

An assessment of the translation-specificity of over-represented multi-word patterns in Swedish fiction texts translated from English

P-O Nilsson

English Department - Göteborg University (Sweden)

This paper discusses in what sense an over-represented multi-word pattern in a corpus of translated texts can be said to be typical of translated text. The purpose of the discussion is to assess, from a quantitative as well as from a qualitative perspective, the status of translational collocation data retrieved through data-driven methods from a comparable and parallel aligned corpus of English and Swedish original and translated texts. The study focuses on the explicitation of clausal relations in translations from English into Swedish. In some cases, lexical and grammatical contrast lead to explicitation, but in others explicitation is due to different factors.

1. Introduction

Corpus-based descriptive translation studies (cf. Baker 1996) has proved to be a fruitful alliance between corpus linguistics and translation studies, particularly in terms of the quantification of results and the processing of large amounts of data, to reveal general tendencies of surface patterning in translated texts. In recent years, the trend within this line of translation studies has been to focus on TL oriented research, combined with a quantitative approach. This type of study is typically represented by work carried out on the *Translational English Corpus* (TEC) (e.g. Laviosa 1998; Olohan 2001), a corpus which is composed of texts translated into English from a variety of source languages. In the case of Swedish – the category of TL texts examined in this paper – early TL-oriented studies were carried out in the mid-1980s by Martin Gellerstam, who used Swedish fiction texts translated from English (e.g. Gellerstam 1986). A small investigation of Swedish comparable corpora was carried out in Nilsson 2002.

The value of large-scale quantitative studies of comparable corpora consisting of non-translated and translated TL texts is that such studies allow one to make significant generalizations about general features of the translated texts, such as overrepresentation of lexical items and combinations of items. What cannot be achieved within an exclusively TL-oriented approach, however, is a systematic TL-SL comparison of elements below text level, since no aligned source texts are available in a comparable TL corpus. This way, the method restricts the perspective to one of fairly general-level statements about the properties of translated texts as compared with non-translated texts (which is often the purpose of such studies).

The data discussed in this paper result from the application of the TL-oriented method in conjunction with a method that should be seen as com-

plementary to it: maintaining the TL orientation through using the comparison of TL texts for the definition of the linguistic items to study, and supplementing these data with the aligned SL text segments for the purpose of explanation. On the basis of these TL and SL data, the paper discusses in what sense one specific over-represented multi-word pattern in a corpus of translated texts can be said to be typical of translated text.

2. Material, method and aim

The material used for this investigation is the fiction part of *The English-Swedish Parallel Corpus* (ESPC), a comparable and parallel aligned corpus of English and Swedish original and translated texts. Designed fairly generally for cross-linguistic research, the ESPC may be contrasted with a corpus such as the TEC (cf. previous section), which is expressly designed with the specific aim of providing a basis for research into texts as translations, i.e. into their roles and functions as translations, and not for general contrastive or typological research. An extensive range of parameters have been taken into account in the construction of the TEC in order to ensure comparability, so that any specific features one encounters in the translated texts can reasonably be assumed to be due to their being translations and not to other factors (cf. Laviosa 1997). In the case of the ESPC, a more general set of corpus design criteria were applied, and it was not necessarily always the translated text which was used as the point of departure for the inclusion of an SL-TL pair of texts in the corpus (cf. Altenberg, Aijmer & Svensson 2001).

Still, it seems reasonable to assume that the translational and non-translational parts of the fiction section of the ESPC are adequately comparable for the research task at hand, which is an investigation of generally occurring high-frequency grammatical items and constructions. The texts belong to the same general category of mainstream literature published by major Swedish publishers from the early 1980s to the late 1990s (with a few exceptions), and a wide range of authors and translators are represented. Moreover, in terms of representativity, the ESPC is a sample corpus, consisting of many extracts (10,000–15,000 words each) rather than a more limited number of whole texts, and this makes for representativity in terms of generally occurring translation features, as contrasted with features typical for individual texts. The degree of even distribution of an item over the corpus as a whole can be assumed to be an indication of its generality in translated texts (cf. Nilsson 2002).

