessarily one of head noun followed by a postmodifier; according to Sinclair, the second noun seems to be most salient. Hence, *of* can introduce a second noun as a potential headword in phrases such as *this kind of problem*, *the bottle of port* (1991: 85), an issue also discussed by Keizer (2007).

The meaning relations expressed by *of* in the “N of N” sequence can to some extent be linked to the meaning of the N1 according to Sinclair (1991: 87 ff). The N1 can be a “focus noun”, specifying some part of the N2 (which is regarded as the head noun), e.g. *the top of the pillar*. The focus can also be on a “specialized part”, e.g. *the first week of the war*, *the horns of the bull*, or on “a component, aspect, or attribute”, e.g. *a gasp of shock*, *the study of geography*. Alternatively, the N1 can be a “support noun”, often reduced in meaning and “offering some kind of support to N2” (ibid: 89), e.g. *the notion of machine intelligence*, *an object of embarrassment*.

Keizer (2007), whose primary interest is in the structural properties of the noun phrase, takes a different perspective on the semantic relations expressed by *of*. “Possession” is seen as a cover term that includes ownership (of material belongings as well as more abstract products), kinship and body parts. Possession is also extended to non-human possessors, e.g. *the mosaics of Venice*, *the problems of the world* (2007: 63). A particular group of nouns in the N1 position are called “relational” (ibid.: 64): relational nouns include kinship terms (*the son of Aron*), body parts, and “nouns which denote parts of (physical or abstract) features of entities”, such as *size*, *middle*, *feature*, *nature* (ibid.); thus nouns that might be seen as either focus nouns or support nouns in Sinclair’s (1991) framework. Other functions of *of*-phrases recognized by Keizer (2007: 65, 73) are partitive (*the first of a few questions*) and appositional (*the job of foreign minister*); see also Quirk et al. (1985: 1284).

For the present analysis the semantic relations were classified descriptively in terms of an analysis of the N1 into the following categories (which might also be seen as semantic sequences; cf. Hunston 2007; Groom 2007):

- part of N2, e.g. *the corners of the tea-towel*
- locative, e.g. *the back of the bus*
- temporal, e.g. *the night of the ball*

- owned/caused by N2, e.g. *the noise of the TV*
- body part, e.g. *the legs of the men*
- feature of N2, e.g. *the names of the places, the colours of the rainbow*
- person, e.g. *the leader of the expedition*
- nominalization, e.g. *the outbreak of the Plague*
- support, e.g. *the idea of the echo*
- ordinal, e.g. *the last of the fishfingers*
- N1 contains N2, e.g. *the pictures of the frog*

Admittedly, there is some overlap between the categories. In particular, it can be argued that the categories “locative”, “temporal” and “body part” are superfluous because they are subtypes of the “N1 is part of N2” relation. However, for the purposes of this study, also taking the Norwegian correspondences into account, it makes sense to single out these subtypes. Several of the categories might be grouped under “possession” in Keizer’s (2007) wide sense, i.e. the top six on the list, and possibly also nominalization (2007: 62). The category of “support noun” has been borrowed from Sinclair (1991). His category of “focus noun”, however, has not been applied, as a more fine-grained system was needed for this study, e.g. one that can single out locative nouns. Support nouns would correspond to Keizer’s appositional relation, and ordinals to the partitive one. The list presented above seems to make sense in light of the correspondences that occur; see further section 5.

4. Material and method

The investigation is based on the fiction part of the English-Norwegian Parallel Corpus (ENPC). The ENPC is a bidirectional translation corpus containing original texts in both English and Norwegian with translations into the other language. See Johansson et al. (1999/2002) for further details of the corpus.

For the present study, the raw text files of English originals and translations were imported into WordSmith Tools, and the string *the * of the ** was searched for. My interest was in recurrent patterns, but since very few sequences recurred in identical form, sorting and selection were made according to the N1 in the sequence following the observation that there was more recurrence of N1s than of N2s. All single occurrences

were removed, as were those occurring in one text only, leaving 430 hits from English originals and 507 from translations. These were matched up with their Norwegian correspondences, stored in a database and annotated for the lexemes occurring as N1 and N2 and the semantic relation between them, correspondence type (congruent, divergent or zero; cf. Johansson 2007: 25), and structural type of the Norwegian correspondence.⁵

In this context correspondences were considered congruent (i.e. structurally equivalent) if they consisted of two definite nouns and an intervening preposition, as in (4). Divergent correspondences are all other structural types, e.g. compounds, as in (2) above, or word class change for either the N1 or the N2, as shown in (5), where the N1 plus the preceding preposition correspond to an adverb.

- (4) No identification on *the back of the photograph*. (NG1)
 Men det finnes ingen identifikasjon på *baksiden av fotografiet*.
 (NG1T)
 Lit: “But there exists no identification on *backside_{def} of photograph_{def}*.”
- (5) A yellowed newspaper cutting lay *on the top of the pile*. (GS1T)
Øverst i bunken lå et gulnet avisutklipp. (GS1)
 Lit: “*Uppermost in pile_{def} lay a yellowed newspapercutting*.”

