

Hope for the future: An analysis of HOPE/HÅP(E) across genres and languages

Signe Oksefjell Ebeling

University of Oslo (Norway)

This article reports on a contrastive study of the cognate nouns and verbs HOPE and HÅP(E) that investigates their lexico-grammatical conditions of use in English vs. Norwegian fiction texts and football match reports. The complex dataset consists of material from a parallel corpus of fiction texts and a comparable corpus of football match reports. An interesting finding is that the verb use outnumbers the noun use in the fiction texts, whereas the noun use outnumbers the verb use in the match reports in both languages. Moreover, the analysis of the lemmas suggests that they have similar potential of use but with slightly different preferences, both across the genres and languages. It is also suggested that the English lemmas are more consistently used in negative contexts than the Norwegian ones. Finally, the method of combining data from two different types of contrastive corpora proved fruitful, as the results become more robust.

Keywords: cognates, comparable corpus, bidirectional parallel corpus, fiction, football match reports, English/Norwegian

1. Introduction and aims

This article presents a contrastive analysis of the cognates HOPE and HÅP(E) through their use in two languages (English and Norwegian) and two genres (fiction and football match reports). In a previous contrastive study of English and Norwegian football match reports it was found that the cognate nouns HOPE and HÅP featured as keywords in texts reporting on defeat (Ebeling, 2019). The reason for this frequent use of HOPE and HÅP in the defeat section of the English-Norwegian Match Report Corpus (ENMaRC) was attributed to the items' frequent use in contexts where hopes are dashed, as in examples (1) and (2).

- (1) However those *hopes* were dashed on 55 minutes when the Gunners added a second. (ENMaRC/CPFC)¹
- (2) Det tente et ørlite *håp* som ble knust desto mer brutalt fem minutter etter. (ENMaRC/VFK)
Lit.: That lit a tiny hope that was dashed even more brutally five minutes later

¹ The ENMaRC corpus text code is the same as the official acronyms of the clubs (CPFC = Crystal Palace Football Club). See Ebeling (2021) for an overview of clubs and acronyms in the ENMaRC.

It was speculated that the use attested in (1) and (2) may be more typical of match reports than of language in general. The fact that HOPE and HÅP feature in reports describing defeats suggests that words may be coloured by their immediate context to the effect that they take on a meaning that is the opposite of what might otherwise be the case (in their typical contexts in other genres). Therefore, part of this study aims to find out to what extent this (negative) use of the nouns HOPE and HÅP and their verb counterparts (HOPE and HÅPE) is overrepresented in the genre of football match reports, or whether it may be seen to extend to other genres.

As mentioned, the second genre under scrutiny here is fiction, and the reason for this choice is twofold. First, it was deemed necessary to investigate the use of the lemmas in a genre that is clearly distinct from written match reports.² Second, from a contrastive perspective, it was deemed necessary to objectively determine the degree of equivalence between the English and Norwegian lemmas on the basis of bidirectional translation data to be certain that we compare like with like. In other words, an English-Norwegian bidirectional corpus had to be consulted, and the only corpus containing enough data from one relatively homogeneous genre is the English-Norwegian Parallel Corpus+ (ENPC+) of fiction texts.

Drawing on data from the ENMaRC and ENPC+, this study seeks to dig deeper into the lemmas HOPE and HÅP(E) by contrasting their lexico-grammatical conditions of use across languages and genres. Preliminary scrutiny of concordance lines suggested that the lemmas have the potential to feature in a number of different (phraseological) contexts in both languages. However, when compared to fiction, match reports seem to make a narrower selection from the full potential of contexts in which the lemmas may be used (cf. Stubbs, 1996: 89).

Based on these observations and previous findings, the study seeks answers to the following questions:

- To what extent are the lemmas used similarly in English and Norwegian?
- To what extent are the lemmas used similarly in match reports and fiction?
- To what extent does the use of different types of “contrastive” corpora contribute to our cross-linguistic knowledge of the lemmas?

The aim is to provide new insights into the actual use and lexico-grammatical features of these lemmas, not only from a cross-linguistic perspective but from a cross-linguistic genre perspective. This ties in with a recent trend in contrastive studies, in which more attention is given to cross-linguistic variation across genres or registers (see e.g. Dupont and Zufferey, 2017; Lefer and Vogeeler, 2014; Neumann, 2014; Teich, 2003). Moreover, the study addresses potential benefits of using both comparable and (bidirectional) translation corpora to widen the horizons of contrastive studies.

The study starts with a general description of the rather complex data under investigation by introducing the corpora used in Section 2. An outline of the contrastive approach and method applied is offered in Section 3, including an account of how the data were extracted and an overview of the material used in the analysis. A cross-linguistic, cross-genre analysis of the actual uses of the lemmas is carried out in Section 4, followed by a discussion of some of the findings in Section 5. Section 6 revisits the research questions and offers some concluding remarks and suggestions for future studies.

² Cf. Biber’s (1993: 336) multidimensional analysis, where fiction and press reportage, of which match reports can be seen as a sub-register, are shown to differ according to several linguistic features.

2. Corpora

As mentioned in Section 1, the material for the present study is extracted from the English-Norwegian Parallel Corpus+ (ENPC+) and the English-Norwegian Match Report Corpus (ENMaRC). The ENPC+ is a bidirectional translation corpus of published fiction and its structure and potential are illustrated in Figure 1 (Johansson and Hofland, 1994). It contains 39 fiction texts originally written in each of the languages with their translations into the other. The texts were published in the period from 1980 to 2012 and include both full-length novels (eight in English and nine in Norwegian) and extracts of 12,000 to 15,000 words (31 in English and 30 in Norwegian). In total, the ENPC+ contains ca. 5.3 million words, i.e. roughly 1.3 million words in each of the sub-corpora: English originals, Norwegian originals, English translations, Norwegian translations. For a more detailed description of the ENPC+, see Ebeling and Ebeling (2013).

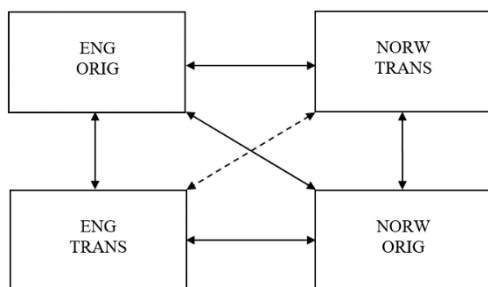


Figure 1. The bidirectional structure of the ENPC+.

This corpus structure enables contrastive studies of a comparable nature, using material from the original texts only, as well as of a parallel nature, using material from the original and their aligned translated texts in both directions. From a translation studies perspective, the potential of comparing original and translated texts in the same language is also a valuable feature of this corpus structure (see e.g. Ebeling and Ebeling, 2017; Ebeling, forthcoming), although not relevant to the present study.

The English-Norwegian Match Report Corpus is a comparable corpus of online written football match reports from the English Premier League and the Norwegian *Eliteserie*. It is comparable according to Johansson (2007: 9), in the sense that it contains original texts in two languages matched by criteria such as genre, time of publication, etc. (see also Ebeling and Ebeling 2020). Its structure is illustrated in Figure 2, corresponding to the boxes connected by the slant solid double arrow in Figure 1.

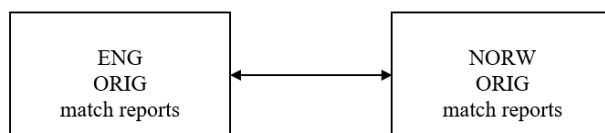


Figure 2. The structure of the ENMaRC.

