



HAL
open science

Typological differences shining through: The case of phrasal verbs in translated English

Bert Cappelle, Rudy Look

► **To cite this version:**

Bert Cappelle, Rudy Look. Typological differences shining through: The case of phrasal verbs in translated English. Gert De Sutter; Marie-Aude Lefer; Isabelle Delaere. Empirical Translation Studies. New Theoretical and Methodological Traditions, Mouton de Gruyter, pp.235-264, 2017. hal-01496966

HAL Id: hal-01496966

<https://hal.science/hal-01496966>

Submitted on 28 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Typological differences shining through: The case of phrasal verbs in translated English

Bert Cappelle and Rudy Loock

Abstract

Are phrasal verbs less numerous in English translations if the source language is a Romance language than if the source language is a Germanic one? This chapter sets out to answer that question. In a subcorpus of English fictional texts translated from Romance languages, *up*, *out* and *down*, which represent phrasal verb use rather well, are indeed underused when compared with non-translated English fiction from the British National Corpus, while no significant difference is to be found for this set of items between non-translated English and English translated from Germanic languages. This finding is strong evidence for source-language interference, as Romance languages on the whole do not have close equivalents to phrasal verbs, while Germanic languages do. This effect appears stronger than any source-language-independent translation universal that could in principle have played a role, such as normalization (exaggerated use of phrasal verbs, which are typical of the English language) or levelling-out (avoidance of phrasal verbs, which are generally felt to be rather colloquial). A comparison of French prefixed verbs with morphologically simplex ones in *Le Petit Prince* further shows that the former are more likely to be translated by phrasal verbs than the latter, again supporting source-language influence, as phrasal verbs resemble prefixed verbs in being composed of a verb and an added element. Our study thus stresses the relevance of taking into account typological differences (and similarities) between source and target language in translation studies.

1. Introduction

1.1. Phrasal verbs

English would be a rather different language without combinations such as *give up*, *cool down*, *run away*, *throw up*, *figure out*, *rub off*, and so on. Phrasal verbs, also known as verb-particle combinations, are among the most frequent constructions in English (Biber et al. 1999, Gardner and Davies 2007), which is why there are countless books, websites and apps on the English language learning market that encourage learners to develop their ‘phrasal verb skills’. In a recent monograph on the diachronic tracing of the construction back to Germanic and Proto-Indo-European preverbs, Thim (2012: 244) states that “throughout the history of English, phrasal verbs have always had a place in the ‘common core’ of the language”.

Not only are they undeniably central to the English language but many of them are also clearly situated towards the colloquial end of the formality spectrum: consider for example *bum around*, *freak out*, *piss off*, not to mention combinations that might be rather unfit for print. On the basis of corpus-evidence, Biber et al. (1999: 409) conclude that “[o]verall, conversation and fiction show much greater use of the most frequent phrasal verbs than news and academic prose.” One can easily confirm Biber et al.’s findings based on the Longman Spoken and Written English Corpus with data from the British National Corpus (Davies 2004-), which reveal that particles in conversation are more than ten times as frequent as in medicine-related academic texts and still more than five times as frequent as in administrative writing. Marks (2005: s.p.) sums up “widespread popular wisdom about phrasal verbs among learners and teachers”, namely “that they are colloquial, casual, informal, characteristic of speech rather than writing and perhaps even a bit sloppy or slovenly, uneducated and not quite proper” [bullets from the original article removed – B.C. and R.L.].¹ Such characterizations are not new. They are echoes from style commentators in the eighteenth century. For example, the famous lexicographer Samuel Johnson, in his *Dictionary of the English Language* (1755), already provided labels such as “improper”, “low”, “common”, “barbarous”, “familiar”, “vulgar” or “less elegant” to some of the phrasal verbs he attested (cf. Wild 2010: 207).

1.2. Intra-language differences, translation universals and source-language interference

The two properties of phrasal verbs as a class – their high frequency and their informality – accords to them the status of ideal test object to evaluate the validity of certain alleged ‘translation universals’ (for introductions of which, see Baker 1993, Halverson 2003, Mauranen and Kujamäki 2004), which generally come in contradiction with the idea that the source language interferes with the linguistic characteristics of target texts. Indeed, since the advent of corpus-based translation studies in the 1990s, researchers have agreed on the existence of the third code, that is, differences between original and translated language (intra-language differences) but have been divided over the way these should be accounted for. While some argue in favor of the existence of translation universals, those “features which typically occur in translated text rather than original utterances and which are not the result of interference from specific linguistic systems” (Baker 1993: 243), others consider intra-language differences as resulting from source language interference. In the literature, most researchers seem to support the translation universals hypothesis (e.g. Baker and Olohan 2000, Olohan 2003, Jiménez-Crespo 2011, Laviosa 1998, Øverås 1998) but other researchers have claimed that source-language interference, also known as ‘shining through’ (Teich 2003), must play (at least) a role (e.g. Koppel and Ordan 2011, Cappelle and Loock 2013). Of course, one might wonder whether interference is not itself a kind of translation universal, albeit one of a very different nature from that of all the others.²

Although initial studies provided evidence for the influence of translation universals, more recent studies have clearly questioned their role in the differences that can be observed between original and translated language (see e.g. Becher 2010, Lind 2007 or Kruger and Van Rooy 2012). Their very existence is controversial and many suggested universals have been severely criticized these last ten years (e.g. Becher 2010, House 2008, Corpas Pastor et al. 2008, Mauranen and Kujamäki 2004), with the term ‘universal’ itself being questioned (see Lind 2007). Among the criticism against translation universals is the fact that source languages are not taken into consideration and that many of the studies based on intra-language differences to prove their existence focus on only one genre, very often literary texts (see e.g. Becher 2001 for a criticism of corpus studies based on the Translational English Corpus) and with (British) English as either source or target language.

Recent studies, based on genre-controlled corpora (see Lefer and Vogeleer 2013 for a special issue on this question for the normalization universal), do show that results are genre-sensitive, and that so-called translation universals might not be so universal. For instance, Delaere et al. (2012) have investigated the use of standard and non-standard Belgian Dutch in translated vs. original texts for several registers (fiction, non-fiction, press texts, administrative texts, external communication) and have shown that intra-language differences are “text type dependent, as some of the results (partially) confirm the general trend of translations being more standardized than non-translation (fiction, external communication, administrative texts) but other results do not (journalistic texts and non-fiction)” (Delaere et al. 2012: 220). Also, Delaere and De Sutter (2013) have investigated whether the normalization universal (which they relate to “risk-aversion behavior”) is source-language and/or register-dependent: their results show that the tendency of translators to normalize Dutch translated texts is not only the result of the translation process. They show that several other factors interact: register, source language, and target audience. For instance, their results show that translators resort more to Standard Dutch in journalistic texts (register effect), but in administrative texts, it is the source language (English vs. French) that has an impact. With a special emphasis on another geographical variety of English (South African English) and using a multi-register corpus, Kruger and Van Rooy (2012) have shown that when tested together and not in isolation, linguistic features associated with translation universals fail to discriminate between original and translated texts. They investigated (i) explicitation, associated with the frequency of optional complementizer *that*, contractions and linking adverbials; (ii) normalization, associated with the frequencies of lexical bundles, coinages and loanwords; and (iii) simplification, associated with lexical diversity and mean word length. What their results, based on a comparable corpus of original and translated South African English, show is that with the exception of two of them (optional *that* and lexical diversity), the linguistic features that were investigated do not show significant differences between original and translated texts. More importantly, their study shows that cross-register differences are also present in their corpus of translated English, suggesting that register variation *is* present in translated language, *contra* what the leveling out universal might suggest.

