

WORD-FORMATION IN TRANSLATED LANGUAGE: THE IMPACT OF LANGUAGE-PAIR SPECIFIC FEATURES AND GENRE VARIATION*

MARIE-AUDE LEFER

Université catholique de Louvain and Institut libre Marie Haps
Centre for English Corpus Linguistics, Collège Erasme
Place Blaise Pascal, 1348 Louvain-la-Neuve, Belgium
Phone: 0032 10 47 40 34
E-mail: marie-aude.lefer@uclouvain.be

Abstract: Corpus-based studies in the field of word-formation have looked at translated language with the aim of (i) assessing the role of the source language (SL) in the over-use of certain derivational affixes in the target language (TL) and (ii) examining translation-related, SL- and TL-independent trends such as the normalization of creative lexis. However, two aspects of the word-formation features of translated language remain under-researched: (i) language-pair specific properties which lead to a marked decrease in the use of certain word-forming devices in the target texts compared to their source texts (here referred to as ‘language-pair specific morphological decrease’) and (ii) the genre-sensitivity of word-forming devices in translated language. To examine these two aspects, we report on a corpus-based case study of the French translation of the English *un-* and *in-* prefixes. The translation data makes it possible to identify the following language-pair-related factors which are responsible for morphological decrease in French target texts: morphological productivity, diverging polysemy and partial phraseological equivalence. The study also shows that morphological decrease is significantly more frequent in fiction than in news, thereby indicating that genre also plays a major role in shaping the word-formation features of translated language.

Keywords: corpora, translation, word-formation, morphological decrease, language pair, genre

1. INTRODUCTION: BACKGROUND AND OBJECTIVES

This article sets out to highlight some of the factors that shape the word-formation features of translated texts. The present study is situated at the cross-roads between corpus-based contrastive linguistics and translation studies, morphology and lexicology. This section introduces the general background of the study and sketches its major aim, namely to assess the potential impact of two types of variables on the reduced use of word-formation devices in translated

* This paper was first presented at the *Methodological Advances in Corpus-based Translation Studies* symposium held at Ghent University College in January 2010. I am greatly indebted to the two anonymous reviewers and Gert De Sutter for their insightful comments on the first draft of this article. All remaining shortcomings are of course my own.

language (here French translated from English): language-pair-related factors (Section 1.1) and variation across genres (Section 1.2). Why I chose to focus on ‘morphological decrease’ rather than on ‘morphological increase’ will be clear in Section 2, where the two notions are defined.

1.1. Translation Studies and Language-pair Specific Features

It is now widely recognized that corpora have dramatically revolutionized the fields of contrastive linguistics and translation studies (Granger et al. 2003; Xiao 2010). The two disciplines make use of similar data to pursue their respective research objectives, which can facilitate their rapprochement (Granger 2003). One of the most obvious signs of this possible rapprochement is the renewed interest that translation studies have taken in the role played by source language systems in corpora of translated texts.

Back in the early 1990s, Baker (1993, 1995) convincingly called for the exploration of universal features of translation, i.e. “features which typically occur in translated text rather than original utterances and which *are not the result of interference from specific linguistic systems*” (1993:243; my emphasis MAL). Baker (1993:240) also stressed that descriptive translation studies as a discipline cannot be reduced to “comparative analyses of source and target texts”, thereby advocating the use of corpora of translated texts (cf. the *English Translational Corpus*¹): “we need to effect a shift in the focus of theoretical research in the discipline, a shift away from comparing either ST [source texts] with TT [target texts] or language A with language B to comparing text production per se with translation” (Baker 1995:233). After some twenty years of intensive and wide-ranging corpus-based research in the field, it now appears that some universals are less plausible than others (see e.g. Becher 2010 on ‘translation-inherent’ explicitation) and that a number of variables other than the translation process itself seem to provide viable alternative explanations for the observed differences between non-translated original texts and translated texts. For example, some scholars have recently re-established the source language as a crucial factor in translation studies (e.g. Becher 2010; Bernardini 2010; De Sutter et al. 2012). This trend ties in with Granger’s (2003:26) observation that “lack of familiarity with TS [translation studies] findings may lead CL [contrastive linguistics] researchers to interpret their data in terms of differences between language systems when they result from translation norms or strategies, while TS researchers may similarly misinterpret their data because of a lack of awareness of a systematic difference between the two language systems established by CL”.

One of the two aims of this article is to assess the impact of language-pair specific (i.e. SL_x - and TL_y -related) properties on the word-formation features of

French texts translated from English. This will be done by taking stock of well-known morphological contrasts between the two language systems and focusing on the decrease, from source texts to target texts, in the use of negative prefixes (English negative *un-* and *in-*) (see Section 2.2 on the notion of morphological decrease).

English–French contrastive grammars and translation textbooks (such as Vinay and Darbelnet 1958/1995; Astington 1983; Chuquet and Paillard 1987; Van Hoof 1989; Van Roey 1990; Paillard 2000) have repeatedly pointed out that certain English derivational affixes are highly productive compared to their French counterparts, hence leading translators to use a variety of non-morphological translation devices in French, such as paraphrases, when translating words containing these affixes. Note that the concept of ‘morphological productivity’ is multi-faceted and complex (see e.g. Baayen 2009 for an overview). It is used here in its most common sense, i.e. the quantitative probability of giving rise to new words. According to these reference textbooks, the paraphrasing pattern is especially frequent with affixes of Germanic origin (e.g. *un-*, *over-*, *-ness*, *-ly*). Examples are provided in (1) to (4) (see Chuquet and Paillard 1987; Van Roey 1990):

- (1) En. recent**ness**
Fr. nature récente ‘recent nature’
- (2) En. admitted**ly**
Fr. il faut le reconnaître ‘one must admit it’
- (3) En. bachelor**hood**
Fr. condition de vieux garçon ‘condition of bachelor’
- (4) En. **un**inspired
Fr. peu inspiré ‘not very much inspired’

The contrastive grammars and translation textbooks mentioned above are mostly introspection-based (except Chuquet and Paillard 1987). However, a recent exploratory study of negative affixes in English, French and Italian (Cartoni and Lefer 2011) has also noted this pattern in corpus data extracted from the *Europarl* corpus (Koehn 2005). It has empirically confirmed that the Germanic prefix *un-* frequently calls for the use of non-morphological translations such as non-prefixed words and paraphrases in French (c. 65% of the cases), mainly because of the lack of a prefixed equivalent in French (see examples 5 and 6; **incible* ‘unfocused’ and **inéthique* ‘unethical’ are not (yet) attested, at least not in reference tools such as monolingual dictionaries). Conversely, the study reveals that derivatives containing the other negative prefix *in-*, which is far less productive than its Germanic counterpart *un-* and than its French cognate *in-* (Lefer 2009), corresponds more systematically to words containing a negative prefix in French, such as *in-* (c. 60% of the cases; see examples 7 and 8):

- (5) [Europarl] En. But I have to say both as a citizen and as a parliamentarian that the discussions so far have seemed to me turgid, **unfocused** and to show a lack of political will and motivation.
Fr. Je dois cependant dire, en ma qualité de citoyen et de parlementaire, que, jusqu'ici, les débats m'ont semblé verbeux, **vagues** et dépourvus de volonté et de motivation politiques.
- (6) [Europarl] En. It is an abhorrent and **unethical** practice which is totally irresponsible from a scientific point of view.
Fr. Cette pratique, répugnante et **contraire à l'éthique**, est totalement irresponsable d'un point de vue scientifique.
- (7) [Europarl] En. The heroine of this book is a raven-haired, stunningly good-looking Member of the European Parliament of **indeterminate** nationality who single-handedly takes on the evil chemical industry.
Fr. L'héroïne de cette bande dessinée est une députée européenne d'une beauté incroyable aux cheveux de jais et de nationalité **indéterminée** qui s'attaque seule à l'industrie chimique malfaisante.
- (8) [Europarl] En. I have heard speeches from Mr Cohn-Bendit and Mr Le Pen which I consider to be **indistinguishable**.
Fr. J'ai entendu les interventions de M. Cohn-Bendit et de M. Le Pen et, d'après moi, elles sont **indifférenciables**.