The match reports are written by the respective football clubs' own journalists and published online on the clubs' web pages immediately after each match. The ENMaRC contains match reports from two seasons, 2016–17 and 2017–18 in the case of the Premier League and 2017 and 2018 in the case of the *Eliteserie* (ES).³ The Premier League part of the corpus contains

³ The football season in England runs from August until May (hence 2016–2017 and 2017–2018) and the season in Norway runs from March until November (hence 2017 and 2018).

reports from 23 teams and amounts to approximately 990,000 words, while the *Eliteserie* part contains reports from 18 teams, amounting to around 315,000 words. Although there are some challenges relating to corpus size when comparing the use of HOPE and HÅP(E) in fiction and match reports, these will be kept to a minimum through the use of normalised frequencies, some (mainly descriptive) statistics and qualitative case studies. Another potential challenge relating to corpus comparability is the time period covered in the ENPC+ vs. the ENMaRC. However, it is not believed that the use of these lemmas has changed much since the earliest texts in the ENPC+ (1980s) to the most recent texts in the ENMaRC (2018).

3. Contrastive approach, method and material

With both a comparable corpus (ENMaRC) and a bidirectional parallel corpus (ENPC+) at hand a generally sound contrastive approach is ensured and the contrastive corpus model ensuing from the combination of the two can be illustrated as in Figure 3.

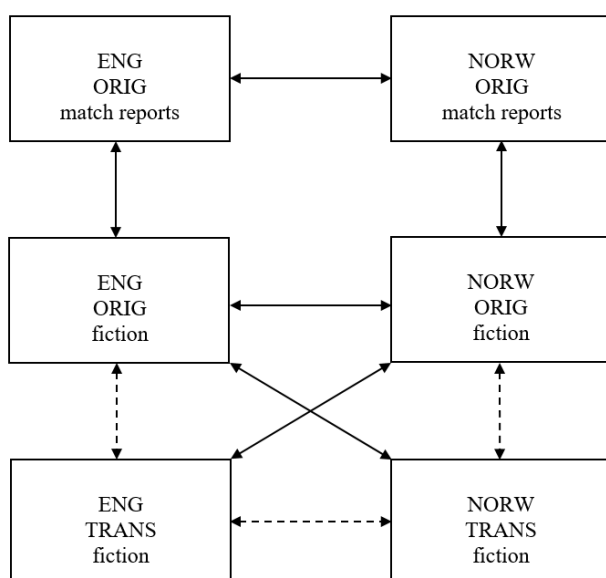


Figure 3. The two-genre comparable-cum-bidirectional corpus model.

The structure and potential of the model can be summed up as a two-genre comparable-cum-bidirectional corpus model. For the purpose of this study, the bidirectional fiction part is mainly used to objectively establish the comparability of the items compared by assessing the items' intertranslatability in a measure of Mutual Correspondence (Altenberg, 1999). Although HOPE and HÅP(E) are etymologically cognate,⁴ and as such fulfil the criterion of the presence of the comparability criterion of a perceived similarity as outlined by Chesterman (1998: 54), their comparability is further strengthened by a Mutual Correspondence of a staggering 95% for the verbs HOPE and HÅPE and an almost equally staggering 91.3% for the nouns HOPE and HÅP in the ENPC+.⁵ This demonstrates that the lemmas are very good cross-linguistic matches of each other and they can safely serve as the starting point of a contrastive analysis. Typical examples are shown in (3) from English into Norwegian and in (4) from Norwegian into English.

⁴ From Middle Low German and Middle Dutch *hope* (*Oxford English Dictionary* and *Det Norske Akademis ordbok*).

⁵ Mutual Correspondence refers to “the frequency with which different (grammatical, semantic and lexical) expressions are translated into each other”, ranging from “0% (no correspondence) to 100% (full correspondence)” (Altenberg, 1999: 254).

- (3) Long after all *hope* had gone Stanton stood there and waited for something to happen... [ENPC+/MoAl1E]⁶
Lenge etter at alt *håp* var ute, sto Stanton der og ventet på at noe skulle skje ... [ENPC+/MoAl1TN]
- (4) — La oss *håpe* at snøen dekker ham til før noen oppdager at han ligger der. [AnHo2N]
“Let’s *hope* the snow will cover him before anybody sees him. [AnHo2TE]

In the cross-linguistic, cross-genre analysis proper, the ENPC+ will not be used to its full potential, and the contrastive analysis will from now on be based on the comparable texts only: fiction and match reports originally written in English and Norwegian (cf. the top four boxes in Figure 3). In the following, the steps taken in the analysis will be described, focusing on the lemmas’ phraseological potential in the two languages and genres.

The first step was to search for all forms of the lemmas using the ENPC+ search interface for the fiction texts⁷ and AntConc (Anthony, 2019) for the match reports. As the corpora are not part-of-speech tagged, manual disambiguation of noun and verb uses had to be performed on the full set of search results (raw numbers): 375 and 112 for the Norwegian forms (*håp|håpet|håper|håpte|håpa*)⁸ in the ENPC+ and ENMaRC, respectively, and 450 and 324 for the English forms (*hope|hopes|hoped|hoping*)⁹.

Table 1 shows the number of occurrences of all noun and verb forms of HOPE and HÅP(E) in the four sub-corpora, both in terms of raw frequencies and normalised frequencies per 100,000 words.

Table 1. Number of occurrences of HOPE and HÅP(E) in the ENPC+ and ENMaRC.

Word forms	ENPC+			ENMaRC		
	Occ. per 100,000 words (Raw freq.)			Occ. per 100,000 words (Raw freq.)		
	Noun	Verb	TOTAL	Noun	Verb	TOTAL
hope hopes hoped hoping	6.2 (83)	27.3 (367)	33.5 (450)	23.7 (235)	8.9 (88)	32.6 (324)
håp håpet håper håpte håpa	8.3 (109)	20.3 (266)	28.6 (375)	20 (63)	15.6 (49)	35.6 (112)

From the “Total” columns, it can be observed that HÅP(E), including all forms, is more frequently attested in match reports than in fiction in Norwegian (35.6 phtw vs. 28.6 phtw), whereas the opposite is the case for English HOPE, although only marginally so (33.5 phtw vs. 32.6 phtw). However, as is evident from Table 1, it is not merely a question of differences between the genres and languages, but also between word classes. This is visualised more clearly in Figure 4, where the marked differences in noun (green) vs. verb (blue) uses are fairly obvious. For the purpose of this visualisation percentages are used to illustrate the proportions of noun vs. verb uses. Although there are some outliers in the material, these do not significantly affect these proportions.¹⁰

⁶ The ENPC+ corpus code identifies the author of the text (MoAl = Monica Ali), text number by that author (1) and language (E). The code of the Norwegian translation (T) of this text is MoAl1TN. See Ebeling and Ebeling (2013) for an overview of texts and text codes in the ENPC+.

⁷ Developed by J. Ebeling and hosted by the University of Oslo (restricted access and password protected).

⁸ The forms *håpt* (past participle), *håpende* (present participle) and *håpene* (definite noun, plural) were not attested in the ENPC+.

⁹ An additional 11 instances of *Hope* as a proper noun have been left out of this study.

¹⁰ For example, for verbs in English fiction there is one outlier. However, a Wilcoxon Rank Sum test (as implemented in R) shows that there is no significant difference in the material with and without this outlier ($p=0.85$); the same applies to the other sub-corpora that have between one and three outliers.

