Contrasting translations

Thomas Egan, Hedmark University of Applied Sciences

Abstract
This paper argues that corpora containing translations of the same source text into two or more target languages represent a sounder base for comparing linguistic expressions in several languages than do corpora containing data from just two languages, whether they be unidirectional or bidirectional. The argument is illustrated with comparisons of some prepositional constructions in English and French using data from the Nor. – Eng. – Fr. – Ger. part of the Oslo Multilingual Corpus, and some constructions in Swedish and Norwegian using the overlapping original texts in the ENPC and ESPC.

1. Introduction

In the last twenty years, since the widespread availability of translation and parallel corpora, much productive work has been done comparing translations with original texts, both the original texts underlying the translations themselves, and other texts in the target language which are employed to check the extent to which the translations are representative of usage in that language. Indeed there have been over 200 such publications based on the corpora whose anniversary is being celebrated in this special edition. A good deal less work has been done on comparing translations into different languages of one and the same set of source texts. Some examples of the latter approach are Johansson (2001), Paulussen (1999), Slobin (2005), Verkerk (2014, 2015) and Viberg (1998, 2003, 2013).

In this paper I explore the advantages of comparing translations into different languages of the same source texts, rather than original texts and their translations, illustrating my argument with comparisons of some constructions in English and French with data from the Norwegian – English – French – German part of the Oslo Multilingual Corpus (see Johansson 2007, Egan 2013a, 2014, 2015a, 2015b)),[1] and some

---

[1] For the sake of convenience the abbreviation OMC will hereafter be used for the Norwegian – English – French – German part of the Oslo Multilingual Corpus (<https://www.hf.uio.no/ilos/english/services/omc/>) even though, strictly speaking, the term OMC denotes a whole family of corpora.

constructions in Swedish and Norwegian using the overlapping original texts in the English-Norwegian Parallel Corpus (ENPC) and the English-Swedish Parallel Corpus (ESPC) (see Hasselgård 2007, Rawoens & Egan 2015).

In section 2 I describe what I see as the weaknesses inherent in using original texts and their translations to compare expressions in two languages. Section 3 contains a comparison of the use of visual and verbal prompts in the elicitation of predications of motion for the purpose of contrasting these in two or more languages. In section 4 I present some sorts of predications which cannot be elicited using visual prompts. Section 5 contains details of the overlapping texts in the ENPC and ESPC and argues that these texts comprise a viable corpus for contrastive studies of expressions in Norwegian and Swedish. Finally, section 6 contains a summary and conclusion.

2. Comparing original texts and translations

When comparing translations with original texts, one is faced with the problem that these are different text types, one of which may display translation effects, the other not. By translation effects (Johansson 1998: 5), or translationese (Gellerstam 1996: 54), are meant the retention in the target language texts of features of the source language that are not equally felicitous in the target. So prevalent are these features that, as has been demonstrated by Cappelle (2012), it is sometimes possible to use them to predict the original language of a translated text. This problem was recognised by the compilers of corpora such as the ENPC and ESPC and was addressed in them by including texts and translations in both directions, thus enabling the comparison of original and translated language in the two languages being compared (see Johansson 2007).

Any comparison between two or more features is dependent on the availability of a viable *tertium comparationis*. According to Johansson (2001: 584), "The advantage of a corpus of original texts and their translations is that the translation is intended to express the same meaning as the original text". However, Altenberg & Granger, writing just a year after Johansson, maintain that:

> It is not realistic to proceed from a *tertium comparationis* that is based on 'identity of meaning'. For one thing, this would be putting the cart before the horse and we would run the risk of methodological circularity: the result of the contrastive

analysis would be no more than the initial assumption (cf. Krzeszowski 1990: 20). For another, the area we want to explore is often fuzzy and impossible to define satisfactorily. (Altenberg & Granger 2002: 16)

Krzeszowski' s original formulation of the first objection reads:

> We compare in order to see what is similar and what is different in the compared materials; we can only compare items that are in some respects similar, but we cannot use similarity as an independent criterion in deciding how to match items for comparison since similarity (or difference) is to result from the comparison and not to motivate it. (Krzeszowski 1990: 20)

However, one could argue that this objection is not equally trenchant if the similarity that is originally postulated ('in some respects similar') is broader, more coarse-grained than the similarity/difference that is the goal of the study. The analysis would then resemble a spiral more than a circle, to pursue the geometric metaphor, and although we follow a curved path in negotiating a spiral, we do not arrive back at our original destination.

The second objection raised by Altenberg and Granger carries, I think, more weight. It is often difficult to formulate even a satisfactory coarse-grained definition of the research object. Ebeling & Ebeling (2013: 21) make the same point, contending that "One of the difficulties in starting with meaning is how to delimit it. Starting with form, the boundaries are already set, while meaning is much more elastic."

Related to the difficulty of pinning down meaning in two texts is a more basic problem involved in comparing originals and translations, which is that these two text types are produced subject to two different sorts of constraints. As researchers we only have access to one of these. That is, we know what the translator is trying to convey, but can only guess at the intentions of the author of the original text. The discrepancy is illustrated in Table 1, where the term '2-text corpora', borrowed from Krzeszowski (1990), is used for translation corpora containing texts in two languages.