In sum, Cartoni and Lefer (2011) seem to empirically confirm a long-standing intuition in English–French contrastive linguistics, namely that cross-linguistic differences in morphological productivity between English and French play a crucial role in shaping the word-formation features of translated French: the prefix *un-*, which is very productive in English, seems to lead to a relatively high proportion of non-morphological translations, while the less productive *in-* corresponds more frequently to a negative prefix in French target texts. Nevertheless, the study also suggests that other variables must be at play: there is a non-negligible number of cases (which are not systematically looked at in the study) where the translator resorts to a non-morphological translation in French despite the fact that the prefixed equivalent is attested and frequent. This is illustrated in (9) and (10) (the adjectives *indifférenciable* 'indistinguishable' and *irréel* 'unreal' are both attested in French).

- (9) [Europarl] En. (...) in Ireland non-consultant hospital doctors worked well in excess of the 65 hours permitted per week and they undertake tasks **indistinguishable** from that of their senior colleagues.
Fr. (...) les médecins hospitaliers non consultants irlandais travaillaient nettement plus que les 65 heures par semaine autorisées et (...) ils remplissaient les **mêmes** tâches que leurs collègues plus expérimentés.

- (10) [Europarl] En. Since the reform of agriculture a number of years ago, to a large extent the price review process we go through has become an **unreal** exercise and does not recognize the problems that face farmers on the ground.

Fr. Depuis la réforme agricole qui a eu lieu il y a plusieurs années, le processus de révision des prix auquel nous nous soumettons s'est mué, dans une large mesure, en un exercice **abstrait** qui ne tient pas compte des problèmes rencontrés par les agriculteurs sur le terrain.

It logically follows that other, as yet largely unexplored, factors influence the use of word-forming devices in translated texts. The present study aims to find an appropriate approach to identify (some of) these factors, especially those that are unquestionably SL- or TL-related, and assess their relative importance by systematically examining the translation strategies adopted by news and literary translators when dealing with *un-* and *in-*words.

1.2. Translation, Word-formation and Genre

As pointed out by Hatim (2001:141), genre “provides translators with a framework within which appropriateness is judged and the various syntactic, semantic, pragmatic and semiotic structures are handled”. In this section, I will first have a look at the distinction between literary and non-literary (e.g. news) translation before examining the importance of this (admittedly broad) variable in contrastive and translation studies, and more particularly those that deal with word-formation.

In addition to fulfilling different functions, literary and non-literary translations differ in a number of respects. In literary translation, “the how one says something can be as important, sometimes more important, than what one says” (Landers 2001:7). Literary and non-literary translation can be differentiated on the basis of the following criteria (which are no more than general tendencies) (Boase-Beier 2006:26–30):

- Literary translation, contrary to non-literary translation, has a tendency to preserve the **style**, not just the message;
- Non-literary translation can covertly be a translation (e.g. translated items of news), while literary translation tends to **overtly** be a translation (cf. “translated by ...”);
- Literary translation undergoes more **editing** than non-literary translation (e.g. publishing houses). It is also probably under less time pressure than non-literary translation such as news translation;

- Literary translation tends to have a **less ephemeral** existence than non-literary translation. It also usually has a wider distribution (although this very much depends on the size of the intended readership).

It can be postulated that translated literary and non-literary texts probably display a number of linguistic differences that mirror the above-mentioned general characteristics.

In contrastive and translation studies based on translation corpora (source texts together with their target texts), the emphasis has mainly been on literary genres (e.g. novels). Few studies have examined cross-genre differences (see e.g. Altenberg 2004/2005; Cosme 2006; Neumann 2008; De Sutter et al. 2012). As pointed out by Olohan², however,

[w]e need to recognize that a substantial part of translation activity in the world involves non-literary texts. If we are to study the activity of translation and translation process generally, we need to delve deeper into those genres and texts too. This will have the consequence that we will be able to make more cross-genre comparisons and study the extent to which features of translation may be influenced by genre, text type, etc. (Olohan 2004:191)

In fact, such cross-genre approaches are essential in word-formation studies. In a seminal article, Plag et al. (1999) have examined the relation between register variation and derivational productivity in the *British National Corpus* (BNC) by measuring and comparing the productivity of 15 suffixes in writing and speech. Their results showed that the productivity and use of English suffixes greatly vary across registers and that productivity is generally higher in writing than in speech. These findings led to a boom in corpus-based cross-register and/or cross-genre studies in the field (see e.g. Grabar et al. 2006 for French). In the same vein, in a contrastive study of prefixation in English and French writing (Lefer 2009), I have shown that the productivity and frequency of English and French prefixes vary across written genres (scientific articles, news editorials, novels) and across academic disciplines (medicine, economics, linguistics). Prefixation is significantly more productive in scientific articles and news editorials than in novels.

In view of its overarching importance, the genre variable needs to be taken into account in word-formation studies. The second aim of the present article is therefore to assess as precisely as possible the impact of genre on the use of word-formation devices in translated texts. More precisely, I will compare the strategies used to translate English *un-* and *in-*words in fiction and news.

The article is structured as follows. Section 2 reviews a number of key contrastive and translation studies that have looked at the word-formation features

of translated language and introduces the terms ‘morphological decrease’ and ‘morphological increase’. The corpus, methodology, translation data typology and statistical test used are then described in Section 3, while Sections 4 and 5 are devoted to the discussion of the results. The article ends with some concluding remarks and avenues for future research (Section 6).

2. THE WORD-FORMATION FEATURES OF TRANSLATED LANGUAGE: INSIGHTS FROM CONTRASTIVE LINGUISTICS AND TRANSLATION STUDIES

2.1. An Overview of Language-pair Specific and Translation-Inherent Phenomena

Two main trends can be identified in corpus-based research on word-formation in translation. Recent corpus-based studies have found traces of **source text interference** in the overuse of some derivational affixes in translated language. Using a bidirectional English–Chinese translation corpus, Wang and Qin (2010:175) have shown that the Chinese derivational noun-forming suffix *-xing* ‘property’ is overused in translated Chinese fiction compared to original Chinese fiction. The authors argue that this tendency is due to the high frequency of *-ity*, *-ness* and *-dom* in English source texts compared to *-xing* in non-translated Chinese. This affix by affix translation pattern seems to be at play in other language pairs as well. For example, Frankenberg-Garcia (2008) has noted that Portuguese *-mente* adverbs are significantly more frequent in translated texts than in non-translated texts when the source language is English. This trend, provided it is generalizable to other language pairs, could be interpreted as a form of **SL-induced morphological increase** in translated language (see Section 2.2).