As far as the influence of the source language is concerned, some studies have shown that the linguistic characteristics of translated texts differ depending on the source language. For instance, Cappelle (2012) shows that fewer manner-of-motion verbs are to be found in English translated from French than in English translated from German, a result that is inter-

preted as resulting from the typological difference between the two source languages, that is the fact that French, like most Romance languages, belongs to the typological group of verb-framed languages, while English and German, like the other Germanic languages, are satellite-framed languages (see below). Among the other studies that have argued in favor of source language interference instead of translation universals to account for observed differences between translated and original language is Dai and Xiao (2011), who have shown that Chinese translated from English contains much more occurrences of the passive voice than original Chinese. Given the fact that the passive voice is much more frequent in translated than in original Chinese, Dai and Xiao conclude, through the observation of their corpus, that more than 8 out of 10 occurrences of the passive voice in translated Chinese result from the interference of the source language. In a similar way, Cappelle and Loock (2013) have shown that English translated from French contains fewer occurrences of existential *there*-constructions (e.g. *there is a dog in the garden*) than original English while French translated from English contains more occurrences of existential *il y a*-constructions (e.g. *il y a un chien dans le jardin*), although the two constructions are translationally equivalent. Once again, in the light of cross-linguistic results obtained for original English and original French (existential constructions are much more frequent in original English), Cappelle and Loock suggest a strong case of interference, as results for translated English and translated French point to opposite directions.

1.3. Aim of the paper

Our paper aims to adopt the same kind of approach as the one to be found in Cappelle (2012), for phrasal verbs this time, so as to check whether their frequency in English translated texts differs depending on the source language, specifically the source language family (Romance vs. Germanic), or whether results are homogeneous, paving the way for an interpretation related to the influence of translation universals. Depending on whether one selects, say, *normalization* (“the tendency to exaggerate the features of the target language and to conform to its typical patterns” (Baker 1996: 183)) or, rather, *levelling out* (“the tendency of translated text to gravitate towards the centre of a continuum” (Baker 1996: 184)) as one’s translation universal of choice, the frequency of occurrence of phrasal verbs in translated (i.e. non-original) English may be hypothesized to be either enhanced or reduced. Indeed, on the one hand, normalization manifests itself in a boost of

common target language features, and one could therefore predict translated texts to display an over-use of phrasal verbs, which are very common items in English. On the other hand, levelling out is the tendency of translated texts to share similar characteristics, thus eradicating any register/genre-related differences, and since phrasal verbs are generally considered to belong to more informal registers, we should expect fewer phrasal verbs in translated texts. That is to say, if a translator aims to make a text sound like a genuinely English one, s/he might use many phrasal verbs, leading to an overrepresentation of phrasal verbs in translated English compared to original English – the effect, as Lind (2007) puts it, that “translated texts, like converts, are more normal than normal”. Alternatively, if a translator is wary of using expressions that are felt to be too colloquial, s/he might use more neutral, morphologically simple verbs over the generally more informal phrasal verbs, leading to an underrepresentation of the latter in translated English compared to original English. We perhaps make it appear here as if the translator makes these decisions consciously, but this need not be the case. Translators are not necessarily aware of selecting options that could be analysed as particularly target-like, including items that sound informal or even slangy or, conversely, of producing a lexically more ‘cautious’ text, whose edges have been smoothed, so to speak.

We strongly believe that contrastive issues should be taken into account in comparing translated and non-translated variants of a single language (Cappelle 2012, Cappelle and Loock 2013, Loock, De Sutter and Plevoets 2013). In the present paper, therefore, we distance ourselves from hypothesized universal laws of translational behaviour (whether consciously applied or not) which do not take the relation between the source language and the target language into consideration.³To the extent that normalization and levelling out can be formulated, respectively, as “make the translation sound like a typical, authentic text in the target language” and “use lexical and grammatical items that fall within the conventional core of the target language’s lexico-grammar”, they are what Chesterman (2004) calls T[target]-universals, which are concerned with the comparison of translations in a language to other texts in the same language. In other words, they do not make reference to differences between the target language and the source language. Yet, if we want to find out whether translated English uses more or fewer phrasal verbs than non-translated English, we should not ignore this important question: How readily does a phrasal verb present itself as a translation of *what is in the source text*?

We will now refine this question against the background of Cappelle’s (2012) study, which linked the underuse of a closely related target-language

feature in English, namely manner-of-motion verbs, to a typological difference between the source language and the target language (section 2). We will then present two corpus-based studies, one in which translated English from two typologically different source language families (Romance vs. Germanic) is compared to non-translated English (section 3) and another in which we look at which kinds of verbs in a French text are likely to tempt the translator to use a phrasal verb (section 4). In our conclusion, we will sum up the main findings and present some methodological reflections (section 5).

2. Framing typology and *translationese*

In Cappelle (2012), it was shown that translated English contains fewer manner-of-motion verbs when the source language is French than when the source language is German. The most likely reason for this, it was suggested, is that French, like most Romance languages, belongs to the typological group of verb-framed languages, while English and German, like the other Germanic languages, are satellite-framed languages (Talmy 2000). This typological difference pertains to whether the most central semantic aspect in an event of change of position or state, among other kinds of events, is preferentially expressed by the verb root or by a sister to it – a prefix or a particle, for instance. For the expression of a motion event, the core semantic element is the so-called path, which refers to direction, source or goal. Thus, in the English sentence *A UFO whizzed by*, it is the particle *by* that encodes the path, just like the separable verb prefix *vorbei* in the German equivalent sentence *Ein UFO sauste vorbei*. Since the verb root itself does not have to express this aspect of meaning, it is ‘freed up’ to encode a more secondary semantic aspect: in this case, the high speed of motion and the accompanying sound this produces. In French, by contrast, as in Romance languages more generally, the direction of motion is typically expressed by the main verb, as in *Un OVNI passa* (‘A UFO passed’, ‘A UFO went by’). The verb, having been assigned with the task of encoding this aspect of meaning, can then no longer express manner of motion, which would then be expressed, if at all, in a constituent functioning as adjunct, such as *à grande vitesse* ‘with high speed’ or *dans un sifflement* ‘with a whistling sound’.

What emerged from that corpus investigation is that the difference in overall framing preferences appears to leave its trace in a corpus of English translations from French. Specifically, this corpus exhibited a higher pro-

portion of path verbs (*leave, rise, etc.*) to manner-of-motion verbs (*crawl, leap, etc.*) than both a reference corpus of non-translated English texts and a corpus of English translations from German. Though this difference may not be noticeable in a single individual text, the cumulative evidence appearing from corpora of translated and non-translated English cannot be ignored: the underuse of manner-of-motion verbs in translations from French is an example of source-language interference.

In what follows, we will attempt to answer the following questions:

1. a. Do English translations from Romance languages contain as many phrasal verbs as texts originally written in English?
- b. Do English translations from other Germanic languages contain as many phrasal verbs as texts originally written in English?

In light of the typological differences between Germanic and Romance languages briefly reviewed above, we predict that the first question will be answered negatively and the second positively. Such findings from translation studies would complement what we know from second language acquisition studies, where it appears that phrasal verbs are underproduced by undergraduate learners whose L1 lacks a similar category (e.g. Hebrew, Chinese) but not by undergraduate learners whose L1 has a close equivalent (e.g. Dutch, Swedish) (Laufer and Eliasson 1993, Liao and Fukuya 2004). In addition, if our expectation for the first sub-question (1a) is borne out, we will address the following question:

2. Assuming that English translations from Romance languages are not completely devoid of any phrasal verbs, which source language expressions are they the translations of?

For this second question, which we will answer with reference to French as a source language, we hypothesize that morphologically complex verbs in the source text (e.g. *re-venir* ‘go back’) are more likely to be translated by a phrasal verb than simplex verbs (e.g. *sortir* ‘leave’, ‘go out’). Such a result would again be in line with our assumption that what is in the source text may be ‘shining through’ in the translation, to use Teich’s (2003) concept, as the presence of a bound derivational morpheme in the French source text could be seen as a trigger for a translation into a particle to form a phrasal verb in the English translation. For English as a source language, previous studies demonstrated such a structure-preserving effect. In a study on the translation of English phrasal verbs into German, Claridge (2002) found that translators typically used structures resembling the source forms, trans-

lating the verb by a literal equivalent and translating the particle by a separable or inseparable prefix. In English translations into Russian, too, phrasal verbs are often translated by prefixed verbs (Mudraya et al. 2005). German and Russian are typologically similar to English in that they make extensive use of particles or prefixes. In other words, they have a frequently used equivalent to phrasal verbs. While French is not considered to be a satellite-framed language, it does contain prefixed verbs and these may account to a large extent for the use of phrasal verbs in the translation. It has been shown in previous research (Paillard and Videau 2008) that the most frequent translation strategy for French verbs starting with *dé-* or *de-* is a phrasal verb in English. Such a finding lends support to the correctness of our hypothesis, but to have full certainty, we should check whether prefixed verbs in French more often lead to a phrasal verb in the target text than non-prefixed ones do.