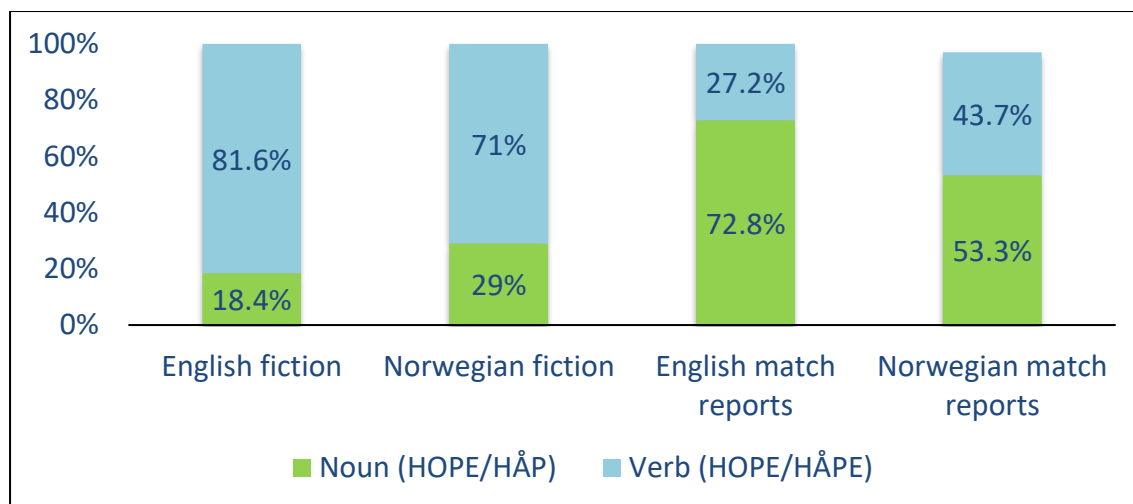


Figure 4. Distribution of noun and verb uses in the ENPC+ (fiction) and ENMaRC (match reports). (See Table 1 for raw numbers and normalised frequencies per 100,000 words.)

As can be observed in Figure 4, the distribution is more similar across the two languages than across the two genres, i.e. verbs are more common in fiction in both languages, whereas nouns are more common in match reports. These genre differences are in fact statistically significant for both nouns and verbs in English fiction vs. match reports and for nouns in Norwegian fiction and match reports.¹¹ This is in line with what Biber *et al.* (1999) note for English fiction and news, of which football match reports can be seen as a sub-category:

The lexical word classes [...] vary greatly across registers: Nouns are most common in news (and to a lesser extent in academic prose); they are by far the least common in conversation. [...] Verbs and adverbs are most common in conversation and fiction. (Biber *et al.*, 1999: 65)

Figure 4 also suggests that these preferences (for verb in fiction and noun in match reports) are more prominent in English than in Norwegian. In other words, there is a narrower difference between the two word classes in the Norwegian material, particularly in the football match reports. This greater presence of verbs in the match reports may be related to what Nordrum (2007) notes in her dissertation on nominalizations in an English-Norwegian-Swedish contrastive perspective, namely that “there is a particularly strong and well-established prescriptive norm in Norway and Sweden favoring a ‘verbal’ or ‘oral’ style” (Nordrum, 2007: 219). This does, however, not explain the larger proportion of nouns in the Norwegian fiction material compared to English. And although the difference is not statistically significant for nouns in English vs. Norwegian fiction ($p=0.3117$), it is an observation that deserves further study in the future.

Following this general overview of noun and verb uses, the analysis now proceeds into the lexico-grammatical features of each word class in a comparison of their uses across the two languages and genres.

¹¹ Not all datasets were normally distributed, thus a Wilcoxon Rank sum test (in R) was chosen for the significance test, returning the following results: $p<0.0001$ for nouns and $p<0.001$ for verbs in English fiction vs. match reports, respectively, and $p<0.05$ for nouns in Norwegian fiction vs. match reports. The difference in verb uses in the Norwegian genres was not statistically significant ($p=0.1413$).

4. Cross-linguistic and cross-genre analysis of the noun and verb lemmas

Section 4.1 starts with an overview of the contextual features relevant to the English and Norwegian nouns in the material before moving on to a comparison of the features that stand out as being typical in each of the sub-corpora, i.e. English football match reports, Norwegian football match reports, English fiction and Norwegian fiction. Section 4.2 follows the same structure for the verbs.

4.1 The nouns HOPE and HÅP

To determine the phraseology of the nouns, the following contextual features were registered:

- **Form:**
 - Singular/Plural
- **Modification:**
 - Premodification (adjective | noun)
 - Postmodification (incl. apposition) (PP | \emptyset -*that* clause | *that*-clause | infinitive clause | relative clause)
- **Syntactic function:**
 - Head of NP and (part of) S | dO | sP
 - Head of NP and part of prepositional complement
- **Context** (negative | not negative)
- **Verb collocate**

Examples (5) and (6) serve to illustrate this classification scheme.

- (5) I had brought with me a new *hope*. [ENPC+/BO1]
- (6) ... og satte inn unggutten Erling Braut Håland i *håp* om å skape mer. I stedet var det ... [ENMaRC/VIF]
 Lit: ‘and brought on the young lad EBH in hope about to create more. Instead was it...’

In (5), *hope* is in the singular, premodified by the adjective *new* and head of a noun phrase functioning as the direct object. There is no evidence of a negative outcome, thus the context is deemed ‘not negative’, and the verb collocate, i.e. the verb in the clause containing *hope*, is *brought*. In the Norwegian example in (6), *håp* is also in the singular form, postmodified by a prepositional phrase (*om å skape mer* ‘of creating more’; lit.: about to create more’) and part of a prepositional complement following the preposition *i*. There are contextual clues suggesting that the context is negative (i.e. *I stedet var det ...* ‘instead it was’; lit.: instead was

it),¹² and there is arguably no verb collocate, as *skape* ‘create’ is part of a clause embedded within the postmodifying PP of *håp* and therefore not directly linked to it.

Table 2 gives a numerical overview of the selected contextual features in terms of proportions (i.e. percentages of total) within each sub-corpus, while Figure 5 visualises these according to the corpus model presented in Figure 3 (comparable parts).

Table 2. Main contextual features of the nouns HOPE and HÅP and their frequency (raw and percentage of total number of occurrences within each sub-corpus).

	English match reports		Norwegian match reports		English fiction		Norwegian fiction	
	Raw	% of total (235)	Raw	% of total (63)	Raw	% of total (83)	Raw	% of total (109)
PP postmodification	116	49.4%	30	47.6%	42	50.6%	54	45.5%
Premodification	40	17%	20	31.7%	15	18.1%	19	17.4%
Modification (pre or post)	155	66%	51	81%	57	68.7%	75	68.8%
HOPE/HÅP as part of S/dO/sP	181	77%	42	66.7%	59	71.1%	61	56%
Genre-specific verb collocates ¹³	100 / 182	55%	28 / 42	66.7%				
General verb collocates ¹³					40 / 56	71.4%	50 / 59	84.7%
Plural HOPE/HÅP	120	51%	0	0%	10	12%	4	3.7%
Negative contexts	116	49.4%	35	55.6%	40	48.2%	26	23.9%
HOPE/HÅP part of prep. complement	53	22.6%	19	30.2%	21	25.3%	42	38.5%

In Table 2, salient contextual features in the sub-corpora, represented as percentages, have been highlighted in bold and have been included in Figure 5. A feature is considered salient either if it is found in a minimum of ca. 50% of the cases, or if it is proportionally more frequent in a particular sub-corpus compared to the others, e.g. HÅP as part of a prepositional complement in the two Norwegian sub-corpora.

¹² It should be noted that, although it is difficult to objectively operationalise the feature of negative vs. non-negative context, the contextual clues are often quite clear in this regard.