Table 1. Sources and targets in 2-text translation corpora

	To be encoded	Encoded by
Translator	Expression in source text	Expression in target text
Original author	?	Expression in source text

We see in Table 1 that the expression in the source text occurs in two columns, in the second column as a prompt for the expression in the translated text and in the third as an utterance to be compared to the latter. The prompt in the third row, the content of which is represented by the question mark, is nebulous compared to its counterpart in the second row. Since the production of a meaningful utterance involves making a series of lexical and grammatical choices, it is vital for the analyst to be aware of the parameters within which these choices are made. However, in the case of the original author, as opposed to the translator, the analyst is in the dark as to the exact nature of the prompts in question, having to reason backwards from their expression in the third column.

The fact that 2-text translation corpora are less than ideal for the comparison of expressions in two languages should not be taken to mean that they cannot be mined for data about one of these languages: quite the contrary, in fact. A series of expressions in the target language may be used to shed light on the polysemanticity of a single item in the source language, for instance, as illustrated in Table 2.

Table 2. Translations in 2-text translation corpora of the same expression in the source language.

	To be encoded	Encoded by
Translator	Expression (a1) in source text	Expression (x) in target text
Translator	Expression (a2) in source text	Expression (x) in target text
Translator	Expression (a3) in source text	Expression (y) in target text
Translator	Expression (a4) in source text	Expression (z) in target text

Table 2 illustrates various translations of one and the same expression in one or more source texts, by one or more translators. It shows that two tokens of expression (a) are translated by expression (x), one by (y) and one by (z). With respect to translations of prepositions, the part of speech which I will be drawing on for the exemplification of my arguments in sections 3 – 5, Garretson (2004) maintained that "… if we take as our default assumption that similar forms will be used to translate similar meanings, we must expect that related meanings will be expressed with the same form more often than unrelated meanings will" (Garretson 2004: 23). The method of analysis illustrated in Table 2 has been employed in studies based on the ENPC and ESPC to throw light on a large variety of linguistic items (for just one of many possible examples, see Aijmer, this volume). The greater the number of source texts, and the

greater the number of translators, the more light the analyst can shed on the expression in the original language as reflected in the various translations. However, if one is interested in comparing expressions in two or more languages, rather than in using translations as a means to explore the polysemanticity of items in a source language, 2-text corpora cannot furnish us with optimal *tertia comparationis*, for reasons outlined above. In the next section I present a couple of alternative sources of data and argue that these do not suffer from the same shortcomings.

3. Contrastive research into spatial relations

Given the problems inherent in the comparison of an expression in an original text with one in a translated text described in the previous section, one might ask whether there are any alternative *tertia comparationis* that the researcher may have recourse to. One way of ensuring that the linguistic items being compared are produced under similar constraints is to provide informants with a *tertium comparationis* from another modality. In research into predications of location and motion, this *tertium comparationis* may take the form of drawings, picture books or video snippets (see, for instance, Engemann *et al*. 2012, Vulchanova *et al*. 2012). The most extensively used text in the elicitation of descriptions of motion in various languages is the *Frog where are you?* story (Berman & Slobin 1994). Figure 1 contains an illustration from this story showing the frog climbing out of its jar.

Figure 1. A picture from the *Frog where are you?* story (Mayer 1969)

Table 3 summarises the process and results of an elicitation experiment, such as that carried out by researchers using visual prompts.

Table 3. Sources and targets in picture-based elicitation experiments

	To be encoded	Encoded by
Informant 1	Picture (a) in source text	Expression in first target text
Informant 2	Picture (a) in source text	Expression in second target text

Comparing Table 3 to Table 1, we see that in Table 3 the constraints on the informants are identical, insofar as both are responding to the exact same stimulus. Moreover, this is a stimulus to which researchers have full access.

Instead of using pictures, still or moving, as prompts, we can use verbal texts. A set of source texts and two sets of target texts make up

what we may call a '3-text corpus' (see Egan 2013a). Informants are still provided with identical prompts, but the *tertia comparationis* are verbal rather than visual. It may be objected that the informants, in this case translators, are subject to more constraints than observers of a picture, insofar as they may not be at liberty to construe the situations or events in the original texts as freely as the latter. They are, however, at liberty to *re*-construe them, should they feel the wish or need to do so. This point may be exemplified by (1) and (2), which consist of [EXIT] predications, similar to that illustrated in Figure 1, produced by English and French translators in response to the Norwegian originals in the OMC (see Egan & Graedler 2015).