In addition to morphological increase, corpus-based research has also put forward a translation universal referred to as the normalization of creative lexis. Note that this strand of research does not necessarily focus on word-formation patterns. However, two studies have examined the normalization of creatively-formed derivatives and compounds, viz. Kenny (2001) and Olohan (2004). Kenny (2001) has investigated the phenomenon of normalization in the English translation of German creative lexis, namely derivatives, conversions, compounds and collocations. In a translation corpus, “normalization may be said to occur when translators opt for conventional target language solutions to problems posed by creative or unusual source text features” (ibid. 66). Kenny’s results, which are based on the *German–English Parallel Corpus of Literary Texts*, indicate that SL creative lexis is often translated by means of more conventional lexis in the target language. For example, the German creative blend

Aber-gläubigkeit (*Aberglaube* ‘superstition’, *Gläubigkeit* ‘religious devoutness, naïve trust’, *aber* ‘but’) is translated into English *superstitions* (ibid. 146), which, contrary to the source creative form, is morphologically simple. Normalization may in fact result in morphological decrease in the target text when the SL creative word-formation pattern is not translated into a similar TL pattern (e.g. Ge. *Hans schluckt adamsäpfelnd.* – En. *Hans swallows. His Adam’s apple bobs.*) Along similar lines, Olohan (2004:108–117) has investigated lexical productivity in translation through the analysis of productively formed *-ish* derivatives in non-translated and translated English. Relying on the fiction parts of the *Translational English Corpus* and the BNC, Olohan has shown that lexis is more conservative and more conventional in translated fiction, thus corroborating Kenny’s findings mentioned above. She found that there are more non-lexicalized, creative uses of *-ish* (i.e. *-ish* forms not listed in reference dictionaries) in non-translated fiction than in translated fiction (e.g. *mummyish*, *Londonish*). Olohan’s study, like Kenny’s, seems to point to a form of **translation-inherent morphological decrease** as creatively formed derivatives in *-ish* are less frequently encountered in translated English.

2.2. ‘Morphological Increase’ vs. ‘Morphological Decrease’ in Translation

The corpus-based studies summarized in Section 2.1 above seem to point to the existence of two opposing trends, both of which display a significant impact on the word-formation features of translated texts, vis-à-vis non-translated original texts and SL texts:

– ‘**SL-induced morphological increase**’, i.e. an increase in the number of morphologically complex words (i.e. derivatives, compounds, conversions, blends, etc.) in translated texts in L_y compared to non-translated texts in L_y , which is due to SL_x interference. In Wang and Qin (2010) for example, it is assumed that this increase stems from SL interference: morphologically complex words in the SL_x tend to be translated into equivalent morphologically complex words in the TL_y (affix-by-affix translation pattern). In other words, morphological increase, when it is SL-induced, results in a morphological *status quo* between the source text and its target text. It is an increase which concerns non-translated L_y vs. translated L_y . It is possible to hypothesize that the amount of SL-induced morphological increase varies across language pairs: a morphologically rich SL_x will have a greater morphological impact on the TL_y than a comparatively morphologically poor SL_z . It remains to be seen whether translation-inherent trends also account for a specific form (or several forms) of morphological increase (which

would then be termed ‘translation-inherent morphological increase’). To the best of my knowledge, this aspect has not yet been investigated;

– ‘**translation-inherent morphological decrease**’, i.e. a decrease in the number of morphologically complex words (i.e. derivatives, compounds, conversions, blends, etc.) in target texts in L_y compared to their source texts in L_x , which is inherent in the translation process itself. This decrease seems to be triggered by creative forms and is a by-product of normalization: SL_x creative lexis (which can sometimes (though by no means always) be morphologically constructed) is translated into more conventional TL_y lexis (which can sometimes (though by no means always) be morphologically not constructed, i.e. simplex). According to Kenny (2001) and Olohan (2004), normalization is a universal feature of translated texts: it occurs independently of the SL and TL. This form of morphological decrease should be further tested by using what Baker (1995) calls a ‘comparable corpus’, i.e. translated and non-translated texts in $L_{y(\text{and } x, z, \dots)}$, preferably representative of different registers and genres, and by counting and comparing the total number of morphologically complex words in each sub-corpus (prefixed and suffixed words, compounds, blends, etc.).

Following the exploratory findings presented in Cartoni and Lefer (2011), I would like to postulate that language-pair-related factors also play a fundamental role in the phenomenon of morphological decrease, a variable that has been overlooked in Kenny (2001) and Olohan (2004). No explicit mention is made of the possible influence of the German morphological system on the translation patterns observed in the English target texts in Kenny (2001). The exact role played by the source languages on the use of *-ish* in translated English in Olohan (2004) is not taken into account either. It might be the case, for example, that German is much more flexible than English in creating new words (see e.g. Vinay and Darbelnet 1958/1995). It might also be that a number of source languages in the *Translational English Corpus* lack attenuative suffixes such as the English *-ish*. I therefore wish to introduce a second type of morphological decrease: ‘**language-pair specific morphological decrease**’, i.e. a decrease in the number of morphologically complex words (i.e. derivatives, compounds, conversions, blends, etc.) in target texts in L_y compared to their source texts in L_x , which is due to morphological contrasts between the two language systems under scrutiny (L_x and L_y) (see Lefer and Cartoni 2011 and Lefer 2011 on contrastive word-formation).

As mentioned previously, the present study focuses on one aspect of the word-formation features of translated texts, viz. language-pair specific morphological decrease. I will not examine the other aspects of morphological increase/decrease. Importantly, I do not claim that there is an overall morphologi-

cal decrease (or increase) in translated texts compared to non-translated texts or in target texts compared to source texts. My aim here is not to draw general conclusions on the overall morphological features of translated language but rather to study in some more depth one facet of this intriguing and under-researched aspect of translated language.

3. DATA AND METHODOLOGY

3.1. Corpus Used

The study relies on part of the English–French PLECI translation corpus, which contains two genres: fiction and news.³ The corpus is sentence-aligned. The PLECI sub-corpus used in this study is made up of c. 150,000 words per language and per genre (see *Table 1*) and totals 575,292 words.

Table 1
Translation corpus used in the study

	Number of running words		Number of texts	Average number of running words per text	Publication date
Fiction	EN (source)	137,885	10	13,788	1985–2003
	FR (target)	143,348		14,335	
News	EN (source)	141,919	unknown	unknown	1999–2005
	FR (target)	152,140			

The literary sub-corpus is made up of ten extracts taken from the beginning of novels published between 1985 and 2003. It represents 9 authors (two extracts are by the same author, viz. Ruth Rendell) and 11 translators (one of the novels, *An Unkindness of Ravens*, was jointly translated by two translators). These extracts were selected because they are the most recent ones in PLECI (some novels included in the corpus were published much earlier, e.g. in the 1940s). The news sub-corpus covers a period of c. 6 years and is made up of (full) newspaper articles published in various English-language newspapers together with their French translations published in *Courrier International*, *Le Monde Diplomatique*, *Le Monde* and *Time*.

3.2. Data Extraction Method and Data Selection Criteria

The English sentences with *un-* and/or *in-*derivatives and the corresponding French segments were automatically extracted with the multilingual concordancer *ParaConc* (Barlow 2008) on the basis of form-based queries including a wild card, namely *un** and *in** (and its allomorphs *il**, *im** and *ir**). It is an established fact that form-based queries generate a certain amount of noise, i.e. irrelevant data. However, its main advantage is its 100% recall rate. As a result, the following data selection filters were applied to weed out unwanted data:

- **noise**, i.e. irrelevant data that do not contain the prefixes *un-* or *in-*, such as in *unique*, *understand*, *inspector*, *invite*, *intergovernmental*;
- ***un-*verbs**, where *un-* conveys reversal or removal (e.g. *She'd **undone** her coat but hadn't taken it off*, see Cartoni and Lefer 2011). The reason for getting rid of the *un-*verbs is that the English *in-* cannot be prefixed to verbs to convey reversal/removal. In view of the fact that the study aims to analyse *in-* and *un-* data sets that are as comparable as possible in terms of potential translation strategies, it was necessary to discard *un-*verbs, which, if included, might have skewed the results of the comparison. This weed-ing-out step required a systematic disambiguation of *unXed* and *unXing* forms (where *X* is a verbal base), as they can act both as verbs and/or adjectives in English and can therefore express either reversal/removal (in verbs) or contradictory/contrary negation (in adjectives) (e.g. *She **unlocked** her car* vs. *an **unlocked** car* 'not locked, open'). The *un-*forms have been included in the final data set and considered as *unXed* or *unXing* adjectives if (i) they are used in attributive position (e.g. *There's still an oval of **un-melted** frost on the roof*) and/or (ii) the *unX* verb is not attested, i.e. not recorded in reference monolingual dictionaries such as the *Macmillan English Dictionary*⁴ (e.g. *uninterested*: **uninterest*) and/or (iii) they are listed as adjectives in monolingual dictionaries;
- **synchronically opaque derivatives**: three defining criteria for derivative status were used to weed out synchronically opaque, highly lexicalized derivatives (see Lefer 2011:653): (i) formal analysability (derivatives are perceived as analysable as a base plus *un-* or *in-*), (ii) semantic transparency (it is possible to deduce the meaning of the derivative from the meaning of its base and the negative meaning of the prefix it contains) and (iii) recurrence (existence of a derivational paradigm where the base of the derivative occurs in several other words with a similar meaning, which is easily apparent to native speakers with no knowledge of the classical languages or French). Following these criteria, opaque derivatives such as *inadvertent* and *insomniac* were discarded from the data set.