Throughout this paper, we assume that phrasal verbs are, by and large, satellite-framed structures. This assumption is justified insofar as particles frequently express the path in a motion event (e.g. *walk in*) or a more metaphorical or abstract change of state in other events (e.g. *cool down*, *wake up*). Even in aspectual combinations like *play along*, the particle encodes an aspect of meaning which in verb-framed languages would more typically be expressed in the main verb, for instance as something close to *accompany someone playing*. Only in highly idiomatic combinations (e.g. *make out* ‘kiss in a sexual way’) is it less clear that the particle expresses the core semantic part of a complex event. Even so, it can hardly be denied that phrasal verbs are compatible with, and indicative of, a language’s satellite-framed typological nature.

3. Phrasal verbs in translated English from typologically different languages: A large-scale quantitative corpus-based approach

3.1. Method

3.1.1. Overall design

The methodology adopted here involves three corpora: (i) a reference corpus of English texts, (ii) a comparable corpus of English texts translated from source languages A1, A2, A3, etc. and (iii) another comparable corpus

of English texts translated from source languages B1, B2, B3, etc. Here, as opposed to Cappelle (2012), we do not focus on the proportion of manner-of-motion verbs to path verbs but on the frequency of occurrence of particles relative to corpus size. Another difference is that we here extend the comparable corpora considerably: we do not just look at translations into English from French and German but at translations into English from Romance languages and Germanic languages.

3.1.2. *Corpora*

Our corpus study rests on an analysis of fictional texts in original and translated English. The reason why we have restricted our analysis to fictional texts is (i) their availability in different corpora (as opposed to press or technical texts for instance) and (ii) the fact that phrasal verbs are frequent in this register (see above). Our reference corpus was the 100 million word British National Corpus (BNC), which we searched via the Brigham Young University web interface (Davies 2004-). We restricted our searches to the subcorpus of fictional texts, whose total size is 15,909,312 words. This component consists of 476 texts, whose average length is 34,656 words (information obtained via <http://www.natcorp.ox.ac.uk/docs/URG/BNCdes.html#wrides>). For translations into English we used the Translational English Corpus (TEC), available at <http://www.llc.manchester.ac.uk/ctis/research/english-corpus/>. From this corpus, we selected two subcorpora with English fictional texts published between 1980 and 1993, which is also the time range from which the texts in the BNC date. For one of the TEC subcorpora, we selected as source languages all available Romance languages (Brazilian Portuguese, Catalan, European Spanish, French, Italian and Latin American Spanish). This subcorpus, which we can refer to as TEC<Rom, has a total number of 1,952,690 words and consists of 32 texts, with an average length of 111,669 words (shortest: 25,915 words; longest: 197,422 words). For the other subcorpus, we aimed (somewhat overoptimistically, as we will point out shortly) at covering all available Germanic languages (Danish, Dutch, German, Icelandic, Norwegian and Swedish) in the entire TEC corpus. This subcorpus (TEC<Ger) totals 1,146,785 words and comprises 14 texts, whose average length is 81,913 words (shortest: 14,288 words; longest: 166,973 words). Because of our preselection of fiction from a specific period and because of the make-up of the TEC corpus itself, our subcorpora are not balanced for individual languages, though. TEC<Rom is made up of mainly

French (61%), with Brazilian Portuguese (12%), European Spanish (11%), European Portuguese (6%), Italian (4%), Latin American Spanish (3%) and Catalan (3%) only being minor parts of this corpus. TEC<Ger has an even less even spread over individual languages, as predominantly consists of German texts (90%), complemented by only one other Germanic language, namely Swedish (10%). Because of this scarcity of data for some Romance and especially Germanic languages, we will not report findings for individual languages. Future research should take a closer look at possible differences among the languages within the two families studied here.

3.1.3. Search items

As it would be quite unfeasible to search and check all verb-particle combinations in these corpora by hand, we carefully selected search items that could serve as proxy for the entire class of phrasal verbs. Since verbs form an open lexical class and particles a closed one, we restricted our searches to the latter. However, especially as the TEC is not tagged, we had to be careful to choose words that do indeed often occur as particles, in order to maximize precision (i.e., the fraction of retrieved items that are relevant), while also making sure that we used items that are representative of the class of particles as a whole, in order to maximize recall (i.e., the fraction of relevant items that are retrieved). The choice of suitable items was important, since many words that can be used as particles occur in fact more frequently as prepositions – as is the case for *in* and *on*, for instance. For our selection of search items, we relied on data in Gardner and Davies (2007: 346), according to which the words *out* and *up*, of all particle candidates, are not only used most frequently as particles (in 97.3% and 87.4% of their total occurrences, respectively) but are also the two most frequently used particles, together accounting for almost half (46%) of all particle occurrences in the BNC. In short, these two words have high precision and reasonably good recall values as representatives of the class of verb particles. Given the unclear grammatical status of *out of* (see Cappelle 2001), we removed this combination from our selection. In order to have a more representative sample of particles, we decided to add one more word to our set of search items, namely *down*. It appears that *down* is another frequent particle, with 79.2% of all occurrences of this item being tagged as a particle, as reported by Gardner and Davies (2007). This is still a reasonably good precision value, and with *down* added to *up* and *out*, we thus obtained a small set of items which jointly form a representative subset of the class

of particles, making up 57.3% of all particle occurrences, to be precise (still according to Gardner and Davies (2007)). So, the three search items, *up*, *out* (minus *out of*) and *down* are all predominantly used as particles and together represent well over half of the occurrences of all particles.

3.1.4. Statistical analysis

Results of the searches of the three items in the three different corpora were analysed statistically using the Log-likelihood test, a test which is similar to the Chi-square test but is not subject to certain assumptions about how the data are distributed (cf. McEnery, Xiao and Tono 2006: 55). We provide the results for the three individual particles, but as we are not interested in how the relative frequencies of individual particles may differ across corpora, we summed over the occurrences of all three test items in each of the three corpora. The results obtained in the BNC were compared with those obtained in the corpus of translations from Romance languages and separately with those obtained in the other translational corpus.

3.2. Results

In the BNC fictional subcorpus, the three search words yielded 124,529 occurrences, or 7827 occurrences per million words. In the corpus of translations into English from Romance languages (TEC<Rom), we retrieved 11,880 occurrences of these words, or only 6084 per million words. By contrast, the corpus of translations into English from (other) Germanic languages yielded 8951 occurrences of the three search items, which amounts to 7805 per million words, virtually the same frequency as in the BNC reference corpus. The details of the findings are given in Table 1 and graphically represented in Figure 1.

Table 1. Number of occurrences of *up*, *out* and *down* (raw and normalized per million words (pmw)) in the fiction component of the BNC and in two comparable corpora of translated texts into English from Romance and Germanic languages.

	BNC		TEC<Rom		TEC<Ger	
	raw	pmw	raw	pmw	raw	pmw
<i>up</i>	54850	3448	5595	2865	4169	3635
<i>out</i>	38507	2420	3747	1919	2753	2401
<i>down</i>	31172	1959	2538	1300	2029	1769
total	124,529	7827	11,880	6084	8951	7805

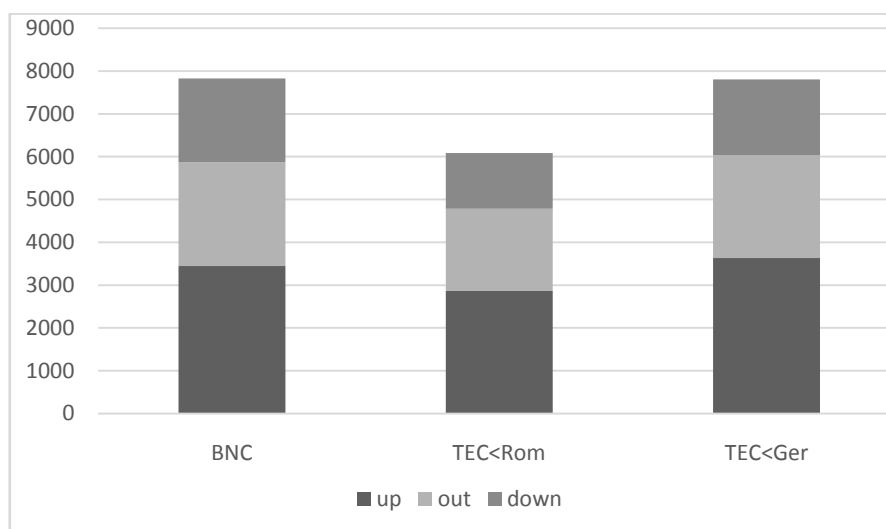


Figure 1. Number of occurrences of *up*, *out* and *down* per million words in the fiction component of the BNC and in two comparable corpora of translated texts into English from Romance and Germanic languages.