(1)   a. Jeg åler meg *ut av* vinduet igjen. (NF1)[2]
        b. Wriggling *through* the window … (NF1TE)
        c. Je me suis glissé à nouveau *par* la fenêtre. (NF1TF)

(2)   a. Hun holdt hesten an da hun var kommet *ut av* den siste kløfta. (HW2)
        b. When she rode *out of* the last crevice, she reined in her horse. (HW2TE)
        c. Elle retint le cheval après avoir *passé* le dernier ravin. (HW2TF)

In (1) both the English and the French translator code the manner of motion in the verb and the path in an adverbial, thus preserving the coding options of the original text. In (2) on the other hand, in which the original text contains a neutral verb of motion and a path adverbial, the English translator employs a manner motion verb and a path adverbial and the French translator the path verb *passer* 'pass'. Note that the inclusion of the Norwegian originals in the examples is not for the purpose of comparing them to the translations, but rather to illustrate the common prompts to which the translators are exposed. Table 4, which summarises the process and results of the translation endeavours, may be compared to Table 3, which it closely resembles, only differing in the nature of the prompts.

---

[2] The letters and number 'NF1' refer to the source text in the Norwegian – English – French – German sub-part of the Oslo Multilingual Corpus, 'NF' being the initials of the author. 'TE' and 'TF' stand for English and French translated text, respectively.

Table 4. Sources and targets in 3-text translation corpora

	To be encoded	**Encoded by**
Translator 1	Expression (a) in source text	Expression in first target text
Translator 2	Expression (a) in source text	Expression in second target text

There is, in fact, one other difference between the processes illustrated in Tables 3 and 4, which may not be immediately obvious. This is the fact that, unlike informants in visual experiments, translators represented in corpora such as the OMC are not participating in an experiment. There is therefore no danger of the results being contaminated by the observer's paradox (Labov 1972).

4. Advantages of using verbal as opposed to visual prompts

While the use of visual prompts in the elicitation of utterances about motion described in section 3 has led to a great increase in our knowledge of how these predications are coded in various languages, they are unfortunately only suited for the investigation of concrete relations, such as location and motion. For more abstract relations, we must turn to verbal prompts (see Egan 2015a). Such relations are moreover very common, as is shown in Table 5, where spatial predications account for just under 50% of the tokens of the Norwegian preposition *gjennom*, the default English translation of which is 'through' (see Egan 2014). To investigate the remaining tokens, we cannot have recourse to visual *tertia comparationis*.

Table 5. Types of 'throughness' encoded by *gjennom* in the ENPC (Egan 2014)

Semantic field	Number of tokens
Motion	146
Perception	62
Other (metaphorical)	46
Time	24
Idiom	13
Medium	8
Means	8
Location	6
Total	313

Leaving aside predications of motion and location, the three most common types of relationship coded by *through* are perceptual, temporal and other (metaphorical). This last category is a very heterogeneous one, consisting of tokens displaying a wide range of metaphorical mappings. The perceptual and temporal categories are more tight-knit, displaying less semantic variation in the source texts. Figures 2 and 3 illustrate the correspondences between English and French translation correspondences of *gjennom* in these two categories, divided in English into translations containing 'through', translations containing other prepositions and translations containing non-prepositional constructions, and in French translations containing the proposition 'à travers', translations containing other prepositions, translations containing the verb 'traverser' and translations containing other non-prepositional constructions.[3]
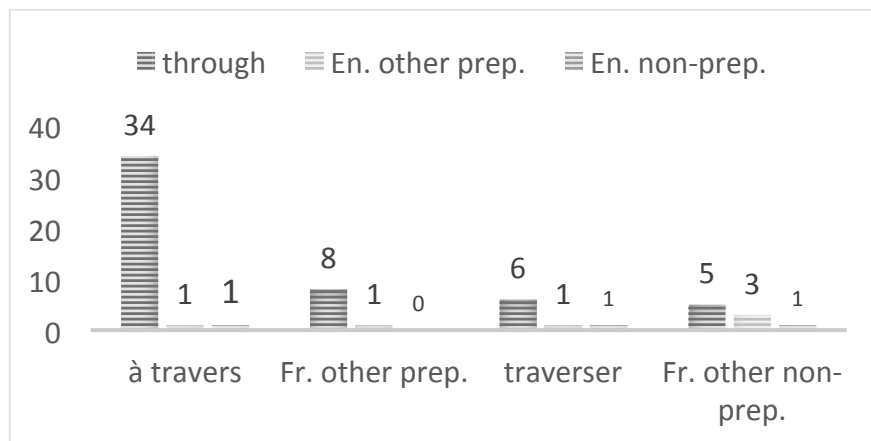

Figure 2. Correspondences between English and French codings of Perception [throughness] (Egan 2014)

In Figure 2, we can see that of a total of 62 tokens, 53 are translated into English by *through* and 36 into French by *à travers*. 34 of the latter

---

[3] The reason why the tokens are divided into four categories in French as opposed to just three in English, is the incidence in French of a large number of tokens containing the verb *traverser*. In a bottom-up investigation, such as this one, the categories must be allowed to emerge from the data, rather than being superimposed on the data in advance of the analysis.

correspond to English *through*. In one case a French preposition other than *through* corresponds to a similar English preposition and in three cases both sets of translators employ a non-prepositional construction. Thus the two sets of translators employ similar strategies in 38 cases (61%). Moreover *through* and *à travers* correspond to one another in codings of [throughness] involving various perceptual senses, such as vision in (3), hearing in (4) and smell in (5).