There are also some cases of zero correspondence, which occurs if either the “N1 of the N2” sequence or the relevant s-unit has no counterpart in the Norwegian text. Cases where either the N1 or the N2 is missing from the Norwegian source or original, however, are analysed as divergent correspondences (of the structural types “N1/N2 missing”).

5. Corpus investigation

This section presents an analysis of the colligational framework “the N1 of the N2” and its Norwegian correspondences. Section 5.1 surveys the noun collocates of the framework before the correspondences are

⁵ Note that the term “correspondence” covers both sources and translations, c.f. Johansson (2007: 23).

explored with regard to congruence and divergence, structural types of Norwegian correspondences, meanings of the N1s, and individual patterns for the most frequent N1s. It is expected that recurrent nouns in the framework will form certain (types of) semantic sequences. Assuming that Norwegian is less fond of complex noun phrases than English is (cf. Elsness 2014), congruent correspondences will be more frequent in Norwegian translations than in Norwegian originals. Furthermore, if Sinclair is right about the greater prominence of the N2 in such sequences, N1 can be expected to be divergent more often than N2. It is likely that individual N1 lexemes or combinations of N1 and N2 as well as certain meaning relations may come with their preferred correspondence types. Preferred correspondence types will thus be related to both patterns and individual nouns.

5.1 Nouns occurring in the “the N1 of the N2” framework

The nouns occurring above 10 times in the N1 position are shown in Table 1. Singular and plural forms of the same noun have been grouped under the same lemma, as have the British and American spellings of *centre/center*.

There is a reasonable degree of overlap between the N1s occurring in this pattern in originals and translations, and those nouns that are shown to occur only in originals in Table 1 are to be found among the top 20 of the translations and vice versa. The majority of the words in both columns are locative and/or denote a part of whatever the N2 refers to. As will be detailed in Section 5.3, “part of N2” and “locative” are indeed by far the most frequent categories of N1.

Table 1. N1 collocates selected by the framework (min. 10 occurrences)

rank	English original		English translation	
	lemma	frequency	lemma	frequency
1	end	39	middle	52
2	rest	26	edge	43
3	back	22	rest	30
4	edge	22	end	23
5	side	18	back	16
6	centre	14	bottom	15
7	top	14	side	14
8	middle	13	top	14
9	bottom	12	sound	13

Only a handful of N2s occur above 10 times, and the majority of those refer to places and times, as shown in Table 2. The most frequent N2s are, in short, nouns that go together with the most frequent N1s, i.e. they denote things that have middles, edges, sides or ends. The impression from Tables 1 and 2 combined is that the framework “the N1 of the N2” is associated with locative, and to some extent temporal, expressions (cf. Biber et al. 1999: 1012).

Table 2. N2 collocates selected by the framework (min. 10 occurrences)

rank	English original		English translation	
	lemma	frequency	lemma	frequency
1	house	15	table	15
2	road	11	room	14
3	room	10	house	11
4	(table	8)	night	11
5	(car	7)	day	10

5.2 Congruence and divergence

The overall frequencies of the correspondence types displayed in Table 3 indicate that the first hypothesis was correct: the proportion of congruent correspondences is higher when English is the source language. The difference between translations and sources is significant (Fisher’s exact test, P value = 0.0009).

The fact that divergent correspondences are more frequent than congruent ones in both directions shows that the construction represents a point where the languages differ in spite of the similarities in syntactic repertoire pointed out in Section 2.

Table 3. Congruence and divergence

	E orig → N tran		E tran ← N orig	
	N	%	N	%
congruent	187	43.5	167	32.9
divergent	232	54.0	331	65.3
zero	11	2.6	9	1.8
total	430	100	507	100

The higher frequency of congruent correspondences in Norwegian translations may be due to the possibility of using the “N prep N” construction in Norwegian in expressions where other structures may be more common. Example (6) may be an illustration of this; although the translation is correct, the compound *leirlederen* is a more idiomatic alternative.⁶

- (6) *The director of the camp*, not wanting to break the news on the phone, had driven to Baltimore to tell them in person. (AT1)
Lederen av leiren, som ikke hadde lyst til å fortelle nyheten i telefonen, hadde kjørt til Baltimore for å fortelle det personlig. (AT1T)
 Lit: “Leader_{def} of camp_{def} ...”

It may be noted that “the N1 of the N2” occurs as the complement of a preposition in about 65% of the cases in both originals and translations. Norwegian translations from English have similar proportions of congruent and non-congruent correspondences both within and outside prepositional phrases (just under half are congruent, and just over half divergent). In contrast, Norwegian sources tend to diverge from English translations if the pattern is part of a PP: 25% congruent and 73%

⁶ In the 70-million-word *Leksikografisk bokmålskorpus* (“Lexicographical corpus of Norwegian”), *leirleder* occurs 8 times and *leder(en) av leiren* not at all. A Google search found both forms, but *leirlederen* was more than ten times as frequent.

divergent, as compared to equal shares of congruent and divergent correspondences if the sequence is not preceded by a preposition.