In terms of method, the focus of the investigation is on generally occurring features rather than on features of individual texts. The first step of the investigation can be said to be corpus-driven (cf. Tognini Bonelli 2000 for a discussion of this concept), in that a specific frequency difference between the translated and non-translated subcorpora defines a linguistic area of study. In this case, the linguistic structure defined is the Swedish multi-word pattern *och + som* ('and' + 'who' / 'whom' / 'whose' / 'which' / 'that' / 'as'), which is over-represented in the translated fiction subcorpus (cf. next section).¹

The motivation for then going on to investigate this specific frequency difference more closely is that it may indicate explicitation of clausal relations, a feature potentially leading to more explicit and redundant text: the Swedish *och* is a coordinating conjunction, and in coordination ellipsis is frequent; if a relative pronoun (*som*) is over-represented as a collocate of a coordinating conjunction, this may be taken as some indication that ellipsis has been “filled in”, leading to increased redundancy in the TL text. An increased level of redundancy and explicitness has been identified as a typical feature of translations before (Blum-Kulka 1986:19). Indeed, in a study of translation between English and Norwegian, Øverås (1998) was able to confirm Blum-Kulka’s explicitation hypothesis, although finding stronger support for it in translations from English into Norwegian than in the reverse direction. In her study, Øverås also states that although she strived to exclude examples reflecting systemic differences between the two languages, there may still be contrastive differences influencing the results, in conjunction with TL community norms. The aim of this paper is to try to make a distinction between instances of explicitation of clausal relations that are determined by linguistic contrast, and instances that are related to other factors, such as the translation process itself or the translation norms of the target community. The hypothesized norm is that translations, at least in some genres², will be more explicit and redundant, more readable and fluent than texts written as originals in the target language.

3. Results and discussion

3.1. Introduction

This section outlines the distribution of occurrences of the investigated TL multi-word pattern *och som*, then isolates a subcategory (coordinated relative clauses) for further inspection. Distributions in TL translated and non-translated texts are compared, and the main types of SL counterparts are described.

In the examples, the source text is presented first, followed by the translation (the target text), which is in turn followed by a gloss back-translation of the relevant target word or construction.³

3.2. TL coordinated relative clauses

The TL multi-word pattern *och som* occurs 61 times in the Swedish original fiction texts in the ESPC, and 100 times in the Swedish translated fiction part.⁴ Among these, 51 instances in the non-translated material and 82 instances in the translated material represent coordinated relative clauses.⁵ There can therefore be said to be a significant degree of over-representation of coordinated relative clauses represented by this specific surface pattern in the translated texts. The investigated TL pattern is exemplified by the Swedish translation in (1):

- (1) That was the house that we grew up in **and that** my father lived in. (JSM1)
 Det var det huset vi växte upp i **och som** min far nu bodde i. ('and that') (JSM1T)

In this example, where the English source text has a coordinated relative clause with an explicit coordinator (*and*) as well as an explicit relative word (*that*), the translator has rendered the SL construction by means of a Swedish relative clause which also has an explicit coordinator (*och*) and an explicit relative word (*som*).

3.3. SL counterparts

The closest corresponding SL counterparts of the TL coordinated relative clause were classified according to whether or not they contained an explicit SL coordinator and / or an explicit relative word. This resulted in the following distribution of the four different possibilities:

Table 1: Distribution of SL surface features corresponding closest to 'och som' ('and' + RELATIVE PRONOUN) in the TL coordinated relative clause.

SL surface feature	Freq.
<i>An explicit SL coordinator and an explicit SL relative word</i>	25
<i>An explicit SL coordinator and no explicit SL relative word</i>	18
<i>No explicit SL coordinator and an explicit SL relative word</i>	19
<i>No explicit SL coordinator and no explicit SL relative word</i>	20
Total	82

The classification reveals a fairly even distribution of instances among the four groups defining the occurrence and non-occurrence of a coordinator and a relative word in the English source texts. As all the Swedish translations of these instances contain a coordinator and a relative word, only the cases in the first of the above SL categories are formally congruent with the TL renderings.