(3)   a. *Gjennom løvverket* skimtet jeg mesterens koie. (NF1)
        b. There was my master's hut visible *through the canopy of leaves*. (NF1TE)
        c. J'entrevois la baraque de Léopold *à travers le feuillage*. (NF1TF)

(4)   a. Måkene skrek inn til dem *gjennom* åpne vinduer. (HW2)
        b. Sea gulls shrieked to them *through the open windo*ws. (HW2TE)
        c. Les mouettes criaient *à travers les fenêtres ouvertes*. (HW2TF)

(5)   a. Han hadde teven av Dina tvers *gjennom* sildetønner …(HW2)
        b. He sensed Dina's aroma even *through the odors of herring barrels…* (HW2TE)
        c. L'odeur de Dina lui parvenait *à travers les tonneaux de harengs* ... (HW2TF)

Figure 3 contains details of the correspondences between the two sets of translations in the case of temporal predications.
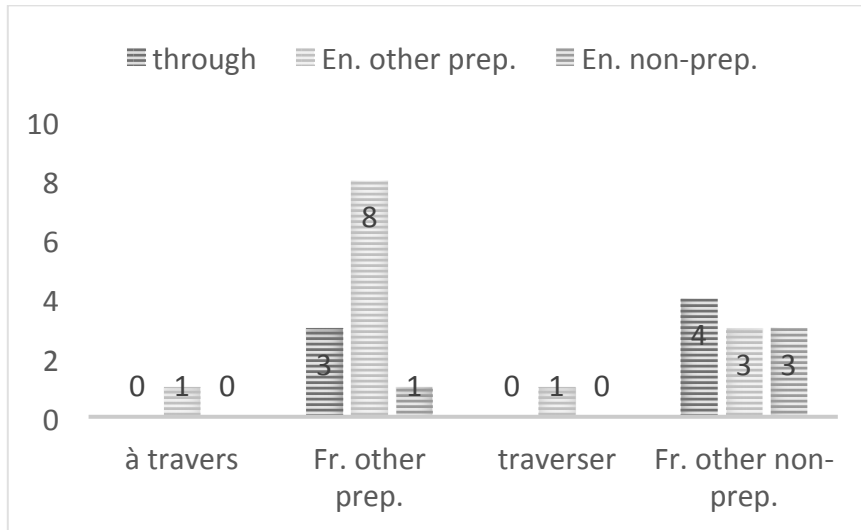
Figure 3. Correspondences between English and French codings of Time [throughness] (Egan 2014)

In Figure 3, we can see that translations of predications of temporal [throughness] differ markedly from those of perceptual [throughness] in that not so much as a single token is translated by both *through* and *à travers*. Of a total of 24 tokens just seven are translated into English by *through* and one into French by *à travers*. English employs alternative prepositions in 13 cases and French in 12. In eight cases both languages employ them, as in examples (6) – (7).

(6)   a. Niels hadde brukt stor kløkt *gjennom flere år*. (HW2)
       b. Niels had operated very cleverly *for many years*. (HW2TE)
       c. Niels avait montré une grande intelligence *pendant plusieurs années*. (HW2TF)

(7)   a. Trolig har han tatt for tunge løft i de unge år, for ut *gjennom manndomsalderen* gikk det så raskt tilbake med førligheten … (BHH1)
       b. He probably overtaxed himself in his youth, for *in his mature years* his vigour declined so quickly … (BHH1TE)
       c. Sans doute avait-il trop forcé au cours de ses jeunes années car, *durant sa maturité*, sa santé déclina si vite... (BHH1TF)

The French translators employ non-prepositional constructions in 11 cases, one of which contains the verb *traverser* and the English translators do so in four cases. Three of these overlap, including examples (8) and (9).

(8)   a. Men bildet dvergen knipset hadde allerede forfulgt Ana *gjennom* hele livet. (JG3)

b. But the photo he'd taken had already hounded Ana *all* her life. (JG3TE)

c. Mais ce portrait avait, en réalité, poursuivi Ana *toute* sa vie. (JG3TF)

(9)   a. 1. Han hadde vært medlem av Nasjonal Samling siden 1934 og betalt kontingent *gjennom* alle år. (BHH1)

b. (1) He had been a member of the National Union since 1934 and paid his membership fee *every* year. (BHH1TE)

c. 1) Il avait été membre du N.S. de 1934 à 1945 et s'était *régulièrement* acquitté de sa cotisation. (BHH1TF)

In (8) both translators employ a predeterminer rather than a preposition, *all* in English and *toute* in French, while in (9) the English translator employs the determiner *every* and the French one the adverb *régulièrement*, to code the fact that the actions in question occurred at regular intervals throughout the periods in question.