¹³ The percentages for verb collocates are calculated on the basis of a reduced number of occurrences, as verb collocates do not feature in instances where the nouns are part of a prepositional complement; cf. example (6). Thus, the total number of occurrences with verb collocates is reduced to 182 in the English match reports, to 42 in the Norwegian match reports, to 56 in the English fiction texts and to 59 in the Norwegian fiction texts.

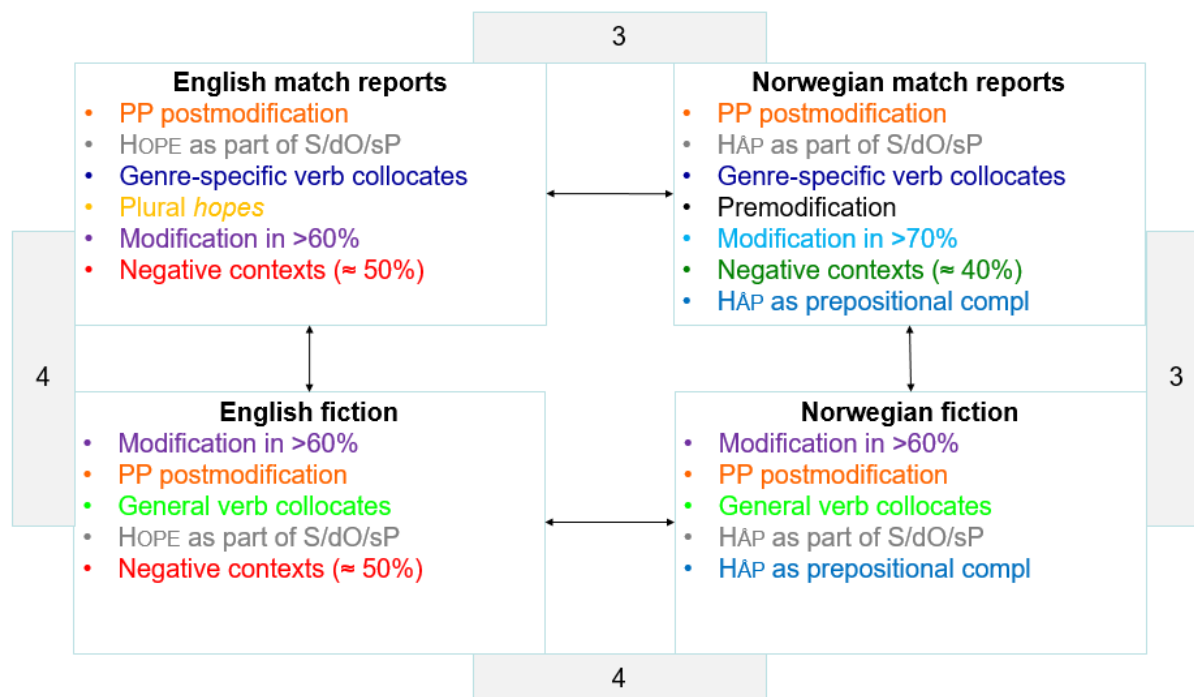


Figure 5. Contextual features of the nouns HOPE and HÅP: Main tendencies.

In Table 2 and Figure 5, we can observe that some characteristics are general for the two nouns in the two genres and languages:¹⁴ postmodification by a PP, the noun is part of the S/dO/sP. A quantification of the similarities and differences is captured in the grey-shaded boxes connecting the different sub-corpora in Figure 5; these show the number of overlapping tendencies. There are three typical features that overlap between English and Norwegian football match reports (the two general features – PP postmodification and part of S/dO/sP – plus arguably more genre-specific verb collocates; see Figures 6 and 7) and three between Norwegian match reports and Norwegian fiction (the two general ones plus HÅP being part of a prepositional complement). There is even more similarity between the use of the nouns in fiction in the two languages, as well as between the genres in English, with four overlapping features each: modification in > 60% and general verb collocates in addition to PP postmodification and part of S/dO/sP for English and Norwegian fiction and modification in > 60%, and negative contexts ≈ 50% in addition to PP postmodification and part of S/dO/sP for English fiction and match reports. Thus, the features that set the genres or languages somewhat apart in the use of the nouns are type of verb collocates, degree of modification, proportion of negative contexts, the use of plural *hopes* and to some extent syntactic function (i.e. the Norwegian noun is more often found as part of a prepositional complement, typically in the sequence *i håp om* ‘in hope about’ ≈ ‘in the hope that’). According to a Log-likelihood test,¹⁵ the difference is statistically significant in the use of plural *hopes* between the two genres in English ($p < 0.0001$ with a high effect size: Odds Ratio=16.25), in the use in negative contexts between the two genres in Norwegian ($p < 0.0001$ with a small effect size: OR=0.18) and in the use of the noun as part of a prepositional complement between English and Norwegian fiction ($p < 0.01$ with a small effect size: OR=2.05).

¹⁴ In terms of dispersion, it should also be noted that most of the features – both for the nouns and verbs (see Table 3 and Figure 8) – are attested in most of the corpus files, albeit with a varying number of occurrences, particularly for features with a low number of attestations overall. A systematic look at dispersion would therefore be welcome in the future, preferably on a larger dataset.

¹⁵ Using the log-likelihood calculator available at <http://ucrel.lancs.ac.uk/llwizard.html>.

Regarding type of verb collocate, the English and Norwegian match reports share the characteristic of making use of verb collocates that are arguably more (football-genre) specific in combination with HOPE (e.g. DASH HOPE, REIGNITE HOPE, QUASH HOPE), while the fiction texts share the feature of making use of more general verb collocates (e.g. HAVE HOPE, BRING HOPE). This difference in verb preferences between the genres becomes apparent in the word clouds in Figures 6 and 7 for English match reports and fiction, respectively.¹⁶



Figure 6. Genre-specific verb collocates (Eng. match reports). **Figure 7.** General verb collocates (Eng. fiction).

Figures 6 and 7 show some overlaps between the most frequent verb collocates in the two genres in English, but, not surprisingly, the verbs are generally different, and, as pointed out above, arguably more genre-specific and action-related in the football match reports. A very similar trend is noted for Norwegian, with verb collocates such as TENNE ≈ ‘ignite/light’, ØYNE ≈ ‘see/nurture’ and SVINNE ≈ ‘vanish’ in the match reports. Examples (7) and (8) serve to illustrate this cross-linguistic tendency of more genre-specific verbs in the match reports.

- (7) ... but another defensive error *killed off* any *hope* of a comeback ... [ENMaRC/AFC]
 (8) Scoringen *tente* et ørlite *håp*. [ENMaRC/STB]
 Lit.: The goal lit a tiny hope

The fiction texts, on the other hand, tend to have more general verb collocates in both languages, with BE/VÆRE and HAVE/HA as the most prominent ones, e.g. examples (9) and (10).

- (9) The poor devil didn't *have* a *hope* in hell. [ENPC+/PeRo1E]
 (10) ... og det *er* vårt *håp* at de beste av våre landsmenn følger vårt eksempel. [ENPC+/BHH1]
 Lit.: and it is our hope that the best of our countrymen will follow our example.

Returning to Figure 5 (and Table 2), we can further note that English HOPE (regardless of genre) occurs in negative contexts in roughly 50% of the cases, as evidenced in both examples (7) and (9), whereas HÅP is less often found in such contexts, particularly in Norwegian fiction with roughly 23% of the cases; neither example (8) nor (10) is deemed negative.