The conspicuous underrepresentation of the three search items in the corpus of English fiction translated from Romance languages, compared to their frequency in the BNC fiction component, is statistically extremely significant (Log-likelihood = 738.32; $p < .0001$). Given the large sample sizes, the small difference between the BNC results and the results from TEC<Ger were in danger of having proved statistically significant as well, but (summarized over the three particles) they are not (Log-likelihood = 0.067; $p = 0.8$).

3.3. Interim discussion

The present findings strongly suggest that an account of the differences between original and translated English which is based on translation universals, in particular normalization or levelling-out, should be treated with utmost circumspection. The results clearly differ depending on the language family to which the source language belongs. Although we have not studied the influence of each individual source language on the linguistic characteristics of the English translated texts, it is possible to consider ‘source language family interference’ as a significant effect in translation, *contra* any translation universal that is claimed to hold irrespective of the source and target language involved. Of course, one might argue that it is only to be expected that texts from the same genetic and typological family as English lead to translations that are closer to non-translated English than texts from genetically more distant and typologically rather different languages do. We fully agree that our results are in line with what was to be expected. Obviously, a finer-grained approach involving a comparison with each individual source language would be desirable, especially as not all languages within a single family may be equally strong representatives of a typological profile (see Iacobini and Masini (2003) for the occurrence of phrasal verbs in Italian, in particular). This is beyond the scope of this paper, which used the TEC to investigate translated English, as the corpus does not contain samples that are large enough for each individual source language.

Still, we would like to stress that our results allow us to cast very serious doubts on the validity of each of the two translation universals discussed in Section 1.3. Source language (family) interference has a stronger influence than these and any other translation universals that could be considered since our results are source-language-family dependent. Quite clearly, translations into English from Romance languages fail to be fully normalized with respect to the presence of phrasal verbs. While such normalization, if that is what it is, seems to be successful for translations from Germanic languages, it falls short for translations from Romance languages. In other words, it does not obtain across-the-board and, by consequence, it would seem wrong to consider normalization a translation *universal*.

It could be objected that the sheer fact that phrasal verbs are used in translations from languages that do not use them (but see Section 4 for an important qualification of this ‘absence’ in the source language) can only be analyzed as the result of a normalization tendency. Such an objection can easily be countered. While it cannot be denied that the presence of

phrasal verbs in English translated from Romance languages could be interpreted as translators' tendency to 'normalize' the target language by using a common feature of it, this does not lead to the translation sounding "more normal than normal", to use a formulation we quoted earlier. In other words, what we dismiss here is normalization as a T(arget language)-universal, that is, a source-language-independent tendency manifesting itself as a stable difference between translations in the target language and originally produced texts in that target language (Baker 1996). It would be very strange, of course, if an English translation from a Romance language contained no, or hardly any, phrasal verbs: after all, the job of the translator is to produce an English text, and phrasal verbs are an integral part of that language's lexicon. The point, though, is that normalization, as a T-universal, predicts that there will be more of these very normal lexical items in translations into English than in original English. This prediction is clearly not borne out. Note also that even with Germanic languages as source languages, there is no overrepresentation of phrasal verbs, which we should have found if normalization leads to exaggeration of target-language features.

Likewise, we can quite safely dismiss levelling-out as a translation universal (cf. again Section 1.3). If we only considered translations into English from Romance languages, this tendency could have served as a suitable explanation for the underrepresentation of phrasal verbs, but we are then left to explain why in translations from other languages, this presumed levelling out does not obtain. So, levelling out, too, loses its appeal as a source-language-independent translation universal. Since our results are source-language-family dependent, *any* such translation universal that could be at play is undermined by interference from the source language.

There just might be a way of salvaging normalization and levelling-out. While, as we saw above, they lead to opposite predictions in the case of phrasal verbs (overrepresentation and underrepresentation, respectively), we could say that, when they are both at work (as they should be, given that they are universals), they cancel each other out. That is, using phrasal verbs in an over-indulgent way (i.e., normalization/exaggeration) may be tempered by the simultaneously operative avoidance of lexico-grammatical structures, including phrasal verbs, that have a too informal ring about them (i.e., levelling-out), and vice versa. The results that we obtained above (underrepresentation in translations from Romance languages; no noteworthy difference in translations from Germanic languages) may be compatible with such neutralisation. They could then be considered, crucially, as the added effect of just one more 'universal' thrown in, namely Toury's (1995: 275) "law of interference", or Tirkkonen-Condit's (2004) "Unique Items

Hypothesis”, according to which linguistic features that are typical of, or even unique to, the target language, when compared to the source language, have a tendency to be under-represented in translations. Even if there is no neutralization at work of any other universals, if they exist at all, the law of interference or the Unique Items Hypothesis can definitely *not* be rejected, in the light of our findings. The law of interference covers “phenomena pertaining to the make-up of the source text [that] tend to be transferred to the target text” (Toury 1995: 275). Toury’s definition could—with the help of some mental gymnastics, admittedly – be applied to cases where the phenomena (here: phrasal verbs) are actually absent from the source text. Indeed, one could argue that the absence of phrasal verbs is part of the make-up of source texts written in (most) Romance languages. Because of this absence in the source language, a more useful concept here is “Unique Item” in the target language. After all, particles, with the exception of Italian, are unique to English as a target language when the source language belongs to the Romance family. If we now adhere to the strictly scientific principle of not multiplying hypotheses beyond those that are needed to explain the observable facts, we can let Occam’s razor do its work and simply dismiss normalization and levelling-out – prime examples of translation universals – as having a major impact on the occurrence of phrasal verbs in translated English.

4. Phrasal verbs in translated English from French: Comparing *Le Petit Prince* and its English translation

In a follow-up study, we tried to answer the following question: If there is a phrasal verb in a text translated into English from a Romance language, what corresponds to this item in the source text? The motivation for asking this question is that Romance languages, with the exception of Italian, do not have any phrasal verbs. If such structures do end up in the translation, this could be evidence that there still is some degree of normalization at work, albeit not to the extent that target-language features are exaggerated in the translation, as the data reported above clearly show.

Although French does not have anything that could be called ‘phrasal verbs’, it does make use of morphologically complex verbs (e.g. *re-venir* ‘come back’, *sur-voler* ‘fly over’) which, just like phrasal verbs in English, belong to the set of satellite-framing structures (Kopecka 2006, Pourcel and Kopecka 2006), as these prefixed verbs can be decomposed into an element

referring to a path (the prefix) and an element referring to (manner of) motion as such. In other words, such verbs have two semantic components, in much the same way that English motional phrasal verbs have. The hypothesis we can therefore formulate is that such complex morphological items in the source text are more likely to act as a trigger for the use of a phrasal verb in the target text than morphologically simplex items in the source text.

4.1. Method

4.1.1. Overall design

For this part of our study, we needed to take a look at what is actually there in the source text, something we did not do in the corpus study reported on in Section 3, where we merely knew that the source text belonged to either a Romance language or a Germanic language. We opted to use a single relatively short text in French and its English translation, the idea being that this research leans rather more to qualitative research, possibly at the expense of full representativeness. The verbs used in the source text were classified as either morphologically simplex (non-derived) or morphologically complex (containing a prefix). We identified the phrasal verbs in the translation with the aim of finding out whether, as we expect, there are more of them whose source expression is a morphologically complex verb than can be expected on the basis of the overall frequency of complex verbs, relative to simplex verbs, in the entire source text.