Figure 1 shows the types of divergences that occur and the percentage of each (from among the total number of divergent correspondences). The N1 is the most frequent locus of divergence: in Norwegian sources (of English translations), the N1 is divergent five times as often as the N2, and in Norwegian translations (of English originals), the N1 is divergent eight times as often as the N2. It is interesting to note that particularly compounds and adverbs corresponding to the English N1 are more frequent in Norwegian originals while the Norwegian translations have larger shares of the N1 missing and of clausal expansions of the “the N1 of the N2” construction.

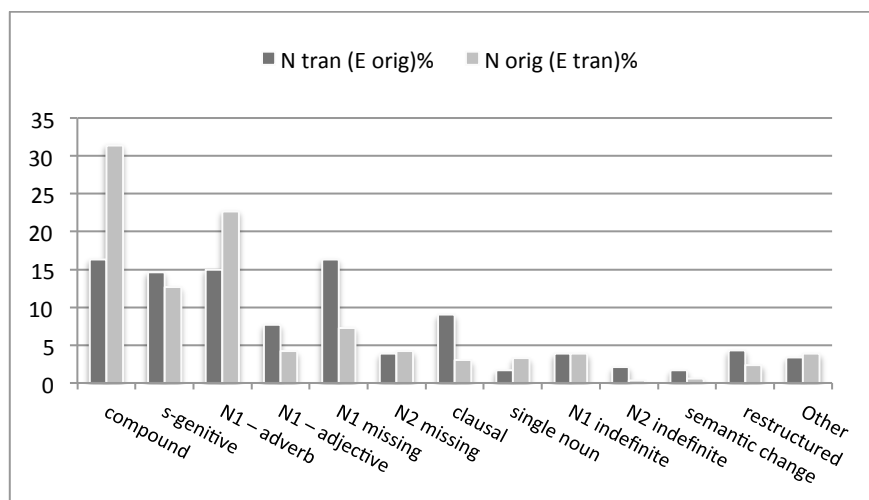


Figure 1: Structural types of divergent correspondences of the “the N1 of the N2” (percentages)

Examples (7)-(13) show the most common types of divergence. In (7) the compound noun *hushjørnet* has been translated by “the N1 of the N2”, while in (8) the English *of*-genitive is rendered by a Norwegian *s*-genitive. Both examples illustrate potentially systematic differences between the two languages that deserve further study: my hypothesis at this point would be that both compounding and the *s*-genitive are more widespread and productive in Norwegian than they are in English.

- (7) She disappeared round *the corner of the house* ... (THA1T)
 Hun forsvant rundt *hushjørnet* ... (THA1)
 Lit: "She disappeared round housecorner_{def}"
- (8) I'm *the head of the household*. (ST1)
 Jeg er *familiens overhode*. (ST1T)
 Lit: "I am family_{def}'s head."

In examples (9) and (10) the English N1 corresponds to an adverb or an adjective in the parallel Norwegian phrase. It should be noted that the phrase *in the middle of* corresponds quite regularly (though not invariably) to the Norwegian adverb *midt* followed by the preposition *i* ('in') or *på* ('on'), and it may be argued that both *in the middle of* and *midt i/på* are complex prepositions. The same pattern is, however, found with other locative nouns, as illustrated by (5) above, where (*on*) *the top of* corresponds to *øverst* ('uppermost'). In the correspondences where the English N1 has been rendered by a Norwegian adjective, shown in (10), it is hard to see any clear patterns.

- (9) What was I doing here in *the middle of the afternoon*? (KF1T)
 Hva gjorde jeg her *midt på ettermiddagen*? (KF1)
 Lit: "What did I here in-the-middle on afternoon_{def}"
- (10) It was *the end of the afternoon*... (JB1)
 Det var *sen ettermiddag*... (JB1T)
 Lit: "It was late afternoon..."

In examples (11) and (12) either the N1 or the N2 is missing from the Norwegian version. The N1s most frequently missing from the Norwegian are locative, as in (11) – the same phenomenon is seen for e.g. *the bottom of*, *the side of*. The missing N2s, on the other hand, tend rather to be either retrievable or inferable from the context, as in (12), where the English translator has deemed it necessary to add the specification of *the street*. This is not, however, simply caused by the translation feature of explicitation, as the reverse phenomenon is equally common in translation from English into Norwegian.