The pre-final data set is presented in *Table 2* (one more filter will be used once the translation data are examined, see Section 3.3). Note that some *un-* or *in-* derivatives occur several times (e.g. *unemployment* in the news sub-corpus, *impatiently* in the fiction sub-corpus). The results presented in this study are based on the analysis of tokens, as the same English type can be translated into a variety of equivalents, which are all of equal interest.

Table 2
Pre-final data set

English prefix	Number of tokens	Number of types
<i>in-</i>	299	125
<i>un-</i>	433	221
Total	732	346

3.3. Translation Data Typology

The French translation equivalents of the English *un-* and *in-* derivatives extracted from PLECI were classified on the basis of the following typology, which is a simplified version of the classification developed in Cartoni and Lefer (2011):

- **Morphological translation**
 - **Translation into a prefixed word** containing a negative prefix (e.g. *in-*, *non-*, *é-*, *mal-*). In example (11), En. *unpredictable* is rendered into Fr. *imprévisible*:
- (11) [PLECI news] En. Because road systems are complex, and because the behaviour of individual drivers is largely **unpredictable**, the overall impact of a particular pricing scheme is anyone's guess.
Fr. A cause de la complexité du réseau routier et parce que le comportement de chaque automobiliste est très **imprévisible**, on ne peut que faire des hypothèses quant à la conséquence globale d'un schéma particulier de péage.
- **Non-morphological translation** (the non-morphological translation strategies described below lead to morphological decrease in the target texts)

- **Translation into a non-prefixed word**, i.e. a word devoid of negative prefixation. In (12), *unpaved* is translated into the French word *cahoteux* ‘bumpy’:
- (12) [PLECI fiction] En. (...) and also took a back road into the city centre – a road that was **unpaved** and also largely deserted.
Fr. (...) il a entrepris d’enchaîner les routes les plus secondaires pour rejoindre le centre-ville, m’entraînant sur des chemins non seulement **cahoteux** mais aussi déserts.
- **Translation into a paraphrase**, i.e. the prefixed derivative is rendered into more than one word in the target text.⁵ Example (13) illustrates this pattern: *unconnected* is rendered into a relative clause *qui n’avaient aucun lien* ‘that had no link’, where negation is conveyed syntactically rather than morphologically:
- (13) [PLECI news] En. Yet Bush has all but declared war on three nations **unconnected** with the events of 11 September – Iraq, Iran, and North Korea.
Fr. M. Bush a réagi en déclarant virtuellement la guerre à trois pays – Irak, Iran et Corée du Nord (l’“Axe du Mal”) – **qui n’avaient aucun lien** avec les événements du 11 septembre.
- **Zero correspondence** (Johansson 2007), when the equivalent of the SL prefixed word is omitted in the target text.
 - **No translation**, when the SL sentence/paragraph containing the prefixed word is left out in the target text.

Cases of ‘zero correspondence’ and ‘no translation’ (11% of the pre-final data set) were discarded from the final data set, as they cannot be used to investigate the word-formation features of translated texts (there is no equivalent in the target text). Cases of zero correspondence (see e.g. 14) should be studied by means of other methods (e.g. think-aloud protocols; cf. Bernardini 2002) to see to what extent they are triggered by untranslatability. Both types occur equally frequently with *un-* and *in-* and in news and fiction (no cross-genre differences were observed).

- (14) [PLECI fiction, zero correspondence] En. She had on a white tee-shirt, pure **unrelieved** white, short-sleeved and showing felt-tipped pen drawings on her arms and hands, the snake again in green, a butterfly with a baby face, a raven woman with aggressive breasts and erect wings (...).
Fr. Elle portait un tee-shirt blanc à manches courtes, et avait dessiné au stylo-feutre sur ses bras et ses mains un serpent vert, un papillon

avec une tête de bébé et une femme corbeau aux seins agressifs et aux ailes déployées.

The final data set, which consists of 653 prefixed tokens⁶, is found in *Table 3*. It is made up of 41% of *in*-tokens and 59% of *un*-tokens.

Table 3
Final data set

English prefix	Number of tokens	
<i>in</i> -	News	164 (25%)
	Fiction	101 (16%)
	Total	265 (41%)
<i>un</i> -	News	209 (32%)
	Fiction	179 (27%)
	Total	388 (59%)
Total		653 (100%)

3.4. Test for Statistical Significance

The Pearson's chi-square test (henceforth χ^2) was used to assess the statistical significance of the observed trends (significance threshold used: $p \leq 0.05$). Although admittedly a rather rudimentary (and hence imperfect) statistical measure (see Kilgariff 2001), it is commonly used to test statistical significance in corpus linguistics (McEnery et al. 2006:55).

4. LANGUAGE-PAIR SPECIFIC FEATURES ACCOUNTING FOR MORPHOLOGICAL DECREASE IN TRANSLATED TEXTS

The three research questions tackled in this section are the following: (i) what language-pair-related factors lead to morphological decrease in translated texts?; (ii) how can these factors be identified?; (iii) what is the exact impact of these factors in terms of the amount of morphological decrease they generate? The genre variable will be dealt with in Section 5.

4.1. Morphological Productivity

The overall results of the analysis indicate that *un*-derivatives lead to the use of non-morphological translations (i.e. non-prefixed words and paraphrases) much more systematically than *in*-derivatives (56% vs 24%) ($\chi^2 = 68.49$, $df = 1$, $p < 0.001$; see *Table 4*).

Table 4

Overall results: morphological vs. non-morphological translations

Prefix	Morphological translations	Non-morphological translations		Total
	Prefixed words	Non-prefixed words	Paraphrases	
<i>in</i> -tokens	202 (76%)	36 (14%)	27 (10%)	265 (100%)
<i>un</i> -tokens	169 (44%)	97 (25%)	122 (31%)	388 (100%)

The most obvious explanation for this diverging pattern is the variable of **productivity**: the English *un*- is more productive than its French counterpart *in*- and therefore triggers the use of non-morphological translations in about half the cases. The English *in*-, by contrast, is less productive than its French cognate *in*- and is easily translated into a negative prefix (76% of the cases). To assess the impact of the productivity variable more precisely, I analysed whether the prefixed counterparts of English *un*- and *in*-words were part of the present-day French word stock by using a bilingual dictionary (Le Grand Robert and Collins version 2.0 online (GR&C), where full text searches are possible⁷) and a French monolingual dictionary (Le Petit Robert). Using a bilingual dictionary for this purpose is very informative as English–French bilingual dictionaries have a tendency to systematically propose prefixed equivalent translations in prefixed word entries (Lefer 2010). I found that 95% of the *in*-derivatives (253 out of 265) and 70% of the *un*-derivatives (271 out of 388) found in the corpus have at least one prefixed equivalent in French. This difference between the two prefixes is significant ($\chi^2 = 65.23$, $df = 1$, $p < 0.001$), confirming that there is an intrinsic difference between these prefixes from an English–French contrastive viewpoint. However, as shown in *Table 5*, there is no systematic link between the use of non-morphological translation devices and the lack of a French prefixed equivalent: 56 *in*-tokens and 113 *un*-tokens with an attested prefixed equivalent in French are translated into a non-prefixed word or a paraphrase in the corpus. This is illustrated in (15) and (16) (*inaudible* and *indéfectible/inébranlable* are attested in French). It can therefore be concluded that productivity does not provide a viable explanation for all instances where *un*- or *in*-tokens do not correspond to a prefixed word in the French target texts.

