

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

IMPACT OF PARALLEL CORPORA AS TRANSLATION MEMORIES ON PHRASEOLOGICAL TRANSLATION QUALITY IN STUDENT TRANSLATIONS OF SPECIALIZED MEDICAL TEXTS

Heidi Verplaetse and An Lambrechts

KU Leuven (Research Group QLVL)

Abstract

The implementation of parallel corpora (source texts and their corresponding translations) as translation memories (TMs) in translation practice is widespread. This implementation may have different effects with regard to translation quality on stylistic, phraseological and terminological levels. In order to analyse these effects we conducted a small-scale experiment with 14 students from the English Translation Workshop: Translating Scientific/Medical Texts at KU Leuven. The students translated two short patient information leaflets (PILs) containing pre-selected phraseological test items (e.g.

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

lexical collocations, grammatical collocations or colligations, complex prepositions, etc.) from English into Dutch. 6 students used a CAT tool (*SDL Trados Studio*) with two specialized TMs and/or any available external resources, whereas 8 students used external resources only. Afterwards the translations were error annotated. The results showed that TM use speeds up the translation process. The CAT tool users in the experiment heavily relied on the TMs they had at their disposal. The lowest number of phraseological errors was generated using a TM in combination with external resources. The most popular bilingual resources among all students were bilingual dictionaries and corpus-based resources (e.g. *Linguee*). If deemed necessary by the students, the TM translations were overruled to create more concise or idiomatic translations. It is important that a critical stance is taken with regard to TM quality and students also need to be made aware of the possibilities additional resources (e.g. corpora) in translation have to offer.

Keywords: corpora, CAT tools, translation quality, translation training, phraseology, LSP, medical translation.

1. INTRODUCTION

Text corpora may improve (student) translations because they contain information which cannot be found in dictionaries, specifically in relation to terminology and idiomatic expressions (cf. phraseology) (Bowker, 1998; Frérot, 2009). Kübler (2011) confirms that parallel corpora, viz. source texts and their corresponding translations, are the

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

perfect tools for translation work: they offer both the terminology and the phraseology required for the translation task. The use and integration of parallel corpora into computer-assisted translation (CAT) tools not only makes it possible to exploit the benefits of corpora mentioned above, but also those of TMs (cf. Verplaetse & Heylen, 2015). One of the benefits is an increased translation speed, which leads to an increase of the translator's productivity and gains. TMs also improve the overall translation quality with an increase of the consistency at the stylistic, phraseological and terminological levels (Austermühl, 2006). However, it must be noted that TM quality should not be overrated: TM maintenance is of utmost importance, as sometimes TM output originates from different TMs and is provided by different translators. This means that the TM output provided depends largely on translator competence (Bowker, 2005, p. 15). Sometimes a translation proposed by the TM may also not fit the context in question (Bowker, 2005, p. 19), so a critical stance with regard to TM content is required. The fact that TMs contain isolated segments instead of full running texts (cf. corpora for the latter) may also negatively impact translation quality. (Bowker, 2005, p. 15). This pleads for the

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

provision of more contextual TMs and/or the use of (monolingual and comparable) corpora as reference material in translation (Frankenberg-Garcia, 2015, p. 254).

In this paper we assess the impact of parallel corpora as TMs on phraseological translation quality in specialized medical texts alongside the impact of any other available resources (dictionaries, glossaries, websites, ...). In doing so, we want to determine if the use of TMs, whether or not in combination with available external resources, positively influences phraseology in translation. The paper starts with a description of the theoretical framework used for the experiment outlined. This is followed by an outline of our methodology, the results, conclusions, limitations and further research.

2. THEORETICAL FRAMEWORK PHRASEOLOGY

Cowie (1994) defines phraseology as “the study of the structure, meaning and use of word combinations” (p. 3168). An even simpler definition is “the study of word combinations” (Zerkina and

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

Kostina, 2015, p. 144), with phraseological units consisting of at least two words. This feature is referred to as polylexicality, one of the criteria of phraseology (Colson, 2008, p. 193). Gries (2008) also adds that these phraseological units comprise at least two words, i.e. the “co-occurrence of a form of a lemma of a lexical item and one or more additional linguistic elements of various kinds” (p. 6).