Thus, suppose there are 900 simplex verbs and 100 complex verbs in the source text and suppose, furthermore that there are 100 phrasal verbs in the translation, then in the still imaginary case that all of these 100 phrasal verbs would have as their source expression the 100 complex source verbs, there would be an extreme case of source language interference. By contrast, if 90 of the 100 phrasal verbs had a simplex verb as their source expression and the other 10 had a complex verb as their source expression, there would be no association whatsoever between the morphological complexity of source verbs and their translations' membership to the class of phrasal verbs.

4.1.2. Corpus

Within a pilot study, the text we took a closer look at is the well-known children's book *Le Petit Prince* by Antoine de Saint-Exupéry and its English translation *The Little Prince* by Katherine Woods, both of which appeared in 1943. The reasons why we selected this text and its translation, which we realize are not very modern ones, are that (i) the original text is widely available and has been translated into many languages, such that our results for the English translated could be compared with those for other target languages and (ii) this is a children's book, for which one can expect there to be a large amount of dynamic scenes, including motion, rather than merely reflective passages. Easily obtainable electronic versions of these texts were sentence-aligned automatically with the help of PlusTools, a Windows Word plugin. The aligned output was then corrected manually. The original text is 14,952 words long and its translation 17,066 words long.

4.1.3. Coding

A master's student from the University of Lille 3, whose mother language is French, was instructed to classify all lexical verbs in the source text as either morphologically simple (disregarding inflectional endings) or morphologically complex (specifically, having a prefix). Deciding whether a verb is morphologically complex or not is not always easy. The student was asked not only to consult an etymological dictionary (namely, the online database Ortolang, <http://www.cnrtl.fr/etymologie/>) but also to use her best judgement in classifying a verb as multi- or monomorphemic but to be consistent in her choices. For instance, the verb *comprendre* 'understand' was morphologically complex in Latin but is no longer felt to be so for contemporary speakers of French and therefore was classified as simplex. By contrast, verbs like *a-néantir* 'destroy' *dé-ranger* 'disturb', *r-assurer* 'reassure' were classified as complex. The analysability of the verbs for present-day speakers of French was used as a criterion to determine the simplex or complex nature of the verbs. A further instruction was that extremely light verbs, such as linking verbs and auxiliaries (e.g. *être* 'to be', *pouvoir* 'can'), should be removed from the analysis.

Semi-automatic identification of phrasal verbs in the target text was carried out by ourselves. We did not do this fully automatically, because we did not have a tagged version of the text and we wanted to make sure that

we only retrieved occurrences of *on*, *off*, *through*, etc. in their use as particles, not as prepositions. Note that the particles *up*, *down* and *out* taken together have 93 occurrences in the translation of *Le Petit Prince*; 93/17,066 equals a normalized frequency of 5499 per million words. This compares rather well to the normalized frequency of these lexical items grouped together in TEC<Rom (which, as we have seen, is 6084 per million words) and confirms the finding from the preceding section that English translations from Romance languages underuse phrasal verbs. We also determined the category of these phrasal verbs' source expressions (simplex or complex). We realize that the analysability of prefixed lexemes in French is an extremely complex issue, and that there is no such thing as a homogeneous French audience; therefore, the MA student's and our assessment of prefixed verbs as being synchronically polymorphemic or not may not generalize perfectly to all speakers of French. According to standard morphological analysis, word parts should be endowed with meaning in order to be identified as morphemes in the language. In the present analysis, this was thought to be the case for *é-* in *s'écrier* 'cry out', for instance, but it is not clear that all speakers of French perceive this lexeme's complexity. In Appendix A, we list all contexts in the source text containing French verbs which were considered to be prefixed and having a phrasal verb in the English translation. In Appendix B, we list all further verb forms in the source text that were considered to be prefixed, but whose translation was not a phrasal verb.

4.1.4. Statistical analysis

All the simplex and complex source verbs were counted by the same master's student who did the coding, while we counted the phrasal verbs which had a simplex verb and those which had a complex verb as a source expression. In the English translation of *Le Petit Prince*, there are 149 phrasal verbs, which were classified depending on their source language structure: 'simplex' (e.g. *find out*<*savoir*), '(morphologically) complex' (e.g. *go on*<*repandre*), 'syntactically complex' (e.g. *put off*<*remettre à plus tard*), 'nothing' in cases of explicitation (e.g. *go on*< \emptyset) and 'other' (e.g. *a month has gone by*<*ça fait déjà un mois*). We disregarded the last three cases, in other words, we retained only those phrasal verbs that corresponded with a simplex or a morphologically complex verb in the source text. We converted the results into a 2x2 contingency table, on which we performed the Chi-Square Test of Association (with Yates correction) using the web cal-

culator developed by Richard Lowry available at <http://vassarstats.net/tab2x2.html>. In addition, the Fischer Exact Probability Test was performed via GraphPad's online application available at <http://graphpad.com/quickcalcs/contingency2/>.

4.2. Results

Out of the 150 phrasal verbs in the translation, 98 have a morphologically simplex verb and 30 have a morphologically complex verb in the French source text (with 12 phrasal verbs being the translation of a syntactically complex structure, 5 having nothing directly corresponding with them in the source text and another 5 corresponding to other structures). In the entire source text, there were another 1592 simplex verbs and another 123 complex verbs whose translation did not correspond with a phrasal verb in the target text. These results are shown in Table 1 (with added percentages based on raw totals) and represented graphically by means of two pie charts in Figure 2 below:

Table 2. Distribution of phrasal verbs versus other translations of simplex versus complex verbs in *Le Petit Prince*.

	Phrasal verb as translation	Other translation
Simplex source verb	98 (6%)	1592 (94%)
Complex source verb	30(20%)	123 (80%)

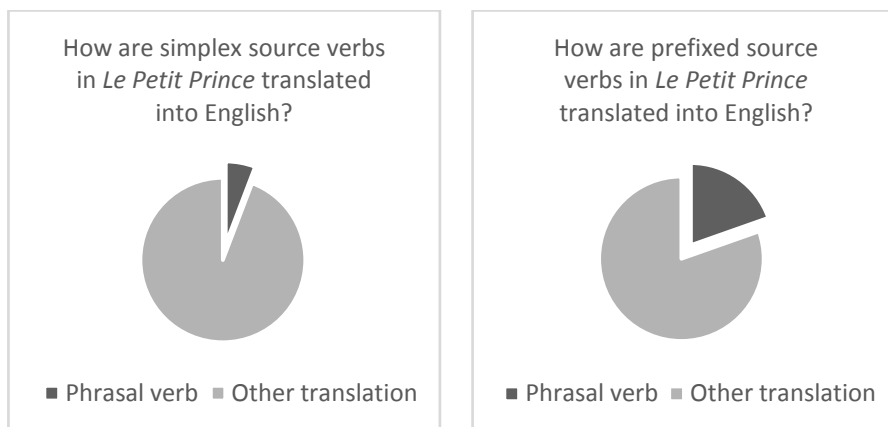


Figure 2. Morphologically simplex versus morphologically complex verbs in *Le Petit Prince* and the share of phrasal verbs corresponding to them in the English translation

Clearly, when the verb is morphologically complex in the French source text, it is much more likely to be translated with a phrasal verb in the English translation than when that source verb is simplex. The observed difference is statistically extremely significant ($\chi^2 = 39.287$; $p < .0001$). Fisher's Exact Probability Test also produces a two-tailed p-value smaller than 0.0001, confirming that association between kinds of source verbs (simplex / complex) and translations (phrasal verbs / other) can be considered to be extremely statistically significant.

4.3. Discussion

The English translation of *Le Petit Prince* contains over a hundred phrasal verbs, even though there were no phrasal verbs in the source text, French being a language that lacks such a structure. This could be taken as evidence that there was some degree of normalization at work. Yet, such a conclusion is trivial. After all, it is nothing more than a translator's job to render a text into the target language and if that target language makes frequent use of phrasal verbs, then the occurrence of such structures is not at all surprising. Moreover, as we saw before, translations into English from Romance languages do not contain more phrasal verbs than non-translated English, but to the contrary.

If phrasal verbs are used in the translation when there is no phrasal verb in the source text, what prompts the translator to introduce these 'unique

items’? We surmised that, again, there may be some source-language influence at work. When the source text contains an expression that displays a satellite-framed encoding strategy, whether for motion or a semantically more abstract (e.g. resultative) event, then the distance to a phrasal verb as translation is somehow smaller than when there is no satellite-framed item in the source text. That is, a prefixed verb in French leads more easily to a phrasal verb in English, since at an underlying cognitive level, these structures show crucial similarities. Table 3 below provides a sample of prefixed verbs in *Le Petit Prince* and their phrasal verb translation in *The Little Prince*.