Table 5

(Lack of) prefixed equivalents and use of morphological and non-morphological translations

Prefix	Prefixed equivalent?	Prefixed words ⁸	Non-prefixed words	Paraphrases	Total non-morphological translations	Total
<i>in</i> -tokens	no	5 (42%)	4 (33%)	3 (25%)	7 (58%)	12 (100%)
	yes	197 (78%)	32 (13%)	24 (9%)	56 (22%)	253 (100%)
<i>un</i> -tokens	no	11 (9%)	50 (43%)	56 (48%)	106 (91%)	117 (100%)
	yes	158 (58%)	47 (17%)	66 (25%)	113 (42%)	271 (100%)

- (15) [PLECI fiction] En. Then the three in the car fall silent with the huge arrival of constant streetlights, the massed flow of the first dual carriageway, and like an **inaudible** hum all round them (...).

Fr. A ce moment-là tous les trois dans la voiture ils se taisent avec l'écrasante arrivée de feux rouges ininterrompus, le flot compact de la première route à deux voies, et comme un bourdonnement **sourd** tout autour d'eux (...).

- (16) (16) [PLECI news] En. Stories about abuses of women may win plaudits, but combating such abuses takes time, resources, imagination, political will – and an **unswerving** commitment to women's rights.

Fr. Les récits de ces violences peuvent mériter des félicitations, mais les combattre requiert du temps, des moyens, de l'imagination et une volonté politique ainsi qu'un engagement **constant**.

In view of this essential observation, I believe it is necessary to zoom in more tightly on other possible language-pair-related factors which may be at play in morphological decrease. I will therefore concentrate on the instances where *in*- and *un*-tokens correspond to (i) a non-prefixed word or (ii) a paraphrase, despite the fact that there is an attested French prefixed equivalent. I will leave aside (i) tokens that have no prefixed counterparts in French (where morphological decrease results from cross-linguistic productivity differences) and (ii) tokens that are translated into a French prefixed word (i.e. cases where the word-formation feature of the source text is retained in the target text). Incidentally, it is noteworthy that some of these latter cases may be the result of SL-induced morphological increase (but it falls outside the scope of this article to look more closely into this). This left us with 169 tokens (113 *un*-tokens and 56 *in*-tokens) for further analysis.

Two types of plausible, language-pair specific explanations are described and illustrated in Section 4.2. The proposed inventory is **corpus-driven**: it was

developed by carefully analysing bilingual concordances and trying to identify the reason(s) that could have led translators not to use a prefixed word in the target text, despite the attested existence of at least one prefixed equivalent. The inventory is therefore highly language-pair dependent and will most probably need to be complemented and/or modified in follow-up studies. A step-by-step methodology was also set up to ensure the consistency of the choices made and the replicability of the study. In order to achieve this, I relied on dictionaries and a monolingual, web-derived corpus.

4.2. Partial Translational Equivalence

The translation data show that a major factor of morphological decrease is **partial translational equivalence** (see Adamska-Salaciak 2010 on the notion of equivalence in contrastive linguistics, translation theory and metalexicography). Translational equivalence can be partial from a semantic, connotative, stylistic and phraseological point of view (Van Roey 1990). I will focus on **partial semantic equivalence** (in particular diverging polysemy) and **partial phraseological equivalence**, as these two factors were found to play a significant role in morphological decrease (see Section 4.3).

4.2.1. Diverging polysemy

An interesting aspect of partial semantic equivalence is the phenomenon of **diverging polysemy** (Van Roey 1990; Fillmore and Atkins 2000; Altenberg and Granger 2002). It is a recurrent cross-linguistic contrast in the data examined here: *in-* or *un-*derivatives that are shared in English and French “have one (or more) meaning(s) in common, but (...) have further extended this (these) common meaning(s) in different, specific ways, so that the translational equivalence is only partial” (Van Roey 1990:90).

In this study, various dictionaries were used to check that the polysemy of English derivatives and their French counterparts was only partially overlapping: GR&C, English and French monolingual dictionaries and a dictionary of false friends (Van Roey et al. 1998). It was found that in such cases, the translator sometimes has to resort to a non-prefixed equivalent or a paraphrase in French because the meanings conveyed by the prefixed equivalent in the TL only partially overlap with the meaning expressed by the derivative in the source text. The French adjective *inconsidéré* (cf. 17), for example, can be used to qualify replies or actions, but not people (in the latter case, GR&C lists the paraphrases *qui manque d’égards/de considération* ‘that lacks consideration’). This cross-linguistic difference in polysemy is reflected in the target text, where the simplex adjective *égoïste* is used instead of *inconsidéré*. In the case of *un-*

healthy, the French prefixed adjective *malsain*⁹ cannot be used in the sense of ‘ill-looking’ (see 18, where the translator opted for the non-prefixed word *maladif* ‘sickly, unhealthy’). Interestingly, in the cases examined in this study, the diverging polysemy of the derivatives can be traced back to the diverging polysemy of their respective bases (here *considéré/considerate* and *sain/healthy*).

- (17) [PLECI fiction] En. He stared at her for a second or two, then tutted to himself and withdrew gracelessly, saying: ‘You can be so **inconsiderate** sometimes.’

Fr. Il la dévisagea une seconde ou deux, puis poussa un soupir et se retira de mauvaise grâce, en déclarant: “Tu peux être parfois tellement **égoïste!**”

- (18) [PLECI fiction] En. He was a skinny, black-haired, bespectacled boy who had the pinched, slightly **unhealthy** look of someone who has grown a lot in a short space of time.

Fr. Maigre, le cheveu noir, le garçon portait des lunettes et avait l’air hâve et légèrement **maladif** de quelqu’un qui a beaucoup grandi en peu de temps.

4.2.2. Partial phraseological equivalence

Partial phraseological equivalence can also result in language-pair specific morphological decrease. Phraseological meaning can broadly be defined as “the meaning of a word as far as it is determined by the preferential or exclusive company it keeps” (Van Roey 1990:107). Partial phraseological equivalence is observed when a SL prefixed item “combine[s] into a syntactic pattern with certain lexical items to the exclusion of others”, whereas its TL prefixed equivalent displays different phraseological restrictions (ibid. 124). In other words, so-called “‘preferred ways’ of saying things” (Altenberg 1998) can (and often do) differ across languages, which might lead to non-morphological translations and hence, morphological decrease. In this study, targeted searches in two tools were used to determine whether the TL prefixed equivalent was phraseologically less appropriate in the target text than the non-morphological equivalent selected by the translator: (i) the web-derived 1.3-billion-word *frWaC* corpus (available on the *Sketch Engine*; see Kilgariff et al. 2004) and (ii) Google searches restricted to the .fr domain.

Two situations arose: (i) the frequency data (*frWaC* and Google) show that the translator has opted for a non-morphological translation that is phraseologically more appropriate than the corresponding prefixed word in the TL (in which case we can safely conclude that partial phraseological equivalence is at

play) and (ii) the frequency data show that the translator has opted for a non-morphological translation that is phraseologically less appropriate than the corresponding prefixed word in the TL (in which case we cannot conclude that partial phraseological equivalence is at play and other plausible explanations must be sought). In (19), the adjective *unique* is used as the translation of *unmatched* (the prefixed adjectives *incomparable* and *inégalé* are listed in the GR&C bilingual entry of *unmatched*). The reason for this is that the collocation *occasion unique* is much more frequent than *occasion incomparable* or *occasion inégalée* (see Table 6).