In the following subsections, each of the categories are exemplified. It should be stressed that the exemplification is not exhaustive, and aims merely at giving an overview – it sometimes focuses on the presence or absence of a coordinator, sometimes on the presence or absence of a relative word.

3.3.1. An explicit SL coordinator and an explicit SL relative word

In both English and Swedish, dependent clauses of the same type may be coordinated. This is exemplified by the relative clauses in (2):

- (2) I saw a girl amongst them who had eyes at the side of her face, bangles of blue copper round her neck, **and who** was more lovely than forest flowers. (BO1)

Jag såg ibland dem en flicka som hade ögon på sidan av ansiktet, armband av blå koppar runt halsen **och som** var ljuvligare än skogens blommor. ('and who') (BO1T)

In the TL text in (2), the option without a coordinator would wrongly identify the antecedent as *armband* ('bangles') rather than *flicka* ('girl') since the Swedish relative *som*, unlike English *who*, does not specify that the antecedent is a person. (As for the choice of a TL coordinating device, a grammatically less explicit possibility in this case would be a comma instead of *och*.) Thus the impossibility of ellipsis of the coordinator can be ascribed to a feature of the contrasting linguistic systems involved. This is however mostly an issue of theoretical interest in order to illustrate the systemic contrast, and does not constitute a real problem for a professional translator, for whom there is no real reason to formally deviate from the SL pattern in an unproblematic case such as this.

More interesting from the point of view of translation is the possibility of omitting the pronoun, since this represents a kind of reduction which is possible (although it may have stylistic effects) and may or may not be realized depending on the kind of decision made by the translator. In this case, the translator quite unproblematically chooses to preserve the same degree of explicitness as in the SL text.

3.3.2. An explicit SL coordinator and no explicit SL relative word

In English as well as in Swedish, a relative clause may occur without a relative word. Omission of the relative pronoun is very common in restrictive relative clauses. In (3), the SL clause has neither a relative pronoun nor a subject, but they are rendered in the translation, along with the coordinator:

- (3) But there are also things that aren't so good, some that I don't like **and** hope to alter. (AH1)

Men det finns också saker som inte är så bra, en del som jag inte tycker om **och som** jag hoppas kunna förändra. ('and that') (AH1T)

In (3), the TL text has been made more explicit through the realization of relative *som* and a subject (*jag*; 'I') although a formally more congruent translation without these elements would have been grammatically possible and fully acceptable. In the choice between an explicit and a less explicit alternative, the translator chooses the more explicit alternative although there is no imperative necessity to do so. Thus, a case of clausal explicitation such as this could reasonably be attributed to other factors than systemic contrast, in this case possibly stylistic considerations having to do with this extract being a piece of dialogue.

A further SL source of the TL pattern is a coordinated postmodifying non-finite clause. Non-finite clauses have a much more restricted use in Swedish, and a relative clause is the typical Swedish counterpart. In (4), the source text has an *-ing* participle clause:

- (4) Another two came when called; four women and two men, all worried and uncertain **and** looking to me for decisions. (DF1)
 Fyra kvinnor och två män som alla var oroliga och osäkra **och som** förväntade sig att jag skulle ta kommandot. ('and who') (DF1T)

Examples (3) and (4) illustrate two different aspects of what leads to the realization of the TL pattern in the corpus: possibilities of (a stylistically or otherwise motivated) choice between explicitation or non-explicitation, against a background of systemic differences that necessitate explicating shifts.

3.3.3. No explicit SL coordinator and an explicit SL relative word

The instances in this category are cases where the English source text has a relative word only, and where the Swedish translator has added a coordinator in the TL rendering, as in example (5):

- (5) In Paris was the much-acclaimed collection of the banker and art connoisseur Everard Jabach, **which** consisted entirely of copies of works he had once owned. (JH1)
 I Paris fanns den mycket omtalade samling som tillhörde bankiren och konstkännaren Everard Jabach **och som** bestod enbart av kopior av verk vilka han en gång hade ägt. ('and which') (JH1T)