- (11) We stared at one another across *the expanse of the room*. (TH1)
 Vi stirret på hverandre gjennom *rommet*. (TH1T)
 Lit: “We stared at each other through room_{def.}”
- (12) Brita and Hildegun stopped at *the corner of the street*, dropped their schoolbags onto the pavement and sat on the wall. (BV1T)
 Brita og Hildegun stanset på *hjørnet*, slapp skoleveskene på fortauet og satte seg på muren. (BV1)
 Lit: “Brita and Hildegun stopped at corner_{def...}”

All in all, the high proportion of divergence and the types of changes that occur in translation between the languages seem to indicate that Norwegian is less favourable than English is to complex noun phrases of the type “definite noun + preposition + definite noun”.

5.3 Meanings of “the N1 of the N2” and their correspondences

This section looks at the broad meaning categories of N1s outlined in Section 3 to see if they vary in their preferred correspondence types. The main results are shown in Table 4, in which zero correspondences have been ignored due to their low numbers (cf. Table 3).

Table 4. Variation across broad N1 meanings (raw frequencies)

	English original	English translation
	Congruent / divergent	Congruent / divergent
part of N2	88 / 56	58 / 108
locative	38 / 83	21 / 126
owned/caused by N2	18 / 29	36 / 32
person	8 / 13	11 / 10
feature	6 / 13	5 / 11
body part	5 / 10	3 / 5
nominalization	6 / 14	11 / 20
support	7 / 8	8 / 13
temporal	6 / 4	8 / 2
N1 contains N2	-	6 / 2
ordinal	5 / 2	0 / 3

Most of the meaning categories have a majority of divergent correspondences with the important exception of “part of N2” in English

originals. In English translations, however, divergent correspondences are almost twice as frequent as congruent ones within this meaning category. This suggests that English and Norwegian have different preferences in expressions of part-whole relationships even though the same structures may be available. Examples (13) and (14) illustrate a congruent and a divergent correspondence of “N1 is part of N2”, respectively.

- (13) And then *the door of the studio* is opened and several people enter... (ABR1)
 Og så blir *døren til atelieret* åpnet, og flere mennesker kommer inn... (ABR1T)
 Lit: “And then is door_{def} to studio_{def} opened ...”
- (14) The planet has already gone a good way across *the surface of the sun* and will soon pass the center. (EFH1T)
 Planeten er kommet et godt stykke lenger inn over *soloverflaten* allerede, snart passerer den midten. (EFH1)
 Lit: “Planet_{def} is come a good bit longer in over sun-surface_{def} already...”

The Norwegian original in (14) uses a compound noun instead of the complex noun phrase framework, which is the most frequent structural type of divergence in the category of part-whole relationships. Simplification of the noun phrase in terms of omission of either N1 or N2 or the use of a noun simplex also occurs, e.g. *the foot of the garden – hagen* (‘the garden’); *the corner of the room – kroken* (‘the corner’); *details of the case – opplysninger* (‘information’).

Sequences with a locative N1 form the meaning category with the most consistently divergent correspondences. The most common type of divergence is correspondence between an English N1 and a Norwegian adverb. This happens in 59% (74 out of 126) of the divergent locative expressions in Norwegian sources of English translations, and in 39% (32 out of 83) of Norwegian translations. Examples are given in (5) and (9) above. The “the N1 of the N2” framework with locative N1 also frequently corresponds to N1 omission in Norwegian translations (21 out of 83) and to compounds (11 out of 83 in translations and 20 out of 126 in originals). Other divergence types are less frequent.

N1s denoting something that is owned or caused by the N2 have a relatively large proportion of congruent correspondences, especially in Norwegian originals where they outnumber the divergent ones. (15) and (16) show a congruent and a divergent example, respectively. The *s*-genitive, as exemplified in (16), is the most common divergent type within this meaning category in both originals and translation, which is not surprising given the general meaning of possession.

- (15) ... and she was pulled from the road into *the shadow of the bushes*. (PDJ3)
 ... og hun ble trukket fra veien inn i *skyggen av buskene*. (PDJ3T)
 Lit: “and she was pulled from road_{def} in to shadow_{def} of bushes_{def}.”
- (16) His elder son, Robert Falcon Scott, became one of the greatest explorers in *the history of the Empire*. (KH1T)
 Den eldste sønnen, Robert Falcon Scott, ble en av de store oppdagere i *Imperiets historie*. (KH1)
 Lit: “...one of the greatest explorers in Empire_{def}’s history.”

N1s denoting a person have almost equal numbers of congruent and divergent correspondences in both directions of translation. No particular type of divergent correspondence dominates, but compounds and *s*-genitives occur in both originals and translations. *S*-genitives are also a common divergent correspondence of N1s denoting body parts, especially in Norwegian translations from English. When the N1 is a nominalization, clausal expansion is the most common type of divergence in Norwegian translations from English, but in the other direction of translation, compounds are most common, as in (17).

- (17) *The drone of the motor* is powerful and pleasant, ... (EH1T)
Motorduren er sterk og god, ... (EH1)
 Lit: “Motordrone_{def} is strong and good...”

In the remaining meaning categories there are no discernible patterns as regards preferred types of divergent correspondences.