Table 3. Sample of morphologically complex verbs in *Le Petit Prince* and their corresponding translation involving a phrasal verb. All prefixes are underlined, as are English particles that are semantically equivalent to them.

<i>s'<u>é</u>crier</i> → cry <u>out</u>	<i>s'<u>en</u>hardir</i> → pluck up courage
<i><u>re</u>faire</i> → do <u>over</u>	<i>s'<u>en</u>fermer</i> → shut oneself up
<i><u>rap</u>porter</i> → bring <u>back</u>	<i><u>re</u>prendre</i> → go on
<i><u>re</u>venir</i> → come <u>back</u>	<i><u>pours</u>uivre</i> → go on

As Table 3 shows, however, only in some cases does the prefix in French actually correspond with a semantically equivalent particle in English. Often, the particle is not at all any close translation of what the prefix expresses (cf. also Paillard and Videau 2008). For instance, in *s'enfermer*, *en*-basically means *in* but this reflexive verb is translated by *shut (oneself) up*, not by *shut (oneself) in*. Thus, it may be that the translator noticed something semantically complex in the source language and was then, at some unconscious level, encouraged to render this complex item by a phrasal verb in the target language, without necessarily feeling the need to preserve the internal semantic makeup of the source item in the translation.

5. Conclusion

In this paper, we have shown that translations into English from Romance languages contain fewer phrasal verbs than translations into English from Germanic languages, which do not differ noticeably from non-translated English in the frequency of phrasal verbs. The most likely explanation for

this is that all Germanic languages are satellite-framed and thus frequently use structures that are very similar to phrasal verbs in English, while Romance languages are verb-framed and do not use such structures as frequently, if at all. This result casts doubt on the validity of normalization as a translation universal, as the frequency of phrasal verbs is apparently not similar to that in non-translated English, or even higher, *independently of the source language used*. For the same reason, this result also undermines the validity of levelling-out, as phrasal verbs are not underused (because of their general informality), *independently of the source language used*—indeed, they are not in the translations from Germanic source languages.

Furthermore, as we also expected, a French simplex verb is less frequently translated into English by a phrasal verb than a French prefixed verb is. The most plausible explanation of this is that French prefixed verbs are satellite-framed structures and thus structurally and especially conceptually close to English phrasal verbs, which then present themselves more easily as translations than in the case of simplex source verbs, even if the prefix is not necessarily rendered by a semantically equivalent particle.

Methodologically, we would like to stress three points that may be useful for similar studies. Firstly, we have demonstrated that when facing a linguistic phenomenon that would require time-consuming manual filtering, it is possible to select a few items which can be retrieved automatically by relying on independent research which shows that these items, when used as search queries, do not yield much noise. As a consequence, especially with quite large corpora (ranging here from over a million to almost sixteen million words), one can use a few well-chosen search items that do not necessitate much, if any, filtering of noisy examples afterwards, as representatives of the phenomenon under research. In our example, corpus frequencies for the words *up*, *out* (at least after extraction of occurrences for *out of*) and *down* can be ‘trusted’ as indicators of the use of phrasal verbs.

Secondly, we have shown how a large-scale, quantitative corpus-based approach can and sometimes has to be combined with a smaller-scale, more qualitative study, where data have to be coded manually and with considerable deliberation, as was done in this study for *Le Petit Prince* and its English translation. Results obtained by such more painstaking research still allow – and should undergo – a minimum of statistical validation, if the data are sufficiently large. We hope to have shown that the combination of different methods is especially fruitful when one method drives hypotheses that can be tested with another.

Thirdly, and most importantly, we believe that the application of a certain method can never be a goal in itself. Any methodology should remain subservient to answering a question or supporting a hypothesis that helps to

advance, by however small a step, our current state of knowledge. In the present paper, we have used a rather simple method to discredit normalization and levelling-out as translation universals and to show that the typological nature of the source language ‘shines through’ in the frequency of use of phrasal verbs in English translations, thus supporting the high importance of source-language influence. We have used an equally simple method to support the linguistically relevant idea that French, in its use of prefixed verbs, is not a *perfect* verb-framed language. While these prefixes do not necessarily carry information on ‘path’ but, rather, are remnants of previous stages in which the language was satellite-framed (Latin and perhaps Old French), these vestiges of satellite-framing appear to be synchronically relevant, at least to L2 users of French translating to English, as prefixed verbs are translated more often than non-prefixed ones by phrasal verbs in English, which we have regarded in this paper as satellite-framed structures. In short, our methods have been used to some advantage to further our current knowledge in the fields of both translation studies and language typology.

Acknowledgements

We would like to thank the organizers and participants of the workshop “New Ways of Analysing Translational Behaviour in Corpus-Based Translation Studies” held as part of the 46th Annual Meeting of the Societas Linguistica Europaea at Split, Croatia (18-21 September 2013) for valuable comments and encouragement. We are grateful to Lindsay Lechnig, a master’s student at the University of Lille 3, for classifying all the verbs used in Saint-Exupéry’s *Le Petit Prince* and performing the counts on them. We also thank the editors and two anonymous reviewers for their thoughtful comments, occasional criticisms, and detailed suggestions, which helped us improve our text. Of course, any remaining errors are ours.

Appendix A: Prefixed verb contexts in *Le Petit Prince* with phrasal verbs in the English translation

Below is a list of sentences or sentence fragments with prefixed verbs in *Le Petit Prince* (with the prefix given in boldface) and corresponding passages containing a phrasal verb in the English translation (with the particle in

boldface). Note that there is not always a direct semantic correspondence between the French prefix and the English particle.

<i>Je refis donc encore mon dessin</i>	<i>So then I did my drawing over once more.</i>
<i>Alors il s'écria: [...]</i>	<i>He cried out, then: [...]</i>
<i>L'astronome refit sa démonstration en 1920</i>	<i>So in 1920 the astronomer gave his demonstration all over again</i>
<i>Et comme il se sentait un peu triste à cause du souvenir de sa petite planète abandonnée, il s'enhardit à solliciter une grâce du roi: [...]</i>	<i>And because he felt a bit sad as he remembered his little planet which he had forsaken, he plucked up his courage to ask the king a favor: [...]</i> [No direct semantic correspondence]
<i>Il faut exiger de chacun ce que chacun peut donner, reprit le roi.</i>	<i>"One much require from each one the duty which each one can perform," the king went on. [On refers to continuation after an interruption, and so has some of the resumptive meaning expressed by re-]</i>
<i>Je vais repartir !</i>	<i>So I shall set out on my way again. [Note that it is the adverb again here that captures the meaning of the prefix in the source text.]</i>
<i>acheva le buveur qui s'enferma définitivement dans le silence.</i>	<i>The tippler brought his speech to an end, and shut himself up in an impenetrable silence. [No direct semantic correspondence]</i>
<i>Le petit prince poursuivit: [...]</i>	<i>The little prince went on with his explanation: [...]</i> [No direct semantic correspondence]
<i>S'il s'agit par exemple de la découverte d'une grosse montagne, on exige qu'il en rapporte de grosses pierres.</i>	<i>For example, if the discovery in question is that of a large mountain, one requires that large stones be brought back from it.</i>
<i>On pourrait entasser l'humanité sur le moindre petit îlot du Pacifique.</i>	<i>All humanity could be piled up on a small Pacific islet. [No direct semantic correspondence]</i>
<i>Où sont les hommes ? reprit enfin le</i>	<i>"Where are the men?" the little prince at last took up the conversation</i>

petit prince.

Bonjour, répondit poliment le petit prince, qui se retourna mais ne vit rien.

Mais le renard revint à son idée:

Les autres pas me font rentrer sous terre.

Le lendemain revint le petit prince.

Il eût mieux valu revenir à la même heure, dit le renard.

Tu reviendras me dire adieu, et je te ferai cadeau d'un secret.

Et il revint vers le renard:

Ils reviennent déjà ? demanda le petit prince...

Comme le petit prince s'endormait, je le pris dans mes bras, et me remis en route.