Another criterion for phraseological units in the broad sense is fixedness (Colson, 2008, p. 193). Phraseological units occur more frequently than might be expected by chance (Gries, 2008, p. 6): from a syntactic point of view, the words comprising a phraseological unit are linked regularly (Mel’čuk, 2012, p. 32).

With regard to the classification of phraseological units, we distinguish referential and textual phraseological units (Granger and Paquot, 2008). A third category, communicative phrasemes, is not taken into account, as these do not often occur in specialized (medical, legal, technical, ...) text types. This category includes, among others, speech act formulae (e.g. greetings) and proverbs (e.g. *a place for everything and everything in its place*) (Granger and Paquot, 2008, p. 12)

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

Referential phrasemes convey content messages, referring to objects, phenomena or real-life facts, whereas textual phrasemes are grammaticalized sequences which structure and organize text content (Granger and Paquot, 2008, p. 12).

Referential phrasemes can, among others, be categorized into (lexical/grammatical) collocations, phrasal verbs, idioms, bi- and trinomials and similes. Examples of textual phrasemes are complex prepositions, complex conjunctions, linking adverbials and textual sentence stems (Granger and Paquot, 2008, p. 20). The relevant categories for this study will be described in the following sections.

Lexical and grammatical collocations

Nouns, adjectives, verbs and adverbs can all be part of lexical collocations. We distinguish different types of lexical collocations with regard to syntactic relationships (based on Benson, Benson and Ilson, 1986; Benson, Benson and Ilson, 1997, cited in Alsulayyi, 2015, p. 33; Benson, 1989, cited in Wu, 1996, p. 476; Hausmann, 1989, cited in Nesselhauf, 2005, p. 22; Nofal, 2012, p. 81-82):

1. noun + noun (e.g. *bowel movement*, *a group of patients*)

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

2. verb + noun (e.g. *to draw blood*)
3. noun + verb (e.g. *blood flows*)
4. adjective + noun (e.g. *nursing infants*)
5. verb + adverb (e.g. *move slowly*)
6. adverb + verb (e.g. *far exceed*)
7. adverb + adjective (e.g. *fully aware*)

Grammatical collocations, also referred to as colligations, consist of a dominant word (e.g. a noun, a verb, an adjective) and a preposition or a grammatical structure (Faghih & Mehdizadeh, 2013, p. 1604). We classify grammatical collocations as follows (based on Nofal, 2012, p. 81; Benson, Benson & Ilsen, 1997, cited in Alsulayyi, 2015, p. 33)

1. verb + preposition (e.g. *engage in*)
2. noun + preposition (e.g. *need for*)
3. preposition + noun (e.g. *in advance*)
4. adjective + preposition (e.g. *prone to*)

Phrasal verbs

We refer to *phrasal verbs* as a separate category (cf. Granger and Paquot, 2008). However, phrasal verbs can also be considered part of the category of grammatical collocations (Nofal, 2012, p. 81).

Phrasal verbs are verb-particle combinations (Nofal,

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

2012, p. 81). They can have an idiomatic or a non-idiomatic (literal) meaning, e.g. *stand up for our rights* versus *stand up from your chair*.¹

Idioms

Idioms can be defined as “relatively frozen expressions whose meanings do not reflect the meanings of their component parts” (Bahns, 1993, p. 57). In other words, from a semantic point of view, they are non-compositional (Granger and Paquot, 2008, p. 22). E.g. *break a leg*, which should not be taken literally, but instead is a fixed expression to wish someone good luck. As such, the expression’s meaning cannot be derived from its constituents. This implies that these expressions need to be learnt as a whole (Biber et al., 1999, p. 988).

As stated above, verb-particle combinations with idiomatic meanings are classified as idiomatic phrasal verbs. Relatively fixed expressions of which the meaning cannot be derived from its constituents and which have no verb-particle structure are

¹ In this paper we consider all verb-particle combinations to be phrasal verbs, regardless of their use, i.e. both the literal (non-idiomatic) or figurative (idiomatic) uses (Dixon, 1982; Greenbaum, 1985, cited in Olteanu, 2012, p. 68).