5.4 Structures and meanings in some correspondence types

This subsection will take a closer look at the four divergent correspondence types that differ the most between originals and translations (cf. Fig. 1), namely

- compound (more divergence in Norwegian originals);
- N1 – adverb (more divergence in Norwegian originals);
- N1 missing (more divergence in Norwegian translations);
- clausal correspondence (more divergence in Norwegian translations).

The exploration will include a scrutiny of the N1s and the meaning relations expressed in such correspondences to find out whether particular nouns or meaning types favour particular Norwegian correspondences.

A striking number of Norwegian compounds correspond to “the edge of the _”: 30 out of 104 in English translations and 7 out of 38 in English originals. An example is given in (18). The source of “the edge of the _” is typically a compound ending in *-kanten* (‘the edge’).

- (18) He leaned forward over *the edge of the table*. (GS1T)
Han bøyde seg fram over *bordkanten*. (GS1)
Lit: “He leaned himself forward over table-edge_{def.}”

Other reasonably frequent N1s with compound correspondences are *side*, *mouth*, *surface*, *wall* and *turn*. *Turn* occurs exclusively in *the turn of the century*, corresponding to *århundreskiftet* (‘the year-hundred-shift’); thus indicating that both “the N1 of the N2” sequence and the Norwegian compound are lexicalized.

The most frequent meaning relation between N1 and N2 in the phrases corresponding to compounds is that of part-whole, i.e. the N1 denotes a part of the N2. There are also some locative N1s, but it may be argued that these nouns also denote a part-whole relationship, e.g. *the side of the ship*. We may note that the components of the Norwegian compounds typically have the reverse order of the nouns in the “N1 of

N2” framework: in a part-whole relationship, the compounds start with the whole (e.g. ‘table-edge’).⁷

The N1s corresponding to a Norwegian adverb are typically locative, focusing on some part of the N2, or giving a more precise location within the N2, as shown in (19) and (20). Temporal N1s may have the same type of correspondence, cf. example (9). As noted above, both locative and temporal expressions with *in the middle of* regularly correspond to the Norwegian *midt i/på*, and these expressions account for many of the N1–adverb correspondences (48 out of 75 in English translations and 9 out of 35 in English originals). Other frequent N1s are *bottom*, *centre* and *front*. The locative noun corresponding to an adverb typically occurs after a preposition, as in (19); thus the adverb corresponds to the preposition plus the locative noun, as shown also by (20).

- (19) A small lamp stood in *the centre of the table*, ... (EG2T)
 En liten lampe med gul silkeskjerm var plassert *midt på bordet*...
 (EG2)
 Lit: “A small lamp with yellow silkshade was placed in-the-middle on table_{def}”
- (20) *At the bottom of the hill* is the Lady with the Fleas. (LSC1T)
Nederst i bakken kommer Damen med maurene, ... (LSC1)
 Lit: “Nethermost in hill_{def} comes Lady_{def} with ants_{def}...”

Correspondences with a missing N1 equivalent are most frequent in Norwegian translations from English. No N1 lemma is omitted or added noticeably more often than others in this group. However, many such N1s are locative, of the type that Sinclair (1991: 87) calls “focus nouns”, which “[specify] some part of the N2”. Sinclair regards the N2 as the headword in such constructions, and such analysis may explain the omissibility of the N1; they are semantically less important than the N2 in the framework “the N1 of the N2”. An example is given in (21); see also (11) above.

⁷ The reversed order of N1 and N2 also occurs with *s*-genitive correspondences, e.g. *the back of the letter* – *brevets bakside* (‘the letter’s backside’); *the colours of the rainbow* – *regnbuens farger* (‘the rainbow’s colours’).

- (21) He shrugged and put it at *the back of the safe*. (FF1)
Han trakk på skuldrene og la kopien inn i *safen*. (FF1T)
Lit: “He shrugged on shoulders_{def} and put copy_{def} in to safe_{def}.”

Correspondence by means of a clausal structure is shown in (22). This type of correspondence is more common in Norwegian translations than in sources, although it occurs in both types of text, cf. Figure 1.

- (22) Mattie looked for *the direction of the voice*. (GN1)
Mattie snudde seg for å se *hvor stemmen kom fra*. (GN1T)
Lit: “Mattie turned herself for to see where voice_{def} came from.”

N1s in frameworks with clausal correspondences include those that can be read as nominalizations, as in *the outbreak of the punk vogue – da punk-bølgen brøt ut* (‘when the punk-wave broke out’), and those that reflect subjective genitives (Quirk et al. 1985: 1278), e.g. *the occupants of the cars – de som satt i bilene* (‘those who sat in the cars’). Whereas it may be hard to find a Norwegian noun phrase corresponding idiomatically (in this context) to *the direction of the voice* in (22), other examples may simply show a preference for clausal rather than phrasal expression, e.g. (23), where the Norwegian verb *hete* (‘be called’) is available to convey the meaning of *the name of*.⁸

- (23) *The name of the gallery* is Sub-Versions, one of those puns that used to delight me before they became so fashionable. (MA1)
Galleriet heter Sub-Versjoner, en type ordspill som pleide å more meg før det gikk inflasjon i dem. (MA1T)
Lit: “Gallery_{def} is-called Sub-Versions...”