- (19) [PLECI fiction] En. It's an **unmatched opportunity** to see what sort of effect continuous exposure to media images might have on dream content.
Fr. C'est une **occasion unique** d'observer quel genre d'effet peut avoir sur le contenu des rêves l'exposition continuelle a des projections d'images.

Table 6
Occasion inégalée vs occasion incomparable vs occasion unique

	<i>frWaC</i> corpus	Google (.fr) (rounded up frequencies)
<i>occasion inégalée</i>	0	250
<i>occasion incomparable</i>	0	650
<i>occasion unique</i>	1,512	815,000

Similarly, in (20) we see that the translator has opted for *la tête et les bras nus* 'naked' and not for *la tête et les bras découverts* 'uncovered' to render *head and arms uncovered*. The collocations *tête découverte* and *bras découverts* are attested in French, but they are less frequent than *tête nue* and *bras nus*.

- (20) [PLECI fiction] En. We sat on the rocks in the sun and she told us we must not leave the beach without covering ourselves or walk in the village in shorts or attempt to go in the church – this one was for me – with **head and arms uncovered**.
Fr. Nous nous assîmes tous les trois au soleil, sur les rochers, et elle nous expliqua qu'il ne fallait ni quitter la plage sans s'être rhabillés, ni se promener en short dans le village, ni pénétrer dans l'église -cela pour moi – **la tête et les bras nus**.

4.3. Summing Up: Impact of Each Factor

Now that the three major language-pair-related factors have been identified, described and illustrated, we will turn to the last question addressed in this section: what is the exact impact of the language-pair-related factors in terms of the amount of morphological decrease they generate? This is shown in *Table 7*, which presents a quantitative breakdown of the impact of each factor in our data set (tokens with a prefixed equivalent are not included here).

Table 7

Factors responsible for language-pair specific morphological decrease: *in*- and *un*-tokens

	<i>in</i> -tokens	<i>un</i> -tokens	Total
Productivity	6 (10%)	106 (48%)	112 (40%)
Diverging polysemy	25 (40%)	54 (25%)	79 (28%)
Partial phraseological equivalence	11 (17%)	6 (3%)	17 (6%)
Unexplained data	21 (33%)	53 (24%)	74 (26%)
Total	63 (100%)	219 (100%)	282 (100%)

It appears that productivity is the most prominent language-pair specific variable to influence morphological decrease (40%), followed by diverging polysemy (28%) and partial phraseological equivalence (6%) (the ‘unexplained data’ category, which represents *c.* one fourth of the non-morphological translations, is discussed in more detail below). In other words, productivity, diverging polysemy and partial phraseological equivalence can be used as plausible explanations of language-pair specific morphological decrease in 74% of the cases where a non-morphological translation is found in the French target texts. However, we see that the relative importance of each factor heavily depends on the SL prefixation pattern (*un*- or *in*-). For example, productivity plays a much more fundamental role in the *un*-data ($\chi^2 = 30.89$, $df = 1$, $p < 0.001$), as already suggested above.¹⁰ The other two language-pair-related factors are also unevenly distributed across the *in*- and *un*-data sets: diverging polysemy and partial phraseological equivalence are more frequent with *in*-derivatives ($\chi^2 = 5.47$, $df = 1$, $p < 0.05$ and $\chi^2 = 18.71$, $df = 1$, $p < 0.001$ respectively). Interestingly, the three language-pair specific variables equally lead to the two non-morphological translation patterns studied here, i.e. non-prefixed words and paraphrases (cf. *Table 8*).

Table 8

Factors responsible for language-pair specific morphological decrease: non-prefixed words and paraphrases

	Non-prefixed words	Paraphrases	Total
Productivity	53 (40%)	59 (40%)	112 (40%)
Diverging polysemy	41 (31%)	38 (26%)	79 (28%)
Partial phraseological equivalence	9 (7%)	8 (5%)	17 (6%)
Unexplained data	30 (22%)	44 (29%)	74 (26%)
Total	133 (100%)	149 (100%)	282 (100%)

It is essential to point out, however, that 26% of the data (the ‘unexplained data’ category in *Tables 7* and *8*) cannot be accounted for by means of language-pair specific variables at this stage. This data subset is illustrated in (21), where the translator favoured a translation by means of the verb *perturber* ‘disrupt’ and a syntactic negation *ne... pas* to render *an uninterrupted night’s sleep*. *Sommeil ininterrompu* is, however, a perfectly valid and frequent collocation in French.

- (21) [PLECI fiction] En. I expect that you will allow me, on that account if no other, an **uninterrupted** night’s – and here he turned off the light – sleep.
Fr. Je te demande instamment de **ne pas perturber**là, il éteignit la lumière... .. mon sommeil.

It is important to stress that these cases are as frequent in news (33 instances) as in fiction (41 instances) and cannot be attributed to variation across genre (see Section 5). It might be the case that some other language-pair-related factors have been overlooked here. It could also be that (at least some of) these (apparently) non-language-pair specific and non-genre-related instances are simply the result of translators’ personal choices and preferences. In addition, we cannot rule out the possibility that these cases represent a form of translation-induced morphological decrease. If this is a valid hypothesis, a similar phenomenon should be observed in the other translation direction (i.e. from French into English) and in other language pairs.

5. THE IMPACT OF GENRE

This section addresses and tentatively answers the following research question: what is the impact of genre on the word-formation features of translated language? As appears from *Table 9*, the translated data in news and fiction are markedly different: there are significantly more translations into prefixed words in news texts than in literary texts (65% vs 46%; $\chi^2 = 23.05$, $df = 1$, $p < 0.001$), while there are more translations into non-prefixed words in fiction than in news (30% vs 13%; $\chi^2 = 26.00$, $df = 1$, $p < 0.001$). Paraphrases, however, are equally frequent in the two genres.

Table 9

Morphological and non-morphological translations of *in*- and *un*-tokens in news and fiction

	News	Fiction
Prefixed words	242 (65%)	129 (46%)
Paraphrases	81 (22%)	68 (24%)
Non-prefixed words	50 (13%)	83 (30%)
Total	373	280

These cross-genre contrasts might be due to the fact that the two genres contain different kinds of *un*- and *in*-derivatives, which call for different translation strategies (which, in itself, would be a cross-genre contrast). Nevertheless, a closer look at the derivatives contained in the two sub-corpora reveals that (i) *un*- and *in*-derivatives without any prefixed equivalent in French are equally frequent in the two genres (69 tokens in news and 60 tokens in fiction corresponding to 40 and 45 types respectively; see *Table 10*); (ii) more generally, the same trends emerge when we take into consideration the yes/no-condition (attestation of a prefixed equivalent in French; see *Table 10*) and (iii) there are slightly more *un*-derivatives in fiction (64%) than in news (56%) ($\chi^2 = 4.13$, $df = 1$, $p < 0.05$). However, if we look at the distribution of the translation patterns across genres for the *un*- and *in*- prefixes considered separately, we observe the same patterns as those outlined above: fiction favours non-prefixed equivalents while news has a predilection for prefixed equivalents. This is shown in *Table 11*.