There is no doubt that the data presented in Figures 2 and 3 show that (translators into) English and French employ very different means of coding [throughness] with respect to perception and time. It is also clear that the sort of data presented in these two figures could not have been reliably elicited using pictures of any sort. Nor could they have been discovered using a 2-text corpus.[4] In any event it is hard to think of search items that would prove efficacious in unearthing them. And even were it possible to think of such items, the procedure for uncovering the relevant tokens would need to be very laborious. Yet the data are worth having insofar as they give us an indication of the correspondences

---

[4] An anonymous reviewer maintains that it would be possible to arrive at these conclusions using several 2-text corpora. While I agree that this might in theory be possible, one would still be faced with the problem of ensuring that the tokens in the two source corpora were as similar as possible.

between codings of certain perceptual and temporal predications in the two languages.

One type of relation that exhibits a great deal of overlap with respect to all sorts of predications is [betweenness], as is demonstrated in Egan (2013a) in a study of translations into English and French of the Norwegian preposition *mellom*. Table 6 contains details of the overlap between *between* and *entre* in coding the various sense of *mellom*.

Table 6. Tokens of *mellom* in the OMC translated by *between* and/or *entre* (Egan 2013a)

Type	Total tokens	between	entre	between + entre
Location	139	101	97	77 55%
Motion	89	50	60	40 45%
Interaction	51	44	36	32 63%
Relationship	47	43	35	33 70%
Comparison	36	33	23	21 58%
Time	17	13	11	11 65%
Idiom	14	8	8	7 50%
Total	393	292	270	221 56%

Of a total of 393 tokens of Norwegian *mellom*, 74% are translated into English by *between* and 69% into French by *entre*. Moreover the two prepositions correspond to one another in 56% of all cases. One may legitimately question the usefulness of this calculation of the overlap between the two prepositions, limited as it is to tokens that translate Norwegian *mellom*. We can however also use the 3-text corpus to work out the degree of mutual correspondence between the English and French prepositions *between* and *entre* using a method based on Altenberg (1999). Altenberg's method involves the division of the total number of occurrences in target texts in a 2-text corpus of item *a* translating item *b* and vice versa by the total number of occurrences of both terms in the two sets of source texts. Multiplying the result of this calculation by 100 gives us the percentage overlap of the two items, which Altenberg labels their 'mutual correspondence'. We can adapt this method to 3-text corpora by replacing the total of *a* translating *b* and vice versa by the total number of mutual occurrences in the two translated texts of *a* and *b* multiplied by two. We have to multiply by two, since the correspondence is in both directions, i.e. we count the *a*s corresponding to *b*s plus the *b*s corresponding to *a*s. We then divide the result by the total number of tokens of both *a* and *b* in the two translated texts rather than the source

texts, which by definition do not contain either item. There are a total of 365 tokens of *between* in the English translations in the OMC and 477 of *entre* in the French translations. 242 of these occur in parallel translations. Using the above formula, their degree of mutual correspondence in the translations, which we may label MC_t, can be calculated as follows:

$$MC_t = \frac{(\text{overlap } \textit{between}/\textit{entre})*2*100}{\text{total } \textit{between} + \text{total } \textit{entre}}$$

$$MC_t = \frac{(242)*2*100}{365 + 477} = 59.61\%$$

This figure of 59.61% may be compared to the total for mutual correspondence of the two items in translations of *mellom* which, according to Table 6, is 56%. These figures may, of course, also be compared to results of calculations using Altenberg's original formula on the correspondence of the two items in 2-text corpora.

I round off this section by summarising what I see as the advantages of 3-text corpora for a comparison of items in two languages. In the first place 3-text corpora afford us a sound(er) basis for comparing two texts since in each case we know both what both language users are aiming to convey, since they are both translators, and how they go about encoding this. Moreover, unlike in the case of experiments using visual prompts, there is no danger of the results being contaminated by the observer's paradox. Lastly, 3-text corpora allow us to calculate the overlap between expressions in two languages using a different method to that of pairs of 2-text corpora. The results of the two methods may be mutually informative.

5. The ENPC and ESPC as a 3-text corpus

The source tests in the ENPC and the ESPC are to a large extent (approx. 80%) the same. As pointed out by Hasselgård (2007):

> The English-Norwegian Parallel Corpus (ENPC) and the English-Swedish Parallel Corpus (ESPC) are sister corpora with common objectives, design criteria, and structures; cf. Aijmer *et al* (1996:79 f). Especially their fiction parts also share a good number of texts. One advantage of this is that it is possible not only to compare

> English to Norwegian and Swedish respectively, but also to compare different translations of the English originals with each other. (Hasselgård 2007, section 1)

By eliminating tokens from texts that are exclusive to one corpus or the other, we can thus use the ENPC and the ESPC for contrastive studies of Swedish and Norwegian, just as we can use the Norwegian – English – French – German part of the OMC for contrastive studies of English, French and German. Hasselgård (2007) deals with adverbs of frequency and usuality. She concludes that the jury is still out on the suitability of ENPC and the ESPC for the sort of contrastive study she conducted, at least if one is searching for "lexical alternatives, expecting the translations used in the ENPC to be automatically usable in Swedish or those in the ESPC to work in Norwegian" (Hasselgård 2007: section 8). However, she is more positive to the usefulness of the 3-text corpus for comparing constructions.