This is similar to example (2) above in that *och* is necessary in order to achieve a more unambiguous assignment of the antecedent (*samling*; 'collection'), this kind of disambiguation being necessary because of the difference in pronoun systems: while *which* may not have a person as antecedent, *som* is unspecified in this respect. Example (6), below, is related to systemic factors in a less absolute sense, although these play a part here as well:

- (6) He 'd never met a woman of that age before who did not know his name and family, **who** did not stop to swap a word or two. (JC1)
 Han hade aldrig förr mött en kvinna i den åldern som inte hade känt till hans namn eller släkt **och som** inte hade stannat för att växla några ord. ('and who') (JC1T)

In the TL rendering, the option without a coordinator is conceivable, although less acceptable, and a comma would need to be inserted in order to avoid the unlikely assignment of *släkt* ('family') as antecedent instead of *kvinna* ('woman'). The chosen alternative – coordination with *och* – disambiguates the sentence and increases readability.

3.3.4. No explicit SL coordinator and no explicit SL relative word

This concluding category involves cases where the English source text has neither a relative word nor a coordinator, and where the Swedish rendering has both. Many of the SL instances are cases of non-finite clauses in post-modifying or other functions. The use of non-finite clauses is much more restricted in Swedish. Example (7) is similar to the example of a postmodifying non-finite clause in (4) above, but it is even less explicit, lacking a coordinator as well as a relative word before the verb, *killing*:

- (7) In Germany they later portrayed the Plague as a maid travelling through the air like a blue flame, killing her victims simply by raising an arm. (BR1)
 I Tyskland avporträtterades pesten senare som en jungfru farande genom luften som en blå eldslåga **och som** dödade sina offer endast genom att sträcka upp armen. ('and who') (BR1T)

This is a clear example of a TL systemic restriction leading to clausal explicitation, since a Swedish non-finite clause would be grammatically dubious.

Example (8), finally, is a case of more stylistically motivated than grammatically conditioned clausal explicitation:

- (8) Instead, he went to a café he liked called My American Cousin. (AT1)
 I stället gick han till ett kafé som han tyckte om **och som** hette My American Cousin. ('and which') (AT1T)

In this case several TL options are possible. The SL *-ed* participle *called* could be translated as the Swedish counterpart *kallat*, but the resulting construction would involve a slight change of meaning and would also belong to a fairly formal Swedish register. Another solution is to include the relative *som* only, which would render the sentence fully comprehensible and stylistically unmarked. As a further solution (the chosen alternative), the translation may include the coordinator *och* as well as the relative *som*. Semantically, the last two alternatives are virtually indistinguishable from one another (cf. Telemann et al. 1999:496-7). Still, the clausal-introducing solution opted for by the translator is the maximally explicit alternative, including both a coordinator and a relative word.

3.4. Co-textual factors

The classification and description of the examples in the above sections focused on the features of the specific clause under investigation. It goes without saying that there is a multitude of co-textual factors that have an impact on the type of rendering chosen, and none of these have been treated here. A few cases are therefore exemplified and briefly discussed in this section.

One important co-textual factor is the SL constituent that results in the TL rendering with which the investigated TL relative clause is to be coordinated. Consider examples (9) and (10):

- (9) His eyes were streaming from the **windborne** grit and dust which pecked and spiralled at his face. (JC1)
 Ögonen rann på grund av dammet och sanden **som vinden förde med sig och som** kom virvlande upp i ansiktet på honom och stack som nålar. ('...which the wind brought with it and which...') (JC1T)
- (10) It might have been some minor disaster **befalling her employers that** Vine and Malahyde had told her about, a bit of the roof blowing off, for instance, or water through a ceiling. (RR1)
 Det kunde lika gärna ha varit någon mindre olycka **som drabbat hennes arbetsgivare, och som** Vine och Malahyde hade meddelat henne, som till exempel att en bit av taket blåst bort, eller att taket läckte. ('...that had struck her employers, and that...') (RR1T)

In example (9), the English premodifier *windborne* is translated as a Swedish postmodifying relative clause, *som vinden förde med sig*. This in turn results in the following relative clause being coordinated, with *och*. In example (10), the postmodifying SL non-finite clause *befalling her employers* is translated as a Swedish postmodifying relative clause with which the following clause is coordinated. From the perspective of the investigated coordinated relative clauses, each of the instances of explicit coordination in examples (9) and (10) can be said to be indirectly conditioned, by co-textual translational shifts: where there was no explicit coordination in the source text, coordination arises in the translation as a result of a change in the preceding co-text. Co-textual factors can also be seen to intervene in other examples: in example (5), for instance, an English *of* construction is translated as a Swedish relative clause, with which the second relative clause is coordinated.