5.5 Variation across particular N1s

This final part of the corpus investigation turns to the most frequent lexical items to be found in the N1 position in the colligational framework under study: *middle*, *edge*, *rest*, and *end*. As shown in Figure 2, the four lexemes vary greatly as regards their typical correspondence

⁸ The Norwegian verb *å hete* does not have a straightforward English equivalent, but means ‘to be called’ (cf. French *s’appeller* and German *heissen*).

patterns. *Middle* has mostly divergent correspondences in both translations and sources; *edge* has more divergent sources than translations; *Rest* has the greatest proportion of congruence of the set, and *end* shows opposite trends in sources and translations.

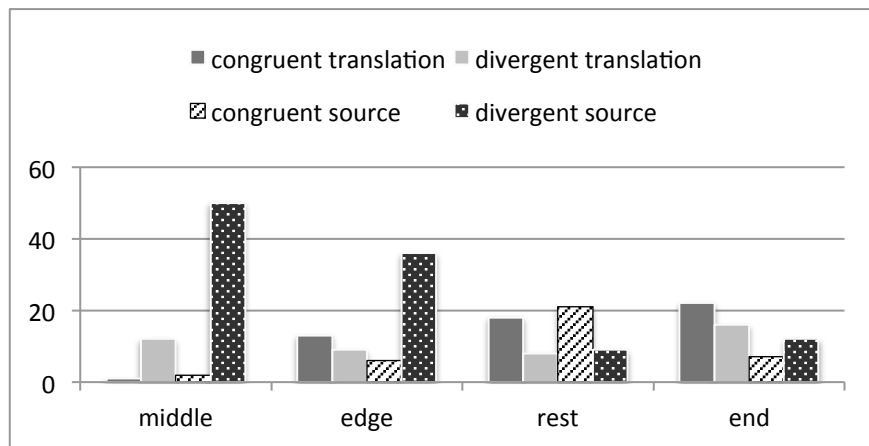


Figure 2. Congruent and divergent correspondences across the most frequent N1s (raw frequencies).

As pointed out repeatedly below, the most frequent correspondence *par excellence* of (*in*) *the middle of the* is *midt i/på*. Although both *in the middle of* and *midt i/på* may be seen as complex prepositions, *in the middle of* also has a congruent correspondence in Norwegian, shown in (24). However, this expression is very infrequent, with one occurrence in Norwegian translations and two in Norwegian originals. Divergent correspondences other than *midt i/på* are also very infrequent.

- (24) The man *in the middle of the semicircle* was taller than the others.
 (MN1T)
 Mannen *i midten av halvsirkelen* var høyere enn de andre. (MN1)
 Lit: “Man_{def} *in* middle_{def} of semicircle_{def} was taller than the others.”

The edge of the N2 is another sequence with mainly divergent correspondences. The most frequent divergent correspondence type of *the edge of the* is compound, e.g. *bordkanten* (‘the table-edge’), *bassengkanten* (‘the pool-edge’). Other divergent correspondence types

are rare, with N1 omission being the second most frequent type, exemplified by (25). As shown in Figure 2, *the edge of the* N2 also has congruent correspondences, illustrated by (3) above.

- (25) Her long, grey braids were barely visible over *the edge of the skin rug*. (TTH1T)
 Den lange, grå fletta hennes stakk såvidt fram fra *skinnfellen*. (TTH1)
 Lit: “The long, grey braid_{def} hers showed barely forth from skin-rug_{def}.”

The rest of the + N2 differs from the two sequences discussed above by having predominantly congruent correspondences. The N2 in this sequence is often temporal, e.g. *resten av dagen / uken* (‘the rest of the day/week’). Divergent correspondences are mainly non-recurrent except N1–adjective in Norwegian translations, as shown in (26). The adjectives corresponding to *rest* as N1 in this pattern are quantifying; thus they reflect the quantifier character of the sequence *the rest of*.⁹

- (26) She had completely forgotten now about *the rest of the class*. (RD1)
 Hun hadde fullstendig glemt alle *de andre elevene*. (RD1T)
 Lit: “She had completely forgotten all the other pupils_{def}.”

The last sequence to be discussed in this section, *the end of the*, is congruent more often in translation and divergent more often in sources. However, the difference is small. The most frequent type of divergence found in translation is clausal expansion (six cases), exemplified by (27). In sources the most frequent divergences are *s*-genitive (four examples), as in *the end of the world – verdens ende* (‘the world’s end’), and N1 corresponding to an adjective, as in (28).