Les hommes, dit le petit prince, ils s'enfourment dans les rapides, mais ils ne savent plus ce qu'ils cherchent.

Reviens demain soir...

Lorsque je revins de mon travail, le lendemain soir, j'aperçus de loin mon petit prince assis là-haut, les jambes pendantes.

je veux redescendre !

Tu vas pouvoir rentrer chez toi...

again. [Note that it is the adverb again here that captures the meaning of the prefix in the source text.]

*"Good morning," the little prince responded politely, although when he turned **around** he saw nothing.*

*But he came **back** to his idea.*

*Other steps send me hurrying **back** underneath the ground.*

*The next day the little prince came **back**.*

*"It would have been better to come **back** at the same hour," said the fox.*

*Then come **back** to say goodbye to me, and I will make you a present of a secret.*

*And he went **back** to meet the fox.*

*"Are they coming **back** already?" demanded the little prince.*

*As the little prince dropped **off** to sleep, I took him in my arms and set out walking once more. [No direct correspondence]*

*"Men," said the little prince, "set **out** on their way in express trains, but they do not know what they are looking for." [No direct semantic correspondence]*

*Come **back** tomorrow evening...*

*When I came **back** from my work, the next evening, I saw from some distance away my little prince sitting on top of a wall, with his feet dangling.*

*I want to get **down** from the wall. [No direct semantic correspondence]*

*Now you can go **back** home...*

<i>Moi aussi, aujourd'hui, je rentre chez moi...</i>	<i>I, too, am going back home today...</i>
<i>Quand je réussis à le rejoindre il marchait décidé, d'un pas rapide.</i>	<i>When I succeeded in catching up with him he was walking along with a quick and resolute step. [No direct semantic correspondence]</i>
<i>Il hésita encore un peu, puis il se releva.</i>	<i>He still hesitated a little; then he got up. [No direct semantic correspondence]</i>
<i>Mais je sais bien qu'il est revenu à sa planète, car, au lever du jour, je n'ai pas retrouvé son corps.</i>	<i>But I know that he did go back to his planet, because I did not find his body at daybreak.</i>
<i>écrivez-moi vite qu'il est revenu</i>	<i>Send me word that he has come back.</i>

Appendix B: Other verbs coded as prefixed in *Le Petit Prince*

Below is a list of all other verb forms coded as prefixed by the MA student; the reader may beg to differ about the prefixed nature of some of these forms, for example in the case of *ajouter*. Note, however, that eliminating tokens from this list would only make our findings stronger, not weaker, as these are the forms for which the translation does not contain a phrasal verb.

<i>ajouta</i>	<i>anéantir</i>	<i>démodent</i>	<i>émerveilla</i>
<i>ajouta</i>	<i>apaisent</i>	<i>dépendra</i>	<i>emportait</i>
<i>ajouta</i>	<i>apercevais</i>	<i>dérangé</i>	<i>emportent</i>
<i>ajouta</i>	<i>apercevrai</i>	<i>dérangea</i>	<i>emporter</i>
<i>ajouta</i>	<i>aperçu</i>	<i>dérangeaient</i>	<i>emporter</i>
<i>ajouta</i>	<i>aperçus</i>	<i>dévisser</i>	<i>emporter</i>
<i>ajouta</i>	<i>aperçus</i>	<i>disparu</i>	<i>emporter</i>
<i>ajouta</i>	<i>aperçut</i>	<i>écrasent</i>	<i>emporter</i>
<i>ajouta</i>	<i>aperçut</i>	<i>efforçai</i>	<i>encombre</i>
<i>ajouta</i>	<i>aperçut</i>	<i>égaré</i>	<i>enferme</i>
<i>ajoutais</i>	<i>aperçut</i>	<i>égaré</i>	<i>enferme</i>
<i>ajouté</i>	<i>aperçut</i>	<i>embaumait</i>	<i>enfonça</i>
<i>ajouter</i>	<i>décrire</i>	<i>embaumait</i>	<i>enfuir</i>
<i>ajustait</i>	<i>défait</i>	<i>embellit</i>	<i>enlève</i>

<i>enroula</i>	<i>reconnaître</i>	<i>ressemble</i>	<i>revoir</i>
<i>éprouvai</i>	<i>redevint</i>	<i>ressemblent</i>	<i>revu</i>
<i>éprouve</i>	<i>refis</i>	<i>ressemblent</i>	<i>s'attendrir</i>
<i>éprouve</i>	<i>rejoindre</i>	<i>ressemblent</i>	<i>s'écria</i>
<i>épuisait</i>	<i>réjouir</i>	<i>ressemblent</i>	<i>s'écria</i>
<i>intimide</i>	<i>remua</i>	<i>ressemblent</i>	<i>s'écria</i>
<i>ralluma</i>	<i>répandait</i>	<i>ressemble-</i>	<i>s'écrient</i>
<i>ralluma</i>	<i>repartir</i>	<i>raient</i>	<i>s'étire</i>
<i>rallumer</i>	<i>repose</i>	<i>retournais</i>	<i>se découra-</i>
<i>rallumer</i>	<i>reprit</i>	<i>retrouvé</i>	<i>gea</i>
<i>rassuré</i>	<i>reprit</i>	<i>réveille</i>	<i>se réveiller</i>
<i>rassurent</i>	<i>ressemblaient</i>	<i>réveillé</i>	<i>surprendre</i>
<i>réchauffait</i>	<i>ressemblait</i>	<i>réveiller</i>	<i>surveillance</i>
<i>recomença</i>	<i>ressemblait</i>	<i>réveillons</i>	<i>surveillée</i>
<i>recomença</i>	<i>ressemblait</i>	<i>revoir</i>	
<i>recompte</i>	<i>ressemble</i>	<i>revoir</i>	

Notes

1. It should be noted that the informality of phrasal verbs is a tendency. As Marks (2005: s.p.) points out, “individual phrasal verbs can have distributions that go against the grain of this generalisation. For example, *carry out* is equally common in newspapers and academic writing, but rare in conversation and fiction, and *point out* is more common in academic writing than in the other three genres.”
2. Source language interference *has* been listed as a potential translation universal (cf. Toury’s 1995 “law of interference”), but translation universals *à la Baker* clearly exclude it (see definition above) and most studies on translation universals do not mention interference as one of them (see Mauranen (2004) for a discussion on the unclear status of interference in relation to translation universals).
3. We leave aside another criticism that we could level to such translation universals, namely that some of them come in pairs that are mutually incompatible. In the case of normalization and levelling out, Cappelle (2012) has shown that these two reasonable-sounding candidates for universality in translation behaviour make contradictory predictions in the case of phrasal verbs. For another example, apart from the tendency of *explicitation*, translated language has also been claimed to display *implicitation* (Klaudy and Károly 2005).

Needless to say, both these hypothesized generalizations cannot hold universally, at least not if *universally* is taken to mean what it is supposed to mean, namely ‘for all translators, for all language pairs, in all texts, in every place’. The problem in weakening the universality constraint and allowing these proposed universals to be mere tendencies, as is customary in translation studies, is that these universals lose their predictive potential and effectively become unfalsifiable, thus having little or no scientific value. See Chesterman (1997) for discussion of these concerns.

References

- Baker, Mona
 1993 Corpus linguistics and translation studies: Implications and applications. In *Text and Technology*, M. Baker et al. (eds.), 233–250. Amsterdam / Philadelphia: Benjamins.
- Baker, Mona
 1996 Corpus-based translation studies: The challenges that lie ahead. In *Terminology, LSP and Translation. Studies in language engineering in honour of Juan C. Sager*, Harold Somers (ed.), 175–186. Amsterdam/Philadelphia: Benjamins.
- Baker, Mona and Maeve Olohan
 2000 Reporting *that* in translated English: Evidence for subconscious processes of explicitation? *Across Languages and Cultures* 1(2), 141–158.
- Becher, Viktor
 2010 Abandoning the notion of “translation-inherent” explicitation: Against a dogma of translation studies. *Across Languages and Cultures* 11(1), 1–28.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan
 1999 *Longman Grammar of Spoken and Written English*. Harlow: Pearson.
- Cappelle, Bert
 2001 Is *out of* always a preposition? *Journal of English Linguistics* 29(4), 315–328.
- Cappelle, Bert
 2012 English is less rich in manner-of-motion verbs when translated from French. *Across Languages and Cultures* 13(2), 173–195.
- Cappelle, Bert and Rudy Loock
 2013 Is there interference of usage constraints? A frequency study of existential *there is* and its French equivalent *il y a* in translated vs. non-translated texts, *Target* 25(2): 252–275.