Hasselgård's study is based on the 24 overlapping fictional texts in the two corpora.[5] However, there are 39 texts in all that feature in both corpora, 24 fiction and 15 non-fiction, and the latter are equally suitable for inclusion in a contrastive study. The actual texts that overlap are:

- In the ENPC:
    - Fiction: all texts minus AB1, BC1, TH1, DL1, DL2 and MM1
    - Non-fiction: all texts minus HB1, ROB1, SJG1, ML1 and LTLT1
- In the ESPC:
    - Fiction: all texts minus DLO1
    - Non-fiction: all texts minus CAOG1, RH1, RL1, JPM1, AS1, ABB1, AZ1, PHA1, STO1, and all Nobel lectures and parliamentary speeches

The complete overlapping corpus of both fiction and non-fiction was used by Rawoens & Egan (2015) in order to compare the Norwegian preposition *mellom* with its Swedish counterpart *mellan*. Figure 4 shows

---

[5]   The full lists of texts in the two corpora may be consulted at <http://www.sol.lu.se/engelska/corpus/corpus/espc.html>   (ESPC)   and <https://www.hf.uio.no/ilos/english/services/omc/> (ENPC).

the total number of tokens of both prepositions when used to translate English *between*.
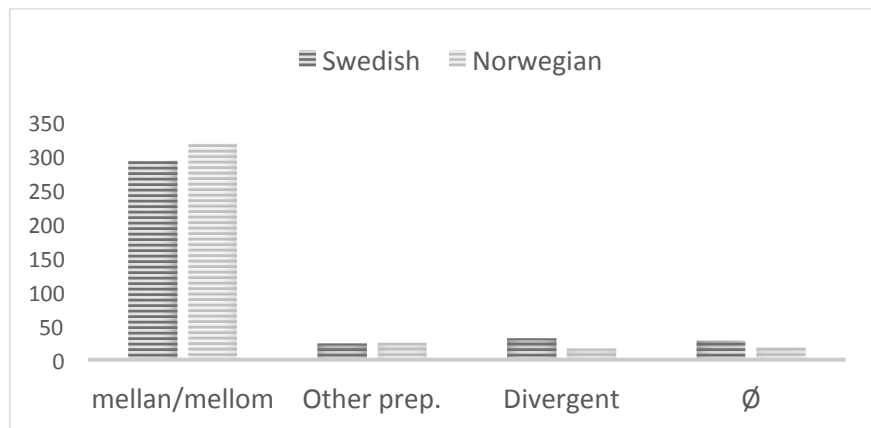

Figure 4. Swedish and Norwegian translation equivalents of English *between* (Rawoens & Egan 2015)

The overlap between the two prepositions in translations of *between* is 73%, while their degree of mutual translation correspondence (MC_t), calculated using the adjusted Altenberg method described in section 4, is 70%. There are more tokens in Norwegian than in Swedish (25 as opposed to 18) in which an alternative preposition is employed to encode a [betweenness] relationship. Especially common in Norwegian is the composite proposition *fra…til* ('from…to'), represented by 11 tokens. In nine cases these predications are rendered by *mellan* in Swedish, as in example (10).

(10) a. Some double stars are so close that they touch, and starstuff flows *between* them. (CSA1)
      b. Vissa dubbelstjärnor ligger så nära varandra, att de snuddar varandra och stjärnmateria flödar *mellan* dem. (CSA1TS)[6]

---

[6] 'TS' stands for Swedish translated text and 'TN' for Norwegian translated text, in accordance with the system for identifying texts in the OMC and utilised in examples (1) – (9).

c. Enkelte dobbeltstjerner står så nær hverandre at de berører hverandre, og stjernestoffet bølger *fra* den ene *til* den andre og tilbake igjen. (CSA1TN)

There are just over twice as many divergent constructions in the Swedish than the Norwegian translations of *between* (30 compared to 14). Seven of these are divergent in both languages. One of these is reproduced here as example (11). The Swedish translator has employed an adverbial meaning 'together' and the Norwegian translator a verb meaning 'help' and a reciprocal pronoun meaning 'one another'.

(11) a. She and Ryan *between* them cope with the baby. (PDJ1)
      b. Och *tillsammans* klarar hon och Ryan av babyn. (PDJ1TS)
      c. Hun og Ryan hjelper *hverandre* med den lille. (PDJ1TN)

Turning to tokens of *mellan* and *mellom* that do not translate *between*, there are 128 of these in the Swedish translations and 125 in the Norwegian translations. In 32 cases *mellan* and *mellom* overlap. One could argue that this overlap represents stronger evidence for the extent of semantic similarity between the Swedish and Norwegian prepositions than translations of *between*, since *mellan* og *mellom* must be considered the default translation correspondences of the English preposition. Two thirds of the tokens of *mellan* and *mellom* occur in congruent translations, with the most frequent original prepositions being *of*, *through*, *from*, *in* and *among*. In (12) both translators have chosen to translate *through* by their equivalent of *between*.