- (27) Now, near *the end of the day*, it was clear that hope had been in vain. (AH1)

⁹ In this respect *the rest of the* is analogous to other quantifying phrases with *of* such as *a lot of* (cf. Renouf & Sinclair 1991), which may be the clearest cases of the N2 acting as notional head of the noun phrase (cf. Owen 2007: 212).

Nå da *dagen nærmet seg slutten*, var det tydelig at håpet hadde vært forgjeves. (AH1T)

Lit: “Now when day_{def} neared itself end_{def}, was it clear...”

- (28) *At the end of the working day*, with a more or less favorable outcome in my briefcase, it feels good to step into the dim interior of the restaurant... (KF1T)

Etter *endt arbeidsdag* med et mer eller mindre lovende resultat i dokumentmappen, føles det godt å trå inn i den halvmørke restauranten... (KF1)

Lit: “After ended workday with a more or less promising result in briefcase_{def}...”

6 Summary of findings and concluding remarks

The present study has been exploratory in that it has tested the possibility of using a colligational framework as the starting point for contrastive analysis. A number of findings have been presented. First of all, “the N1 of the N2” seems to select the N1 more systematically than the N2. The selected N1s typically focus on a part of the N2 (cf. Sinclair 1991), often with locative/temporal meaning. Possessive meaning, expressed by *of*-genitives, is also frequent. Support nouns, however, which were common in the pattern “a _ of” (Renouf & Sinclair 1991), are not particularly frequent in the present material.

Norwegian correspondences of the framework are diverse, and differ between the directions of translation. Congruent correspondences, defined as “definite noun + preposition + definite noun”, account for 44% of Norwegian translations and 33% of sources, and the difference is greater if the pattern is part of a prepositional phrase (section 5.2). Of the structural types of divergence, compound nouns were most frequent, closely followed by N1–adverb correspondences and *s*-genitives. It is noteworthy that most of the divergent correspondence types avoid the postmodification inherent in “the N1 of the N2” framework. This is in line with Elsness’s (2014) findings that Norwegian uses fewer (clausal) postmodifiers than English, and that the content of the modification is more frequently expressed outside the noun phrase in Norwegian (2014: 117).

Some of the structural types of divergent correspondences differ across sources and translations: compounds and N1–adverb are more common in Norwegian sources; N1 omission and clausal correspondence are more common in Norwegian translations. The Norwegian preference for adverbs instead of nouns to denote location seems to be relatively systematic, and agrees with the general hypothesis that Norwegian prefers verbal to nominal style to a greater extent than English (cf. Nordrum 2007).

The broad meaning categories of the framework “the N1 of the N2” has also proved important for the preferred type of Norwegian correspondence. The most striking example is the category of locative N1s, which very often correspond to Norwegian adverbs, particularly if they are preceded by a preposition (in which case the Norwegian adverb corresponds to preposition + N1, e.g. *at the foot* – *nederst* (‘nethermost’)). Sequences where the N1 is a part of N2 can have congruent correspondences in Norwegian; however, it seems that Norwegian originals prefer other ways of expressing this relation, particularly compound nouns. It should be noted that Norwegian has a prepositional genitive analogous to the *of*-genitive, which accounts for the comparative frequency of congruent correspondences in the categories part-whole and owned/caused by N2.

As shown in Section 5.5, individual lexemes may also come with their preferred correspondence types: (*in*) *the middle of* regularly corresponds to *midt i/på*; a correspondence that seems quite conventionalized and is found e.g. in the bilingual dictionary *Engelsk stor ordbok* (Kunnskapsforlaget 2001). Another frequent N1, *edge*, regularly corresponds to the second part of a compound noun. *The rest of* corresponds congruently and frequently to *resten av*, while *the end of* has less clear correspondence patterns. A less frequent sequence, *the name of*, was shown to correspond to a clausal expansion with the verb *hete*, thus showing how languages may possess different lexicogrammatical resources for expressing the same meaning.

This investigation has been exploratory, and a number of issues remain for future studies, such as investigations of related colligational patterns (e.g. “a(n) N1 of a(n) N2”) and their Norwegian correspondences. There is good reason to believe that such colligational frameworks are register-sensitive, cf. Biber et al (1999: 1014); thus a study involving more registers would also be worthwhile.

A potential problem of the present study is that the search was unidirectional, going only from English to Norwegian. This is probably unavoidable in an investigation based on function words: the languages cannot be expected to have function words that are similar enough to perform the same functions in phraseological frameworks. In this case, the main problem was the morphological difference in how definiteness is marked, with the Norwegian definite suffix even varying for gender and number. In addition, there is probably no Norwegian preposition with the same degree of multifunctionality as English *of*, so there is no way of searching lexically for Norwegian sequences equivalent to “the N1 of the N2”. The tagged ENPC allows searches for two nouns with an intervening preposition, but for the present framework, this search string has a very low degree of precision; in most of the hits, the PP was not part of the same phrase as the N1, and due to the lack of morphological annotation, the search string cannot discriminate between definite and indefinite nouns. Furthermore, with the present extent of corpus annotation it is impossible to search for Norwegian compound nouns, which might have been an interesting starting point of a different investigation. There may thus be a need for developing more sophisticated ways of searching for patterns across languages.