4.2 The verbs HOPE and HÅPE

Following the procedure applied to the nouns in section 4.1, the following contextual features were recorded for the verbs:

¹⁶ The word clouds were generated in WordArt.com on the basis of a list of all verb collocates occurring more than once in the corpora; see the Appendix for the number of actual occurrences in each case and that determine the size of the verbs in the clouds.

- **Verb form:**
 - o Tense, aspect, modality, voice
- **Verb complementation:**
 - o Ø-*that/at* clause | *that/at*-clause | infinitive clause | PP (*for/på*) | intransitive use (no complementation) | prop word | parenthetical use (... , *I hope* | ..., *håper jeg*)
- **Subject:**
 - o Pronoun | NP (including proper nouns)
- **Context** (negative | not negative)

The classification framework is illustrated in examples (11) and (12), where the former is a relatively typical example of English fiction: the personal pronoun *I* is the subject, *hope* is in the present tense, it is followed by a Ø-*that* clause and the context is not negative. The translation into Norwegian is included in example (11) and demonstrates a highly congruent – almost word for word – rendering, to illustrate that this is also typical of Norwegian fiction. Similarly, example (12) is a relatively typical example of the English football match reports, with a full NP as subject, *hoping* is the main verb in a modal perfect progressive verb phrase followed by a Ø-*that* clause and the context is negative.¹⁷

(11) *I hope* everything goes well for you. [ENPC+/AnCI1E]

Jeg *håper* alt går bra for deg. [ENPC+/AnCI1TN]

(12) The Head Coach *would have been hoping* his team could hold out until half-time... [ENMaRC/WFC]

An overview of the distribution of the contextual features in each sub-corpus is given in Table 3, while Figure 8 visually summaries the main tendencies.

Table 3. Main contextual features of the verbs HOPE and HÅPE and their frequency (raw and percentage of total number of occurrences within each sub-corpus).

	English match reports		Norwegian match reports		English fiction		Norwegian fiction	
	Raw	% of total (88)	Raw	% of total (49)	Raw	% of total (367)	Raw	% of total (266)
Simple present tense	3	3.4%	28	57.1%	136	37.1%	126	47.4%
Simple past tense	4	4.5%	7	14.3%	90	24.5%	87	32.7%
Past perfect	7	8%	8	16.3%	19	5.2%	22	8.3%
Non-finite (inf./-ing) clause	32	36.4%	6	12.2%	84	22.9%	30	11.3%
Progressive aspect	34	38.6%	N/A	-	39	10.6%	N/A	-
Ø- <i>that/at</i> clauses	25	28.4%	20	40.8%	216	58.9%	106	39.8%
<i>That/at</i> -clauses	4	4.5%	11	22.4%	30	8.2%	87	32.7%

¹⁷ It is interesting to note that HOPE occurs in the progressive aspect much more frequently in the match reports than in the fiction texts, in 34 out of 88 cases (38.6%) vs. 39 of 367 (10.6%), respectively. Several scholars have pointed to an increased use of the progressive with stative verbs or in combination with modal verbs in recent years (Aarts *et al.*, 2010; Leech *et al.*, 2009). However, it is hard to determine, and also beyond the scope of this study, whether there is a genre or diachronic effect here.

Inf.-clause complementation	31	35.2%	6	12.2%	49	13.4%	10	3.8%
Prep.complementation	24	27.3%	11	22.4%	32	8.7%	24	9%
NP Subject	58	65.9%	3	6.1%	36	9.8%	24	9%
1 st p pl Subject	14	15.9%	23	46.9%	20	5.4%	18	6.8%
1 st and 3 rd p sg Subject	5	5.7%	16	32.6%	277	75.5%	193	72.6%
Negative contexts	18	20.5%	11	22.4%	55	15%	44	16.5%

As was the case in Table 2 for the nouns, the most salient contextual features in the sub-corpora (percentages) have also been highlighted in bold in Table 3 for the verbs and are included in Figure 8.

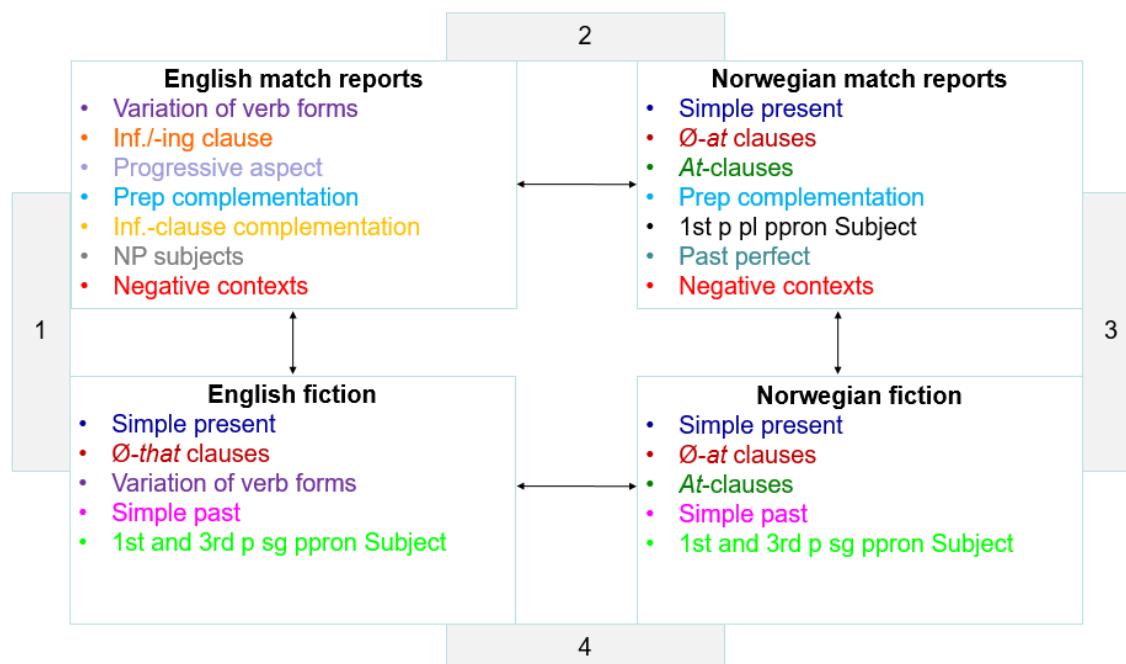


Figure 8. Contextual features of the verbs HOPE and HÅPE: Main tendencies.

Figure 8 shows that there is less overlap of typical features between the sub-corpora here than was the case for the nouns (see Figure 5), notably with only one overlapping feature for English fiction and match reports and only two for English and Norwegian match reports. There is most similarity between the fiction texts in English and Norwegian, which are characterised by *Ø-that/at* clauses, a combination of the simple present and past tense, and 1st and 3rd person singular pronouns as Subjects. The Norwegian texts also show a fair amount of overlap across the two genres, whereas there seems to be more of a text type effect in English. There are few features that are typical of both English match reports and English fiction.¹⁸ In fact, they only share one of the characteristic features that can be gleaned from Table 3 for the verb HOPE, namely variation of verb forms. By variation is here understood a combination of tense, modality and aspect, as well as the use of non-finite forms. A couple of examples are given in (13)–(15), with a modal perfect, a present perfect progressive and a non-finite *-ing*, respectively.

¹⁸ Statistically significant differences were recorded for the following features, according to a LL test: simple present tense ($p < 0.0001$; OR=0.03), simple past tense ($p < 0.0001$; OR=0.06), non-finite clauses ($p < 0.0001$; OR=0.52), *Ø-that* clauses ($p < 0.0001$; OR=0.16) and negative contexts ($p < 0.01$; OR=0.44). However, it is important to note that, in some cases, these tests are based on very small numbers.