(12) a. … and picking their way with distaste *through* the puddles on the pavement. (PM1)
      b. De kryssar med avsmak *mellan* vattenpölarna på trottoarerna. (PM1TS)
      c. Så tar de seg forsiktig frem *mellom* søledammene på fortauet med sterkt misbilligende ansiktsuttrykk. (PM1TN)

Had the two translators chosen to employ *genom/gjennom*, the most common Swedish and Norwegian correspondences of *through*, the target texts would have differed from the source text in predicating that the

subjects actually put their feet in the puddles rather than wove their way around them.[7]

In one third of the cases where *mellan* and *mellom* cooccur but do not translate *between*, the source text contains a divergent construction, that is one without a preposition. One English construction that stands out as being avoided by both the Swedish and Norwegian translators is the coding of bilateral relations by means of a hyphen, as in 'East-West suspicions' in (13). Other original tokens of the same type are 'The Iran-Iraq war' (CS1), 'the Sino-Soviet split' (MAW1) and 'U.S.-British relations' (AH1).

(13)   a. *East-West* suspicions have erected a visible barrier…(CS1)
       b. Den ömsesidiga misstänksamheten *mellan* öst och väst har upprättat en synlig barriär… (CS1TS)
       c. Mistenksomhet *mellom* øst og vest har reist en synbar skranke… (CS1TN)

In both (13b) and (13c) the suspicions are said to pertain *between* East and West. The coding of this relation by means of a preposition rather than a hyphen in both Swedish and Norwegian illustrates a feature common to these two languages. Such shared features are easy to spot in a 3-text corpus, since they stand out like individual trees in a wood. In a 2-text corpus consisting of texts from closely related languages, such as Swedish and Norwegian, on the other hand, their very similarity renders their identification more difficult.

6. Summary and conclusion
In this paper I have argued that 3-text translation corpora afford us possibilities for conducting contrastive research that are different to, and in some respects superior to, those afforded by (pairs of) 2-text corpora. Contrasting translations is different to contrasting original texts and translations insofar as the two items being compared have been produced in response to the exact same set of prompts, which function as the *tertium comparationis*. It is the identity of these prompts which led me to

---

[7] There are some similar examples of *mellom* translating *through* in Egan (2012: 48), including 'Diana picked her way *through* the women and answered her front door.' (ST1)

argue that they are also more suitable for comparison than a corpus consisting of original texts and translations of these, irrespective of whether the latter consists of translations in one or two directions. We have also seen that sets of translations in a 3-text corpus may furnish us with an alternative method for gauging the degree of mutual correspondence between linguistic items in the two languages concerned. This method is based on that proposed by Altenberg (1999). Finally, I hope I have shown that the common original texts in the ESPC and ENPC are suitable for contrastive analyses of expressions in Swedish and Norwegian.

References

Aijmer Karin, Bengt Altenberg and Mats Johansson, Eds. 1996. *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies Lund 4-5 March 1994*. Lund: Lund University Press

Altenberg, Bengt. 1999. "Adverbial connectors in English and Swedish: Semantic and lexical correspondences." *Out of Corpora. Studies in Honour of Stig Johansson*. Eds. Hilde Hasselgård and Signe Oksefjell. Amsterdam: Rodopi. 249-268.

Altenberg, Bengt and Sylviane Granger. 2002. "Recent trends in cross-linguistic lexical studies." *Lexis in Contrast: Corpus-based Approaches.* Eds. Bengt Altenberg and Sylviane Granger. Amsterdam: John Benjamins. 3-48.

Berman, Ruth and Dan Slobin. 1994. *Relating Events in Narrative: a Crosslinguistic Developmental study*. Hillsdale, NJ: Laurence Erlbaum.

Cappelle, Bart. 2012. "English is less rich in manner-of-motion verbs when translated from French." *Across Languages and Cultures* 13(2): 173-195.

Ebeling, Jarle and Signe Oksefjell Ebeling. 2013. *Patterns in Contrast.* Amsterdam: John Benjamins.

Egan, Thomas. 2012. "*Through* seen through the looking glass of translation equivalence: a proposed method for determining closeness of word senses." *Corpus Linguistics: Looking Back – Moving Forward*. Eds. Sebastian Hoffman, Paul Rayson and Geoffrey Leech. Amsterdam: Rodopi. 41-56.

Egan, Thomas. 2013a. "*Tertia Comparationis* in Multilingual Corpora." *Advances in Corpus-based Contrastive Linguistics. Studies in Honour of Stig Johansson*. Eds. Karin Aijmer and Bengt Altenberg. Amsterdam: John Benjamins. 7-24.

Egan, Thomas. 2013b. "Between and through revisited." *Corpus linguistics and variation in English: Focus on Non-native Englishes*, a special issue of the online-journal VARIENG: Studies in Variation, Contacts and Change in English. Eds. Magnus Huber and Joybrato Mukherjee.