- Corpas Pastor, Gloria, Ruslan Mitkov, Naveed Afzal and Viktor Pekar
 2008 Translation universals: do they exist? A corpus-based NLP study of convergence and simplification. In *Proceedings of the Eighth Conference of the Association for Machine Translation in the Americas (AMTA-08)*, Waikiki, Hawaii, 21–25.
- Chesterman, Andrew
 1997 Explanatory adequacy and falsifiability in translation theory. In *Transferre Necessè Est: Proceedings of the 2nd International Conference on Current Trends in Studies of Translation and Interpreting*, Kinga Klaudy and János Kohn (eds.), 219–224. Budapest: Scholastica.
- Chesterman, Andrew
 2004 Beyond the particular. In *Translation Universals: Do they exist?*, Anna Mauranen and Pekka Kuyamaki (eds.), 33–49. Amsterdam / Philadelphia: Benjamins.
- Claridge, Claudia
 2002 Translating phrasal verbs. In *Teaching and Learning by Doing Corpus Analysis: Proceedings of the Fourth International Conference on Teaching and Language Corpora, Graz, 19-24 July, 2000*, Bernhard Kettemann and Georg Marko (eds.), 361–373. Amsterdam: Rodopi.
- Dai, Guangrong and Richard Xiao
 2011 SL ‘shining-through’ in translational language: A corpus-based study of Chinese translation of English passive, *Translation Quarterly* 62: 85–108.
- Davies, Mark
 2004- *BYU-BNC*. (Based on the British National Corpus from Oxford University Press). <http://corpus.byu.edu/bnc/>.
- Delaere, Isabelle, Gert De Sutter and Koen Plevoets
 2012 Is translated language more standardized than non-translated language? Using profile-based correspondence analysis for measuring linguistic distances between language varieties, *Target. International Journal of Translation Studies* 24(2): 203–224.
- Delaere, Isabelle and Gert De Sutter
 2013 Applying a multidimensional, register-sensitive approach to visualize normalization in translated and non-translated Dutch. In *Interference and Normalization in Genre-Controlled Multilingual Corpora* (Belgian Journal of Linguistics 27), Marie-Aude Lefer and Svetlana Voegelé (eds.), 43–60. Amsterdam: Benjamins.
- Gardner, Dee and Mark Davies
 2007 Pointing out frequent phrasal verbs: A corpus-based approach. *Tesol Quarterly* 42(2), 339–359.
- Halverson, Sandra L.
 2003 The cognitive basis of translation universals. *Target* 15 (2), 197–241.
- House, Juliane
 2008 Beyond intervention: universals in translation, *Trans-kom* 1, 6–19.

- Iacobini, Claudio and Francesca Masini
 2003 The emergence of verb-particle constructions in Italian: locative and actional meanings. *Morphology* 16(2), 155–188.
- Jiménez-Crespo, Miguel A.
 2011 The future of “universal” tendencies: A review of papers using localized websites, *Proceedings of UCCTS 2010, Using Corpora in Contrastive and Translation Studies*, Edge Hill University, UK. Online available at: <http://www.lancs.ac.uk/fass/projects/corpus/UCCTS2010Proceedings/>.
- Klaudy, Kinga and Krisztina Károly
 2005 Implication in translation: An empirical justification of operational asymmetry in translation. *Across Languages and Cultures* 6(1), 13–28.
- Koppel, Moshe and Noam Ordan
 2011 Translationese and its dialects. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, 1318–1326. Portland, Oregon: Association for Computational Linguistics.
- Kruger, Haidee and Bertus Van Rooy
 2012 Register and the features of translated language. *Across Languages and Cultures*, 13(1): 33–65.
- Kopecka, Anetta
 2006 The semantic structure of motion verbs in French: Typological perspectives. In *Space in Languages: Linguistics Systems and Cognitive Categories*, Maya Hickmann and Stéphane Robert (eds.), 83–101. Amsterdam / Philadelphia: Benjamins.
- Laviosa, Sara
 1998 Core patterns of lexical use in a comparable corpus of English narrative prose. *Meta* 43(4), 557–570.
- Laufer, Batia and Stig Eliasson
 1993 What causes avoidance in L2 learning: L1-L2 difference, L1-L2 similarity, or L2 complexity? *Studies in Second Language Acquisition* 15(1), 35–48.
- Lefer, Marie-Aude and Svetlana Vogeleeer (eds.)
 2013 *Interference and Normalization in Genre-Controlled Multilingual Corpora* (Belgian Journal of Linguistics 27). Amsterdam: Benjamins.
- Liao, Yan and Yoshinori J Fukuya
 2004 Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning* 54(2), 193–226.
- Lind, Sarah
 2007 Translation Universals (or laws, tendencies, or...?) *TIC Talk 63* (Newsletter of the United Bible Societies Translation Information

- Clearinghouse). Online available at: http://www.ubs-translations.org/tt/past_issues/tic_talk_63_2007/.
- Loock, Rudy, Gert De Sutter and Koen Plevoets
 2013 Teasing apart Translation Universals and Source-Language Interference: A case study on derived adverbs in English and French. Talk presented at ICLC 7 – UCCTS 3, Ghent, 11-13 July 2013.
- Marks, Jonathan
 2005 The truth revealed: phrasal verbs in writing and speech. *MED Magazine* 34. Online available at: <http://www.macmillandictionaries.com/MED-Magazine/October2005/34-Feature-PV-Spoken-Written.htm>.
- Mauranen, Anna
 2004 Corpora, universals and interference. In *Translation Universals: Do they Exist?*, Anna Mauranen and Pekka Kujamäki (eds.), 65–82. Amsterdam / Philadelphia: Benjamins.
- Mauranen, Anna and Pekka Kujamäki
 2004 *Translation Universals: Do They Exist?* Amsterdam / Philadelphia: Benjamins.
- McEnery, Tony, Richard Xiao and YuKio Tono
 2006 *Corpus-based Language Studies: An Advanced Resource Book*. London and New York: Routledge.
- Mudraya, Olga, Scott S. L. Piao, Laura Löfberg, Paul Rayson and Dawn Archer
 2005 English-Russian-Finnish cross-language comparison of phrasal verb translation equivalents. In *Proceedings of Phraseology 2005: An Interdisciplinary Conference, 13-15 October 2005, Louvain-la-Neuve, Belgium*, C. Cosme, C. Gouverneur, F. Meunier and M. Paquot (eds.), 277–281.
- Olohan, Maevé
 2003 How frequent are the contractions? *Target* 15(1), 59–89.
- Øverås, Linn
 1998 In search of the third code: An investigation of norms in literary translation. *Meta* 43(4), 557–570.
- Paillard Michel and Nicole Videau
 2008 Les verbes français préfixés en *dé-* et leurs traductions en anglais. In *Préfixation, prépositions, postpositions*, M. Paillard (ed.), 75–91. Rennes: Presses Universitaires de Rennes.
- Pourcel, Stéphanie and Anetta Kopecka
 2006 Motion events in French: Typological intricacies. Unpublished ms., Brighton: University of Sussex and Nijmegen: Max Planck Institute for Psycholinguistics.
- Talmy, Leonard
 2000 *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press.
- Teich, Elke
 2003 *Cross-Linguistic Variation in System and Text*. Berlin: Mouton de Gruyter.

Thim, Stefan

2012 *Phrasal Verbs: The English Verb-Particle Construction and its History*. Berlin: Mouton de Gruyter.

Tirkkonen-Condit, Sonja

2004 Unique items: over- or under-represented in translated language? In *Translation Universals: Do They Exist?*, Anna Mauranen and Pekka Kujamäki (eds.), 177–184. Amsterdam / Philadelphia: Benjamins.

Toury, Gideon

1995 *Descriptive Translation Studies and Beyond*. Amsterdam / Philadelphia: Benjamins.

Wild, Catherine

2010 *Attitudes towards English usage in the late modern period: The case of phrasal verbs*. Doctoral thesis. University of Glasgow.