- (13) ... and was unable to execute his volley in the way he *would have hoped* ...
[ENMaRC/CFC]
- (14) *Have you been hoping* for something more exciting? [ENPC+/ABR1]
- (15) Bradley sent his team out for the restart *hoping* they could find a way of causing more problems for the Watford defence. [ENMaRC/SCAFC]

Figure 8 also reveals a low degree of overlap between English and Norwegian match reports,¹⁹ as they only share two typical contextual features for HOPE and HÅPE, namely prepositional complement and negative context, both of which are captured in example (16).

- (16) Det ble ikke den festkvelden vi hadde *håpet på* Aker Stadion søndag kveld.
[ENMaRC/MFK]
Lit.: It did not become the night of celebration we had hoped for at Aker Stadion
Sunday night

With reference to example (16) it is interesting to note that while genre seems to be a decisive factor for the verb to (more typically) be used in negative contexts, language was a decisive factor for the nouns, where the English match reports and English fiction texts were seen to be more in agreement regarding this feature.

5. Discussion

On the basis of preliminary observations of the data it was suggested in Section 1 that the match reports would make use of a narrower selection of contexts in which the ‘hope’ lemmas are used (cf. Stubbs 1996: 89). The case studies presented in Sections 4.1 and 4.2 do not seem to substantiate this hypothesis. In fact, from the potential uses attested in the corpora, the match reports feature a broader repertoire of typical phraseological contexts compared to fiction. Thus, instead of featuring in a narrower selection of contexts from the pool of potential uses, the lemmas are rather shown to typically feature in a different selection of contexts in the match reports. In terms of number of characteristic phraseologies recorded for both the nouns and verbs, genre seems to play a slightly more important role than language.

It could be argued that the contextual features recorded for the nouns and verbs are relatively straightforward to determine, perhaps with one exception: ‘context’. A binary distinction of negative|not negative was applied to extended contexts in which HOPE and HÅP(E) occurred. One potential challenge, also referred to above, was how to operationalise this in the analysis of individual instances, as HOPE and HÅP(E) are arguably reserved for positive connotations, indicating that the negative flavour with which these items are sometimes imbued seems almost contradictory. Example (17) is a case in point, where *surge of hope* carries positive expectations that are later shown not to be fulfilled, when it turns out that it was not Emma who called, but Andrew. It may be speculated that this is a deliberate choice on the part of the writer to create an effect, i.e. an element of surprise, or as Louw (1993: 30) puts it “irony in the text or insincerity in the writer”, with reference to the concept of semantic prosody (see further below).

¹⁹ Syntactic differences between the two languages can be seen to account for some of this, as Norwegian does not have forms corresponding to non-finite *-ing* clauses and a grammaticalised progressive aspect. Statistically significant differences can be noted for the following features: simple present tense (p<0.0001; OR= 0.03), past perfect (p<0.05; OR=0.28), *Ø-at/that* clause (p<0.01; OR=0.40), *at/that*-clause (p<0.0001; OR=0.12), NP Subjects (p<0.0001; OR=6.13) and 1st p. pl subject (p<0.0001; OR=0.19). However, it is important to stress that in some cases the number of occurrences are few here and we should perhaps not put too much weight on these tests.

- (17) He had a *surge of hope* that it was Emma, until he picked it up and heard Andrew babbling excitedly... [ENPC+/MiWa1E]

Sometimes these contextual clues lie outside the scope covered by the default length of a concordance line, and a wider context has to be examined. It is also tempting to suggest that the past tense can be seen as a trigger for hopes being shattered. However, the evidence for this is inconclusive, as the past tense is also regularly used in non-negative contexts – either neutral or positive, as in (18) – where there is nothing in the surrounding context to suggest that the existence of hope came to an end.

- (18) There was *hope* everywhere. [ENPC+/JSM1]

The contextual features recorded for the nouns and verbs bear a strong resemblance to features covered by the categories that are part of Sinclair’s (1996, 1998) Extended Units of Meaning model, viz. collocation, colligation, semantic preference and semantic prosody. And although the material at hand does not uncover strong unanimous lexico-grammatical patterns for neither the nouns nor the verbs that would suggest that HOPE and HÅP(E) clearly function as cores of extended units of meaning, it does reveal tendencies regarding semantic prosody, which is the only obligatory element in the model apart from the core (Sinclair, 1998: 20; Ebeling and Ebeling, 2013: 58). Traditionally semantic prosody refers to semantic colouring through surrounding context and may contribute to a positive or negative reading of words that are in themselves neutral. Louw (1993: 157) defines semantic prosody as “a consistent aura of meaning with which a form is imbued by its collocates”. Put differently, and according to Louw and Milojkovic (2016) collocates contribute to a “context of situation revealing attitude (semantic prosody)” (Louw and Milojkovic, 2016: 54). In the literature, it has been argued that such an attitude may be binary, i.e. positive vs. negative, or non-binary (and more specific), expressing e.g. ‘difficulty’ in the case of *the naked eye* (Sinclair 1996: 33) and ‘occupation’ in case of *train as a* (Hoey, 1997: 5). For the purpose of this study, the binary opposition negative vs. non-negative has been applied.²⁰ In the context of the current investigation it is also important to mention that several scholars have observed that semantic prosody may be both language-specific (e.g. Stewart, 2009: 32)²¹ and register-specific (Xiao and McEnery, 2006: 114ff; Hunston, 2007: 263ff).

For the items under investigation here, there seems to be (at least) two things at play regarding semantic prosody: language and genre. In the case of the nouns, the strongest indication of a negative-like prosody is found in English, regardless of genre (see Table 2 and Figure 5). For the verbs, on the other hand, it is the match reports that show the strongest tendency towards a negative prosody, regardless of language (see Table 3 and Figure 7). The verbs in fiction do not seem to take on a particular prosody at all, as most instances seem to contain a neutral use of HOPE and HÅPE, as in examples (19) and (20).

- (19) Though what you *hope* to find there I have no idea. [ENPC+/PeRo2E]

- (20) Fortsatt *håpet* jeg på Kari Thue. [ENPC+/AnHo2N]
Lit.: Still I hoped for Kari Thue

The Norwegian noun behaves differently from the English noun in being less consistently used in negative contexts. However, it is clear that HÅP is closer to an established negative prosody in the match reports (occurring in negative contexts in ca. 40% of the cases), e.g. example (21),

²⁰ It is important to note that the validity of semantic prosody as a concept has been questioned over the years, but it would take us too far afield to go into this discussion here. But see e.g. Whitsitt (2005), Hunston (2007), Morley and Partington (2009) and Stewart (2010) for some (critical) discussions.

²¹ In the case of mismatched prosodies across languages, see also Partington (1998), Tognini-Bonelli (2002) and Ebeling (2014), and references therein.

than it is in the fiction texts, where it occurs in negative contexts in roughly 23% of the cases (see Table 3); this difference was shown to be statistically significant, albeit with a small effect size.

(21) Alt *håp* om poeng ser nå ut til å være over. [ENMaRC/OBK]

Lit.: All hope of points now looks to be over

In a few instances HÅP is used in contexts with a positive outcome, as example (22) arguably illustrates – a hope has been restored after it had been dashed –, but it is by far most commonly used in more neutral contexts, expressing a hope with expectations of a positive outcome, but where the outcome is in fact unknown, as in (23).