Egan, Thomas. 2014. "Encoding *throughness* in English and French." *Diachrony and Synchrony in English Corpus Studies*. Eds. Alejandro Alcaraz Sintes and Salvador Valera Hernandez. Frankfurt: Peter Lang. 233-257.

Egan, Thomas. 2015a. "Manner and Path: evidence from a multilingual corpus." *CogniTextes* 12.

Egan, Thomas. 2015b. "Motion *to* and motion *through*: evidence from a multilingual corpus". *From Clerks to Corpora: Essays on the English Language Yesterday and Today*. Eds. Philip Shaw, Britt Erman, Gunnel Melchers and Peter Sundkvist. Stockholm: Stockholm University Press. 285-302.

Egan Thomas and Anne-Line Graedler. 2015. "Motion *into* and *out of* in English, French and Norwegian." *Nordic Journal of English Studies*, 14 (1): 9-23.

Engemann, Helen, Anne-Katharina Harr and Maya Hickmann. 2012. "Caused motion events across languages and learner types: A comparison of bilingual first and adult second language acquisition." *Space and Time in Languages and Cultures: Linguistic Diversity*. Eds. Luna Filipović and Katarzyna M. Jaszczolt. Amsterdam: John Benjamins. 263-287.

Garretson, Gregory. 2004. "The meanings of English "of": uncovering semantic distinctions via a translation corpus." Unpublished MA thesis, Boston University.

Gellerstam, Martin. 1996. "Translations as a source for cross-linguistic studies." *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies, Lund, 4–5 March 1994*. Eds. Karin Aijmer, Bengt Altenberg and Mats Johansson. Lund: Lund University Press. 53-62.

Hasselgård, Hilde. 2007. "Using the ENPC and the ESPC as a Parallel Translation Corpus: Adverbs of Frequency and Usuality." *Nordic Journal of English Studies,* 6(1).

Johansson, Stig. 1998. "On the role of corpora in cross-linguistic research." *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*. Eds. Stig Johansson and Signe Oksefjell. Amsterdam: Rodopi. 3-24.

Johansson, Stig. 2001. "The German and Norwegian correspondences to the English construction type *that's what*." *Linguistics* 39 (3): 583-605.

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

Krzeszowski, Tomasz P. 1990. *Contrasting Languages: The Scope of Contrastive Linguistics*. Berlin: Mouton de Gruyter.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania.

Mayer, Mercel. 1969. *Frog, where are you?* New York: Dial Books.

Paulussen, Hans. 1999. "A corpus-based contrastive analysis of English *on/up*, Dutch *op* and French *sur* within a cognitive framework." Unpublished doctoral dissertation, Ghent University.

Rawoens, Gudrun and Thomas Egan. 2015. "Coding *betweenness* in Swedish and Norwegian translations from English." *Meta: Journal des traducteurs* 60 (3).

Slobin, Dan. 2005. "Relating Narrative Events in Translation." *Perspectives on Language and Language Development: Essays in Honor of Ruth A. Berman.* Eds. Dorit Diskin Ravid and Hava Bat-Zeev Shyldkrot. Dordrecht: Kluwer. 115-129.

Verkerk, Annemarie. 2014. "Where Alice fell into: Motion events from a parallel corpus." *Aggregating Dialectology, Typology, and Register Analysis: Linguistic Variation in Text and Speech.* Eds. Benedikt Szmrecsanyi and Bernhard Wälchli. Berlin: Walter de Gruyter. 324-354

Verkerk, Annemarie. 2015. "Where do all the motion verbs come from? The speed of development of manner verbs and path verbs in Indo-European." *Diachronica*, 32 (1): 69-104.

Viberg, Åke. 1998. "Contrasts in polysemy and differentiation: Running and putting in English and Swedish." *Corpora and Cross-linguistic*

Research: Theory, Method, and Case Studies. Eds. Stig Johansson and Signe Oksefjell. Amsterdam: Rodopi. 343-376.

Viberg, Åke. 2003. "The polysemy of the Swedish verb komma 'come': A view from translation corpora." *Meaning Through Language Contrast: Vol 2*. Eds. Katarzyna M. Jaszczolt and Ken Turner. Amsterdam: John Benjamins. 75-105.

Viberg, Åke. 2013. "Seeing the lexical profile of Swedish through multilingual corpora: The case of Swedish *åka* and other vehicle verbs." *Advances in Corpus-based Contrastive Linguistics. Studies in Honour of Stig Johansson*. Eds. Karin Aijmer and Bengt Altenberg. Amsterdam: John Benjamins. 25-56.

Vulchanova, Mila, Liliana Martinez and Valentin Vulchanov. 2012. "Distinctions in the linguistic encoding of motion: evidence from a free naming task." *Motion Encoding in Language and Space*. Eds. Mila Vulchanova and Emile Van Der Zee. Oxford: Oxford University Press. 11-43.

Corpora
ENPC: http://www.hf.uio.no/ilos/english/services/omc/
ESPC: http://www.sol.lu.se/engelska/corpus/corpus/espc.html
OMC: http://www.hf.uio.no/ilos/english/services/omc/