Table 10

Morphological and non-morphological translations of *in*- and *un*-tokens in news and fiction:
(lack of) prefixed equivalents in French

Prefixed equivalent?	Translation pattern	News	Fiction	χ^2 value
No	prefixed words	10	6	$\chi^2 = 0.19$, df = 1, p > 0.20
	paraphrases	35	24	$\chi^2 = 0.12$, df = 1, p > 0.20
	non-prefixed words	24	30	$\chi^2 = 3.86$, df = 1, p < 0.05
Total no		69	60	$\chi^2 = 0.86$, df = 1, p > 0.20
Yes	prefixed words	232	123	$\chi^2 = 21.52$, df = 1, p < 0.001
	paraphrases	46	44	$\chi^2 = 1.54$, df = 1, p > 0.20
	non-prefixed words	26	53	$\chi^2 = 21.51$, df = 1, p < 0.001
Total yes		304	220	$\chi^2 = 0.86$, df = 1, p > 0.20
Total		373	280	

Table 11

Morphological and non-morphological translations in news and fiction: *in*- vs *un*-

Prefix	Translation pattern	News	Fiction	χ^2 value
<i>in</i> -	prefixed words	138 (84%)	64 (63%)	$\chi^2 = 14.89$, df = 1, p < 0.001
	paraphrases	14 (9%)	13 (13%)	$\chi^2 = 1.28$, df = 1, p > 0.20
	non-prefixed words	12 (7%)	24 (24%)	$\chi^2 = 14.40$, df = 1, p < 0.001
Total		164 (100%)	101 (100%)	
<i>un</i> -	prefixed words	104 (50%)	65 (36%)	$\chi^2 = 7.09$, df = 1, p < 0.01
	paraphrases	67 (32%)	55 (31%)	$\chi^2 = 0.07$, df = 1, p > 0.20
	non-prefixed words	38 (18%)	59 (33%)	$\chi^2 = 11.23$, df = 1, p < 0.001
Total		209 (100%)	179 (100%)	

Interestingly, the fact that prefixed translations are favoured in news compared to fiction actually mirrors a contrast put forward in Lefer (2009), namely that prefixed words tend to be more frequent in written genres such as news editorials than in fiction, which is an in-between genre between spoken and written language (cf. dialogues; see Biber et al. 1999:16). Although it is difficult at this

stage to precisely pinpoint the reasons behind the observed cross-genre differences, the results suggest that genre variation shapes the word-formation features of translated language.

When we consider the three language-pair specific variables identified in Section 4, we find that there is no cross-genre variation (the χ^2 values all range between 0.14 and 3.17, $df = 1$, $p > 0.05$), which confirms that those patterns are triggered by SL- and TL-specific features rather than by genre differences (see *Table 12*).

Table 12
Factors responsible for language-pair specific morphological decrease across genres

	News	Fiction
Productivity	59	53
Diverging polysemy	30	49
Partial phraseological equivalence	9	8
Unexplained data	33	41
Total	131	151

More research is undeniably needed in the field to more precisely measure the exact impact of genre on the translation patterns of morphologically complex words and on the word-formation features of translated texts in general. The case study presented in this article does not allow us to draw any firm conclusions but the results clearly point to fundamental differences between news and fiction that would need to be further investigated by taking into consideration more genres and other word-formation patterns.

6. CONCLUSION

Baker (2004:170) has underlined the importance of “coming up with plausible explanations for the patterns we manage to identify” in translated language. It seems that this can only be achieved by confronting data from a wide range of cross-linguistic corpora, namely bilingual comparable corpora, bidirectional translation corpora and monolingual comparable corpora of translated and non-translated language (Bernardini 2010), and by cross-fertilizing insights from translation studies and contrastive linguistics (Granger 2003). This type of approach will make it possible to consider SL- and TL-related factors alongside translation-inherent trends and will go some way towards disentangling them.

The present article has proposed a three-tiered classification of the phenomena related to the word-formation features of translated texts: SL-induced morphological increase (affix-by-affix translation pattern), translation-inherent morphological decrease (normalization) and language-pair specific morphological decrease. The study of the French translation of the English *in-* and *un-* prefixes has tried to demonstrate the importance of considering the impact of SL- and TL-specific features on morphological decrease in translation by bringing to light the role of productivity, diverging polysemy and partial phraseological equivalence. It has also sketched the impact of variation across genres in the use of word-forming devices in translated language.

The approach adopted here now needs to be applied to other word-formation patterns and language pairs in a wide range of registers and genres, which will require more elaborate statistical techniques. More generally, the study has opened up numerous avenues for future research. First, it would be interesting to complement the study by examining differences between individual translators (do literary translators share a tendency to translate *un-* and *in-* prefixed words into French non-prefixed words?). It would also be of interest to test the validity of Tirkkonen-Condit's (2004) Unique Items Hypothesis in the field of word-formation, e.g. by studying the use of productive Germanic affixes such as the English *un-* or *-ness* in English target texts translated from a range of Romance languages. Other promising lines of inquiry would be to start off from morphologically complex forms in target texts and to investigate cases of morphological increase. Finally, it would also be important to distinguish between the translation patterns of lexicalized, productively-formed and creatively-formed complex words (see Bauer 2001 on the demarcation between lexical creativity and morphological productivity), as these may very well diverge, at least partially. Despite the many paths yet to be explored, I hope that this study has contributed to a better understanding of some of the factors governing the word-formation features of translated texts and has shown that the impact of source language systems would be worth re-assessing in corpus-based translation studies.

Notes

¹ See <<http://ronaldo.cs.tcd.ie/tec2/jnlp/>> for an online access to the corpus (last accessed 15 December 2011).

² See also Mason (2001) on the need to consider the importance of genre in corpus-based translation studies.

³ PLECI stands for *Poitiers-Louvain Échange de Corpus Informatisés*. For more information about the PLECI corpus, see <<http://www.uclouvain.be/en-cecl-pleci.html>> (last accessed 15 December 2011).

⁴ See <<http://www.macmillandictionary.com/>> (last accessed 15 December 2011).

⁵ In some rare cases, the paraphrase contains a word with a negative prefix.

⁶ They correspond to 324 prefixed types: 122 *in*-types and 202 *un*-types.

⁷ It is a very powerful tool as it gives access to all uses of a particular word or phrase in the two sides of the dictionary.

⁸ There are 5 *in*-tokens and 11 *un*-tokens that are translated into a prefixed word while no prefixed equivalents are listed in the GR&C. Two patterns emerge: (1) dictionary gaps (e.g. *indiscriminately* – *indistinctement*) and (2) cases of transposition (e.g. *inevitability* N – *inévitabile* A, *unselective* A – *indifféremment* ADV; where **inévitabilité* N and **insélectif* A are not attested in French), which are typically absent from dictionary entries.

⁹ *Mal-* is a negative prefix but it is no longer productive in present-day French.

¹⁰ Even though it is a minor phenomenon in the data set examined, it is interesting to note that suffixes, in particular *-ly* and *-ness*, also contribute to the need for a non-morphological equivalent in French (this is apparent in 3 *in*-tokens and 11 *un*-tokens).