(22) Likevel hadde jeg det bedre en stund. Både fordi Henrik hadde gitt meg tilbake et *håp* jeg ikke lenger hadde ... [ENPC+/MN1]

Still I felt better for a while. Both because Henrik had given me back a hope I no longer had ... [ENPC+/MN1T]

(23) For det var det eneste svaret som ga noe *håp*. [ENPC+/JoNe1N]

Because it was the only answer that gave any hope. [ENPC+/JoNe1TE]

Even in contexts in which the immediate collocates are of a positive nature, as in (24) where the positive adjective *godt* ‘good’ premodifies *håp*, the outcome is not specified as positive in the surrounding context.

(24) “Jeg tror jeg verken skal bekrefte eller avkrefte annet enn at vi i Kripos har godt *håp* om at denne saken går mot en snarlig oppklaring.” [ENPC+/JoNe2N]

“I don’t think I have to confirm or deny anything except that we at Kripos are fairly confident [Lit.: ... Kripos have a good hope ...] that we will soon have this case solved.” [ENPC+/JoNe2TE]

Returning to the starting point of this study, and to the question of whether the prominent use of HOPE/HÅP(E) in negative contexts in match reports (Ebeling, 2019) reflects a true tendency of this genre in both languages and whether such use extends to other genres, we can conclude that the investigation uncovers some conflicting evidence in this respect. Both genre and language seem to have an impact, thus lending some support to the observations referred to above, namely that semantic prosody may be both language- and register-specific. However, it is also interesting to note that not only may semantic prosody depend on language and genre/register, it may also be dependent on word class. This is in accordance with previous studies that have shown similar trends, e.g. Stubbs (1995) in the case of the noun and verb CAUSE.²²

In summary, then, the attraction to negative contexts seems to be language-specific for the nouns HOPE/HÅP, whereas it seems to be genre-specific for the verbs HOPE/HÅPE.

6. Concluding remarks and suggestions for further study

In addition to investigating the potentially negative bias of HOPE/HÅP(E), this study set out to answer a set of research questions regarding the use of these cognates in two languages and two genres on the basis of two different kinds of contrastive corpora. To answer the specifically cross-linguistic question – to what extent are the lemmas used similarly in English and

²² The slight difference noted in the semantic prosody between the verb and noun CAUSE is very much tied to one specific inherent meaning of the noun, namely the ‘aim/principle’ reading, as in *The only cause they had in common was a refusal to eat meat.* (ENPC/PDJ3). The importance of taking separate meanings into account when investigating semantic prosody has been addressed in a recent master’s thesis by Russnes (2020).

Norwegian? – the investigation reveals that, in very general terms, the lemmas show similar potential of use but have slightly different preferred uses. An almost identical answer is suggested for the specifically cross-genre question – to what extent are the lemmas used similarly in match reports and fiction? In this case the lemmas have a relatively similar potential of use, but with different preferences in the two genres studied.

In more specific terms, the nouns are shown to be more similar across the genres in English with more overlapping contextual features than in Norwegian. In fact, Norwegian fiction is more similar to English fiction than to Norwegian match reports (see Table 2 and Figure 5). The analysis of salient contextual features further suggests that the Norwegian match reports seem to have adopted a slightly different set of salient features compared to the other sub-corpora. As far as the verbs are concerned, their use is more similar across the languages than across the genres, (particularly in fiction). Moreover, Norwegian fiction and match reports are more similar than English fiction and match reports (see Table 3 and Figure 8). Thus, in this case, it is the English match reports that seem to have adopted the most special uses. A final and general observation from the case studies is that, overall, the use of these lemmas is most similar in English and Norwegian fiction. At this stage we may only speculate as to the reason for this, but could it be that we are dealing with an established (fiction) genre versus an emerging and less established genre of online match reports, which, as a result, produces more variation?

Regarding the third research question – to what extent does the use of different types of contrastive corpora contribute to our cross-linguistic knowledge of the lemmas? – it has been shown that the combination of comparable data in two genres and translation data in one genre has:

- Provided a firm basis for a contrastive analysis of the items compared;
- Highlighted language similarities/differences/preferences;
- Highlighted cross-linguistic genre similarities/differences/preferences.

These points may form the basis of a slightly modified version of Aijmer and Altenberg's (1996) frequently quoted words on the usefulness of parallel corpora, thus:

The use of different contrastive corpora has given new insights into the languages and genres compared – insights that would have gone unnoticed in a study of only one of the corpora alone, i.e. either the ENPC+ or the ENMaRC.²³

The advantages of drawing on several primary contrastive sources are evident, and in particular a combination of both bidirectional translation data and monolingual comparable data between two or more languages (representing different genres) has been shown to be fruitful. Such corpora have previously been shown to complement each other in the sense that bidirectional translation corpora (i.e. parallel corpora) arguably provide the researcher with a more objective *tertium comparationis*, while comparable corpora are indispensable in a contrastive study in order to provide data sets that are both more extensive and that include text types that are not typically translated between languages (cf. Johansson, 2007; Ebeling and Ebeling, 2020).

Against this backdrop, it is important to stress that the potential of neither corpus type has been exploited to the full in the current paper. Further insights could be gained on the basis

²³ The original Aijmer and Altenberg quote reads as follows: bilingual corpora “give new insights into the languages compared – insights that are likely to be unnoticed in studies of monolingual corpora” (Aijmer and Altenberg, 1996: 12).

of the ENPC+ in a systematic investigation of correspondences of noun and verb patterns in a contrastive perspective (only fiction), for example:

- Analyse the translation paradigms of the English verb pattern *hope* + \emptyset -*that* clauses, on the basis of examples such as (25);

- o \emptyset -*that* clause \rightarrow *at*-clause, even if \emptyset -*at* clauses are possible in Norwegian

(25) He *hopes* [\emptyset] they'll have enough warm clothes to last the coming winter.
[ENPC+/StGa1E]

Han *håper at* de har nok varmt tøy til å klare seg gjennom vinteren som står for døren.
[ENPC+/StGa1TN]

Lit.: He hopes that they ...

- Analyse the translation paradigms of the Norwegian noun pattern PREP *håp* PP + *at* clause, on the basis of examples such as (26).

- o PREP *håp* PP + *at* clause \rightarrow non-finite *-ing* clause + \emptyset -*that* clause, where a syntactically similar pattern is ruled out in English

(26) ... *i håp om at* den andre skulle si navnet sitt; [ENPC+/EFH1]

Lit.: in hope about that ...

... *hoping* [\emptyset] the other man would say his name. [ENPC+/EFH1T]

Furthermore, the advantage of size that is often attributed to comparable monolingual corpora when compared to parallel corpora is not present here, with the ENMaRC being smaller than the ENPC+. However, this may be amended in the future, as more match reports can be added and more sizeable comparable fiction data can be culled from larger monolingual corpora in the two languages.

Finally, it should be noted that there are some challenges involved when trying to carry out a comprehensive case study on the basis of material from a variety of sources. In particular, it is challenging to organise and analyse data from multiple languages and genres in a clear and consistent manner. Nevertheless, rather than shy away from the relatively complex nature of such data, researchers should perhaps complement traditional contrastive analysis techniques with more sophisticated models for handling complex data, as it would give us the opportunity to gain even more rewarding insights into cross-linguistic, cross-genre uses of language.

References

- Aarts, B., Close, J. and Wallis, S. 2010. Recent Changes in the Use of the Progressive Construction in English. In *Distinctions in English Grammar*, B. Capelle and N. Wada (eds), 148–168. Kaitakusha: Tokyo, Japan.
- Aijmer, K. and Altenberg, B. 1996. Introduction. In *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies. Lund 4–5 March 1994*, K. Aijmer, B. Altenberg and M. Johansson (eds), 11–16. Lund: Lund University Press.
- Altenberg, B. 1999. Adverbial Connectors in English and Swedish: Semantic and Lexical Correspondences. In *Out of Corpora. Studies in Honour of Stig Johansson*, H. Hasselgård and S. Oksefjell (eds), 249–268. Amsterdam: Rodopi.
- Anthony, L. 2019. AntConc (version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available from <http://www.laurenceanthony.net/software> [Last accessed 5 May 2020].

- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.
- Chesterman, A. 1998. *Contrastive Functional Analysis*. Amsterdam: Benjamins.
- Det Norske Akademis ordbok (Norwegian Academy Dictionary). 2020. Oslo: Det Norske Akademi for Språk og Litteratur. <https://www.naob.no/ordbok> [Last accessed 17 August 2020].
- Dupont, M. and Zufferey, S. 2017. Methodological Issues in the Use of Directional Parallel Corpora. *International Journal of Corpus Linguistics* 22(2), 270–297.
- Ebeling, J. and Ebeling, S.O. 2013. *Patterns in Contrast*. Amsterdam: Benjamins.
- Ebeling, S.O. 2014. Cross-linguistic Semantic Prosody: The Case of *Commit, Signs of* and *Utterly* and their Norwegian Correspondences. In *Corpus-based Studies in Contrastive Linguistics, Oslo Studies in Language* 6(1), 2014, S.O. Ebeling, A. Grønn, K.R. Hauge and D. Santos (eds), 161–179.
- Ebeling, S.O. 2019. The Language of Football Match Reports in a Contrastive Perspective. In *Corpus Approaches to the Language of Sports. Text, Media, Modalities*, M. Callies and M. Levin (eds), 37–62. London: Bloomsbury Academic.
- Ebeling, S.O. 2021. Minutes of Action! A Contrastive Analysis of Time Expressions in English and Norwegian Football match reports. To appear in *Time in Languages, Languages in Time*, A. Čermáková, T. Egan, H. Hasselgård and S. Rørvik (eds). Amsterdam: Benjamins.
- Ebeling, S.O. Forthcoming. The Function of Recurrent Word-combinations in English Translations from Three Different Languages. To appear in a special issue of *Meta: Translators' Journal*, ed. By C. Ji and M. Oakey.
- Ebeling, S.O. and Ebeling, J. 2017. A Functional Comparison of Recurrent Word Combinations in English Original vs. Translated Texts." *ICAME Journal* 41, 31–52.
- Ebeling, S.O. and Ebeling, J. 2020. Contrastive analysis, *tertium comparationis* and corpora." *Nordic Journal of English Studies* 19(1), 97–117.
- Hoey, M. 1997. From Concordance to Text Structures: New Uses for Computer Corpora. In *Practical Applications in Language Corpora*, B. Lewandowska-Tomaszyk and P.J. Melia (eds), 2–23. Łódź: Łódź University Press.
- Hunston, S. 2007. Semantic Prosody Revisited. *International Journal of Corpus Linguistics* 12(2), 249–268.
- Johansson, S. 2007. *Seeing through Multilingual Corpora: On the Use of Corpora in Contrastive Studies*. Amsterdam: Benjamins.
- Johansson, S. and Hofland, K. 1994. Towards an English-Norwegian Parallel Corpus. *Creating and Using English Language Corpora: Papers from the Fourteenth International Conference on English Language Research on Computerized Corpora, Zurich 1993*, In U. Fries, G. Tottie and P. Schneider (eds), 25–37. Amsterdam: Rodopi.
- Leech, G., Hundt, M., Mair, C. and Smith, N. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: CUP.
- Lefer, M-A. and Voegeler, S. (eds). 2014. Genre- and Register-related Discourse Features in Contrast. Special issue of *Languages in Contrast* 14(1).
- Louw, B. 1993. Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies." In *Text and Technology: In Honour of John Sinclair*, M. Baker, G. Francis and E. Tognini-Bonelli (eds.), 157–175. Amsterdam: Benjamins.
- Louw, B. and Milojkovic, M. 2016. *Corpus Stylistics as Contextual Prosodic Theory and Subtext*. Amsterdam: John Benjamins.
- Morley, J. and Partington, A. 2009. A few Frequently Asked Questions about Semantic – or Evaluative – Prosody." *International Journal of Corpus Linguistics* 14(2). 139–158.
- Neumann, S. 2014. *Contrastive Register Variation. A Quantitative Approach to the Comparison of English and German*. Berlin: De Gruyter Mouton.
- Nordrum, L. 2007. *English Lexical Nominalizations in a Norwegian-Swedish Contrastive Perspective*. Doctoral dissertation, English department, Göteborg University. Available at https://gupea.ub.gu.se/bitstream/2077/17181/5/gupea_2077_17181_5.pdf [Last accessed 17 August 2020].
- Partington, A. 1998. *Patterns and Meaning. Using Corpora for English Language Research and Teaching*. Amsterdam: Benjamins.

- Oxford English Dictionary* (OED) Online. 2020. Oxford: Oxford University Press <<http://oed.com/>> [Last accessed 17 August 2020].
- R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/> [Last accessed 17 August 2020].
- Russnes, M.S. 2020. *Committed to Crime – A Corpus-based Study of the Semantic Prosodies of Separate Meanings within Lexical Items*. Unpublished MA thesis, University of Oslo.
- Sinclair, J. 1996. The Search for Units of Meaning. *Textus IX*. 75–106.
- Sinclair, J. 1998. The Lexical Item. In *Contrastive Lexical Semantics*, E. Weigand (ed.), 1–24. Amsterdam: John Benjamins.
- Stewart, D. 2010. *Semantic Prosody. A Critical Evaluation*. New York / London: Routledge.
- Stubbs, M. 1995. Collocations and Semantic Profiles. On the Cause of Trouble with Quantitative Studies. *Functions of Language*, 2:1, 23–55.
- Stubbs, M. 1996. *Text and Corpus Analysis*. Oxford: Blackwell.
- Teich, E. 2003. *Cross-Linguistic Variation in System and Text. A Methodology for the Investigation of Translation and Comparable Texts*. Berlin: Mouton de Gruyter.
- Tognini-Bonelli, E. 2002. Functionally Complete Units of Meaning across English and Italian: Towards a Corpus-driven Approach.” In *Lexis in Contrast. Corpus-based Approaches*, B. Altenberg and S. Granger (eds), 73–95. Amsterdam: John Benjamins.
- Whitsitt, S. 2005. A Critique of the Concept of Semantic Prosody. *International Journal of Corpus Linguistics* 10/3. 283–305.
- Xiao, R. and McEnery, T. 2006. Collocation, Semantic Prosody, and Near Synonymy. A Cross-linguistic Perspective. *Applied Linguistics* 27 (1): 103–129.

Appendix

Table A. Number of occurrences of verb collocates with a frequency of more than one in the English match reports and fiction texts.

English match reports		English fiction	
Verb	Raw freq.	Verb	Raw freq.
BE	10	BE	12
BOOST	4	BRING	2
DASH	17	GIVE	3
DEAL	3	GIVE UP	4
END	4	HAVE	10
EXTINGUISH	6	HOLD OUT	5
GIVE	43	MAKE	2
GO	5	RETURN	2
HANG	2		
HAVE	2		
HIT	2		
KEEP	15		
KILL OFF	8		
OFFER	5		
PROVIDE	2		
QUASH	2		
RAISE	5		
REIGNITE	3		
SUFFER	4		
TAKE	3		
THWART	2		

Signe Oksefjell Ebeling

Author's address

Signe Oksefjell Ebeling
Department of Literature, Area Studies and European Languages
University of Oslo
P.O. box 1003, Blindern
NO-0315 Oslo
Norway
s.o.ebeling@ilos.uio.no