4. Summary and conclusion

Although the exemplification in the above sections is far from exhaustive, the examples can be said to indicate a kind of paradigm of English structures that have resulted in Swedish coordinated relative clauses. First, there are systemic SL features that predictably lead to explicating clausal shifts because of linguistic contrast. These contrasts include both lexical differences (the sets of relative pronouns in the two languages) and grammatical differences (e.g. the respective norms of usage for non-finite clauses). Second, there are areas where explicating shifts are not entirely necessary, but stylistically preferable. Third, there is a less predictable category of

clausal explicitation where no systemic contrast can be seen to operate, but where a translator simply chooses the most explicit of several available TL alternatives. Finally, once co-text is taken into account, an additional set of dynamic factors that appear to have an influence emerges.

The purpose of this study was to evaluate the status of a set of linguistic data retrieved through a TL-oriented and data-driven method, in terms of the translation-bound nature of these data. It is not entirely clear to what extent the over-representation of the coordinated relative clause pattern investigated here is due to “explicitation proper” (which may be defined as explicitation that can be attributed to factors other than systemic contrast; for a discussion, cf. Vehmas-Lehto, forthcoming). It is clear, however, that most of the explicitation tends to be related to systemic contrast: although the results were not quantified according to the exact type of explicitation, the corpus evidence suggests that systemically conditioned explicitation is most common in the investigated category, which was to be expected considering the contrastive differences between the two languages. Predictable sources of over-representation are their respective relative pronoun systems and usage norms for non-finite clauses.

From the point of view of over-representation, the different senses in which those instances of the retrieved multi-word pattern that represent coordinated relative clauses can be said to be specific to translation may be tentatively outlined as follows:

In those cases where a TL coordinated relative clause is a rendering of a corresponding SL coordinated relative clause, the contribution of this relation to an increased frequency of the TL coordinated relative clause in the translated part of the corpus can be attributed to SL transfer. Consequently, a high frequency of SL coordinated relative clauses has a direct influence on TL frequency.

In those cases where a TL coordinated relative clause is a rendering of an SL construction that is not allowed by the TL system or would be stylistically inadequate if closely translated, the contribution of this relation to an increased frequency of the TL coordinated relative clause in the translated part of the corpus can be attributed to translational shifts carried out because of systemic linguistic contrast or stylistic incompatibility. (There is a cline of acceptability between systemic linguistic contrast and stylistic incompatibility rather than an absolute borderline.)

In those cases where a TL coordinated relative clause is a rendering of an SL construction that is allowed by the TL system but where the TL clause is more explicit than the SL construction through the realization of a coordinator and / or a relative word, the contribution of this relation to an increased frequency of the TL coordinated relative clause in the translated part of the corpus can be attributed to other factors than the above, such as target community norms, factors inherent in the translation process or individual translator preference.

Co-textual factors, finally, may be seen as cutting across these three *transfer – systemic/stylistic – other* categories, in that some translational renderings indirectly give rise to ensuing coordinated relative clauses. As

demonstrated above, there are several such indirect, co-textual factors – SL premodifying adjectives, postmodifying *of* constructions and non-finite clauses being translated as TL relative clauses – that contribute to an increased frequency of coordinated relative clauses in the translated texts.