References

- Adamska-Sałaciak, A. 2010. Examining Equivalence. *International Journal of Lexicography* Vol. 23. No. 4. 387–409.
- Altenberg, B. 1998. On the Phraseology of Spoken English: The Evidence of Recurrent Word Combinations. In: Cowie, A.P. (ed.) *Phraseology: Theory, Analysis and Applications*. Oxford: Oxford University Press. 101–122.
- Altenberg, B. 2004/2005. The Generic Person in English and Swedish. A Contrastive Study of *one* and *man*. *Languages in Contrast* Vol. 5. No. 1. 93–120.
- Altenberg, B. & Granger, S. 2002. Recent Trends in Cross-Linguistic Lexical Studies. In: Altenberg, B. & Granger, S. (eds) *Lexis in Contrast. Corpus-based Approaches*. Amsterdam: John Benjamins. 3–48.
- Astington, E. 1983. *Equivalences. Translation Difficulties and Devices French–English, English–French*. Cambridge: Cambridge University Press.
- Baayen, R.H. 2009. Corpus Linguistics in Morphology: Morphological Productivity. In: Lüdeling, A. & Kytö, M. (eds) *Corpus Linguistics. An International Handbook*. Berlin: Mouton de Gruyter. 899–919.
- Baker, M. 1993. Corpus Linguistics and Translation Studies. Implications and Applications. In: Baker, M., Francis, G. & Tognini-Bonelli, E. (eds) *Text and Technology. In Honour of John Sinclair*. Amsterdam: John Benjamins. 233–250.
- Baker, M. 1995. Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. *Target* Vol. 7. No. 2. 223–243.
- Baker, M. 2004. A Corpus-Based View of Similarity and Difference in Translation. *International Journal of Corpus Linguistics* Vol. 9. No. 2. 167–193.
- Barlow, M. 2008. Parallel Texts and Corpus-Based Contrastive Analysis. In: de los Ángeles Gómez González, M., Mackenzie, J. L. & González Álvarez, E.M. (eds) *Current Trends in Contrastive Linguistics. Functional and Cognitive Perspectives*. Amsterdam: John Benjamins. 101–121.
- Bauer, L. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
- Becher, V. 2010. Abandoning the Notion of “Translation-Inherent” Explication. Against a Dogma of Translation Studies. *Across Languages and Cultures* Vol. 11. No. 1. 1–28.
- Bernardini, S. 2002. Think-Aloud Protocols in Translation Research: Achievements, Limits, Future Prospects. *Target* Vol. 13. No. 2. 241–263.
- Bernardini, S. 2010. *Parallel corpora and the search for translation norms/universals*. Paper presented at the MATS 2010 symposium, Ghent University College, Belgium.

- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Boase-Beier, J. 2006. *Stylistic Approaches to Translation*. Manchester: St. Jerome.
- Cartoni, B. & Lefer, M.-A. 2011. Negation and Lexical Morphology across Languages: Insights from a Trilingual Translation Corpus. *Poznan Studies in Contemporary Linguistics* Vol. 47. No. 4. 795–843.
- Chuquet, H. & Paillard, M. 1987. *Approche linguistique des problèmes de traduction anglais-français*. Paris: Ophrys.
- Cosme, C. 2006. Clause Combining across Languages. A Corpus-Based Study of English–French Translation Shifts. *Languages in Contrast* Vol. 6. No. 1. 71–108.
- De Sutter, G., Delaere, I. & Plevoets, K. 2012. Lexical Lectometry in Corpus-based Translation Studies. Combining Profile-based Correspondence Analysis and Logistic Regression Modelling. In: Oakes, M. & Meng, J. (eds) *Quantitative Methods in Corpus-based Translation Studies*. Amsterdam: John Benjamins.
- Fillmore, C.J. & Atkins, B.T.S. 2000. Describing Polysemy: The Case of *crawl*. In: Ravin, Y. & Leacock, C. (eds) *Polysemy*. Oxford: Oxford University Press. 91–110.
- Frankenberg-Garcia, A. 2008. ‘Suggesting rather special facts’: A Corpus-based Study of Distinctive Lexical Distributions in Translated Texts. *Corpora* Vol. 3. No. 2. 195–211.
- Grabar, N., Tribout, D., Dal, G., Fradin, B., Hathout, N., Lignon, S., Namer, F., Plancq, C., Yvon, F. & Zweigenbaum, P. 2006. Productivité quantitative des suffixations par *-ité* et *-able* dans un corpus journalistique moderne. Paper presented at the TALN 2006 conference, Leuven: 10–13 April 2006.
- Granger, S. 2003. The Corpus Approach: A Common Way Forward for Contrastive Linguistics and Translation Studies? In: Granger, S., Lerot, J. & Petch-Tyson, S. (eds) *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam: Rodopi. 17–29.
- Granger, S., Lerot, J. & Petch-Tyson, S. (eds) 2003. *Corpus-based Approaches to Contrastive Linguistics and Translation Studies*. Amsterdam: Rodopi.
- Hatim, B. 2001. *Teaching and Researching Translation*. London: Longman.
- Johansson, S. 2007. *Seeing through Multilingual Corpora. On the Use of Corpora in Contrastive Studies*. Amsterdam: John Benjamins.
- Kenny, D. 2001. *Lexis and Creativity in Translation. A Corpus-based Study*. Manchester: St. Jerome.
- Kilgariff, A. 2001. Comparing Corpora. *International Journal of Corpus Linguistics* Vol. 6. No. 1. 1–37.
- Kilgariff, A., Rychly, P., Smrz, P. & Tugwell, D. 2004. The Sketch Engine. In: *Proceedings of Euralex 2004*. Lorient. 105–116.
- Koehn, P. 2005. *Europarl: A parallel corpus for statistical machine translation*. Paper presented at MT Summit 2005.
- Landers, C.E. 2001. *Literary Translation. A Practical Guide*. Clevedon: Multilingual Matters.
- Lefer, M.-A. 2009. *Exploring lexical morphology across languages: A Corpus-based study of prefixation in English and French writing*. PhD thesis. Louvain-la-Neuve: Université catholique de Louvain.
- Lefer, M.-A. 2010. Word-formation in English–French Bilingual Dictionaries: The Contribution of Bilingual Corpora. In: Dykstra, A. & Schoonheim, T. (eds) *Proceedings of the XIV Euralex International Congress*. Leeuwarden: Fryske Academy. 810–823.
- Lefer, M.-A. 2011. Contrastive Word-formation Today: Retrospect and Prospect. *Poznan Studies in Contemporary Linguistics* Vol. 47. No. 4. 645–682.
- Lefer, M.-A. & Cartoni, B. 2011. Prefixes in Contrast. Towards a Meaning-based Contrastive Methodology for Lexical Morphology. *Languages in Contrast* Vol. 11. No. 1. 86–104.

- Mason, I. 2001. Translator Behaviour and Language Usage: Some Constraints on Contrastive Studies. *Hermes* Vol. 26. 65–80.
- McEnery, A., Xiao, Z. & Tono, Y. 2006. *Corpus-based Language Studies: An Advanced Resource Book*. London & New York: Routledge.
- Neumann, S. 2008. *Contrastive register variation. A quantitative approach to the comparison of English and German*. Habilitationsschrift. Universität des Saarlandes.
- Olohan, M. 2004. *Introducing Corpora in Translation Studies*. London: Routledge.
- Paillard, M. 2000. *Lexicologie contrastive anglais-français. Formation des mots et construction du sens*. Paris: Ophrys.
- Plag, I., Dalton-Puffer, C. & Baayen, H. 1999. Morphological Productivity across Speech and Writing. *English Language and Linguistics* Vol. 3. No. 2. 209–228.
- Tirkkonen-Condit, S. 2004. Unique Items – Over- or Under-Represented in Translated Language? In: Mauranen, A. & Kuusimäki, P. (eds) *Translation Universals. Do They Exist?* Amsterdam: John Benjamins. 177–184.
- Van Hoof, H. 1989. *Traduire L'Anglais. Théorie et Pratique*. Paris: Duculot.
- Van Roey, J. 1990. *French–English Contrastive Lexicology. An Introduction*. Louvain-la-Neuve: Peeters.
- Van Roey, J., Granger, S. & Swallow, H. 1998. *Dictionnaire des faux-amis français-anglais, English–French*. Paris & Bruxelles: Duculot.
- Vinay, J.-P. & Darbelnet, J. 1958/1995. *Comparative Stylistics of French and English: A Methodology for Translation*. Amsterdam/Philadelphia: John Benjamins.
- Wang, K. & Qin, H. 2010. A Parallel Corpus-based Study of Translational Chinese. In: Xiao, R. (ed.) *Using Corpora in Contrastive and Translation Studies*. Newcastle upon Tyne: Cambridge Scholars Publishing. 164–181.
- Xiao, R. (ed.) 2010. *Using Corpora in Contrastive and Translation Studies*. Newcastle upon Tyne: Cambridge Scholars Publishing.