In spite of its limitations, the present study has confirmed that English and Norwegian differ in their use of noun phrases. The differences are to some extent structural, but more importantly, the languages seem to select differently from relatively similar lexicogrammatical resources: Norwegian uses compound nouns more than English does, and favours a number of expressions that are structurally less complex than the English “the N1 of the N2” framework.

The most important implication of this study is arguably its demonstration of the feasibility of exploring structures and meanings cross-linguistically on the basis of function words. Such a way of exploiting parallel corpora represents a bottom-up approach to the relationship between form(s) and meaning(s) and can “give new insights into the languages compared – insights that are likely to be unnoticed in studies of monolingual corpora” (Aijmer & Altenberg 1996: 12). Renouf & Sinclair observe that the study of collocational frameworks can “raise consciousness of the many different and eminently sensible ways we might develop to present and explain language patterning” (1991: 143). This is certainly the case in cross-linguistic study too. Let this final

corpus example illustrate how and why different languages have different preferences in expressing similar meanings:

- (29) It's the way of the world. (RD1)
Sånn er det bare. ("Such is it simply")

References

- Aijmer, Karin & Bengt Altenberg. 1996. Introduction. In K. Aijmer, B. Altenberg & M. Johansson (eds.), *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies*. Lund: Lund University Press, 11-16.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Brems, Lieselotte. 2015. Some binominal size noun constructions in English and French: a contrastive corpus-based perspective. *Nordic Journal of English Studies* 14:1, 90-115.
- Elsness, Johan. 2014. Clausal modifiers in noun phrases: a comparison of English and Norwegian based on the Oslo Multilingual Corpus. S.O. Ebeling, A. Grønn, K. R. Hauge & D. Santos (eds), *Corpus-based Studies in Contrastive Linguistics*. Oslo Studies in Language 6(1), 91-118.
- Engelsk stor ordbok. Engelsk-norsk / norsk-engelsk*. 2001. Oslo: Kunnskapsforlaget.
- Faarlund, Jan Terje, Svein Lie, Kjell Ivar Vannebo. 1997. *Norsk Referansegrammatikk*. Oslo: Universitetsforlaget.
- Groom, Nicholas. 2007. *Phraseology and Epistemology in Humanities Writing*. PhD thesis, University of Birmingham. Accessed from <http://ethos.bl.uk>, 30 July 2015).
- Hunston, Susan. 2008. Starting with the small words. Patterns, lexis and semantic sequences. *International Journal of Corpus Linguistics* 13:3, 271-295.
- Johansson, Stig. 2007. *Seeing through Multilingual Corpora. On the Use of Corpora in Contrastive Studies*. Amsterdam: Benjamins.
- Johansson, Stig. 2012. Cross-linguistic perspectives. In Merja Kytö (ed.), *English Corpus Linguistics: Crossing Paths*. Amsterdam / New York: Rodopi, 45-68.

- Johansson, Stig, Jarle Ebeling & Signe Oksefjell. 1999/2002. English-Norwegian Parallel Corpus. Manual. Department of British and American Studies, University of Oslo. <http://www.hf.uio.no/ilos/tjenester/kunnskap/sprak/omc/enpc/ENPCmanual.html>, accessed 29 July 2015.
- Keizer, Evelien. 2007. *The English Noun Phrase. The Nature of Linguistic Categorization*. Cambridge: Cambridge University Press.
- Micoli, Annalisa. 2010. “‘We can pull it off, if we hold on... Men skal det ta av, hvis vi står på?’ *Off* versus *av* and *on* versus *på*: a corpus-based contrastive investigation”. MA thesis, University of Oslo. Available at <http://urn.nb.no/URN:NBN:no-25295>.
- Nordrum, Lene. 2007. *English Lexical Nominalizations in a Norwegian-Swedish Contrastive Perspective*. PhD thesis, Göteborg University. Available at <https://gupea.ub.gu.se/handle/2077/17181>.
- Owen, Charles. 2007. Notes on the ofness of *of* – Sinclair and grammar. *International Journal of Corpus Linguistics* 12:2, 201-222.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
- Renouf, Antoinette and John Sinclair. 1991. Collocational frameworks in English. In K. Aijmer and B. Altenberg (eds). *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London: Longman, 128-144.
- Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Corpora

- English-Norwegian Parallel Corpus, see <http://www.hf.uio.no/ilos/tjenester/kunnskap/sprak/omc/enpc/>
- British National Corpus, see <http://www.natcorp.ox.ac.uk/>
- Leksikografisk bokmålskorpus, see <http://www.hf.uio.no/iln/tjenester/kunnskap/samlinger/bokmal/veiledningkorpus/>