Bibliography

- Altenberg, Bengt, Aijmer, Karin & Mikael Svensson (2001). *The English-Swedish Parallel Corpus (ESPC). Manual of enlarged version*. On line at: <http://www.englund.lu.se/research/corpus/corpus/esp.html> (consulted 09.12.2002)
- Baker, Mona (1996). "Corpus-based translation studies: The challenges that lie ahead." Harold Somers (ed.) (1996). *Terminology, LSP and Translation. Studies in language engineering in honour of Juan C. Sager*. Amsterdam/Philadelphia: Benjamins.
- Blum-Kulka, Shoshana (1986). "Shifts of Cohesion and Coherence in Translation." Juliane House & Shoshana Blum-Kulka (eds) (1986). *Interlingual and Intercultural Communication. Discourse and cognition in translation and second language acquisition*. Tübingen: Narr, 17-35.
- Gellerstam, Martin (1986). "Translationese in Swedish Novels Translated from English." Lars Wollin & Hans Lindquist (eds) (1986). *Translation Studies in Scandinavia*. Malmö: Liber, 88-95.
- Kjellmer, Göran (1987). "Aspects of English collocations." Willem Meijs (ed.) (1987). *Corpus Linguistics and Beyond. Proceedings of the Seventh International Conference on English Language Research on Computerized Corpora*. Amsterdam: Rodopi.
- Laviosa, Sara (1997). "How Comparable Can 'Comparable Corpora' Be?" *Target* 9, 289-319.
- Laviosa, Sara (1998). "Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose." *Meta* 43. On line at: <http://www.erudit.org/revue/meta/1998/v43/n4> (consulted 09.12.2002)
- Nilsson, Per-Ola (2002). "Investigating characteristic lexical distributions and grammatical patterning in Swedish texts translated from English." Andrew Wilson et al. (eds) (2002). *A Rainbow of Corpora: Corpus Linguistics and the Languages of the World*. Munich: Lincom-Europa.
- Olohan, Maeve (2001). "Spelling out the optionals in translation: a corpus study." Paul Rayson et al. (eds) (2001). *Proceedings of the Corpus Linguistics 2001 conference*. Lancaster: UCREL, 423-432.
- Renouf, Antoinette & John Sinclair (1991). "Collocational frameworks in English." Karin Aijmer & Bengt Altenberg (eds) (1991). *English corpus linguistics: studies in honour of Jan Svartvik*. London: Longman, 128-143.
- Teleman, Ulf et al. (1999). *Svenska akademiens grammatik, 4: Satsler och meningar* [Swedish Academy Grammar, Volume 4.]. Stockholm: Svenska akademien.
- Tognini Bonelli, Elena (2000). "Functionally Complete Units of Meaning across English and Italian: Towards a Corpus-driven Approach." Sylviane Granger & Bengt Altenberg (eds) (2000). *Lexis in Contrast*. Amsterdam/Philadelphia: Benjamins.
- Vehmas-Lehto, Inkeri (forthcoming). "Translation-proper explicitations – a fallacy or reality?" To appear during 2002.
- Øverås, Linn (1998). "In search of the third code: An investigation of norms in literary translation." *Meta* 43. On line at: <http://www.erudit.org/revue/meta/1998/v43/n4/003775ar.pdf> (consulted 09.12.2002)

¹ “Multi-word pattern”, which is a concept at a fairly low level of abstraction, may also be defined in slightly more restricted terms as “collocation” – one definition of a collocation is that it is “a sequence of words that occurs more than once in identical form (...) and which is grammatically well-structured” (Kjellmer 1987, quoted in Renouf & Sinclair 1991:128).

² There is no corresponding frequency difference for the two-word pattern in the non-fiction part of the ESPC, where the non-translated subcorpus has 95 instances and the translated subcorpus 106.

³ Each example is accompanied by a code which indicates the specific text from which it is taken. A list of codes and texts, including the names of authors and translators, can be found at <http://www.englund.lu.se/research/corpus/corpus/webtexts.html>.

⁴ The Swedish translated fiction subcorpus is around 10 % larger than the Swedish non-translated fiction subcorpus, which means that the frequencies are roughly comparable.

⁵ The remaining instances – 10 in non-translated texts and 18 in translated texts – are examples of other structures (e.g. cases where *som* is a conjunction, as in *och som sådant*; ‘and as such’). These are not treated here.