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

classified as idioms in this paper, e.g. *on edge*.

Complex prepositions

Complex prepositions consist of one (or two) prepositions with a noun, adverb or adjective intervening (Granger and Paquot, 2008, p. 23), e.g. *at initiation of, in response to*.

3. METHODOLOGY

The experiment was conducted with 14 MA students taking the English/Dutch Translation Workshop: Translating Scientific/Medical Texts at KU Leuven, campus Sint-Andries in Antwerp, Belgium. 6 students were experienced CAT tool users (*SDL Trados Studio*) as they took an optional CAT tool module. 8 students had no experience with CAT tools.

The students were asked to translate two patient information leaflets (henceforth abbreviated as PIL) from English into Dutch under different translation conditions, viz. with specialized TMs and/or external resources. External resources are any resources the students could find. During the experiment the students had to write down in a *Word*

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

document which external resources they used. Table 1 shows the set-up of the experiment.

Table 1. Set-up of the experiment

	CAT tool users	non-CAT tool users
PIL 1	TMs only	external resources only
PIL 2	TMs + external resources	external resources only

The PILs to be translated were PILs written by the researchers. It was a deliberate choice to use self-written PILs in this experiment. As we tested phraseological items mainly, it was not possible to use existing PILs, because we needed to collect as many relevant phraseological test items as possible in one short PIL. This is also why our PILs do not follow the structure of the EMA template with pre-defined headings, viz.: 1. What is X and what is it used for, 2. What you need to know before you <take> <use> X, 3. How to <take> <use> X, ...² In addition, had we used an existing PIL, there would have been the risk of students finding plenty of text portions literally in the TMs or in external resources, as fragments of the TMs can also be found in several

² Available at:

<http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000134.jsp>.

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

corpus-based resources, e.g. *Glosbe*, *Wordscope*, *Mymemory*, *Linguee*, *Reverso Context*.

The selected phraseological test items to be incorporated in the self-written PILs were chosen randomly and were drawn from PILs found on the website of the European Medicines Agency.³ In order to determine whether the TMs selected were likely to be useful translation resources for the CAT tool users, the frequencies of the phraseological test items were looked up in the translation memories which were going to be used in the experiment by the CAT tool users, viz. the translation memories from the EMA (Tiedemann, 2009) and the European Centre for Disease Prevention and Control (ECDC) (Steinberger et al., 2012).

PIL 1 (216 words) contained 16 phraseological test items. Table 2 shows the phraseological test items and their frequencies in the EMA TM (downloaded in April 2017), as well as their phraseological classifications. In general, the frequencies of the phraseological test items in the ECDC TM were very low, which is why only the frequencies in the

³ Available at: <<http://www.ema.europa.eu/ema/>>.

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

EMA TM are incorporated in Table 2. Because of the low frequencies in the ECDC TM, it is also assumed that the CAT tool users mostly used the EMA TM for their translation tasks.

Table 2. Phraseological items, classifications and TM frequencies for PIL 1

	Phraseological item	Phraseological classification	TM	Frequency in TM	Remarks
1	by means of	complex preposition	EMA	41	
2	blood drawn	lexical collocation: verb + noun (<i>draw blood</i>)	EMA	1	- frequency of <i>blood drawn</i> in the EMA TM: 1 - frequency of <i>draw blood</i> in the EMA TM: 0
3	act against	grammatical collocation: verb + preposition	EMA	21	
4	about to have	Idiom	EMA	10	
5	apart from	complex preposition	EMA	17	
6	prone to	grammatical collocation: adjective + preposition	EMA	74	
7	bowel movement	lexical collocation (noun + noun)	EMA	12	<i>bowel movements</i> occurs 24 times in the EMA TM
8	flare up	phrasal verb (idiomatic)	EMA	6	

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

9	seek help	lexical collocation: verb + noun	EMA	2	
10	child-bearing age	lexical collocation (adjective + noun)	EMA	47	47 is the sum of the frequencies of <i>child-bearing age</i> (14) and <i>childbearing age</i> (33) in the EMA TM
11	breast discharge	lexical collocation (noun + noun)	EMA	6	
12	at initiation of	complex preposition	EMA	20	
13	anxiety disorders	lexical collocation (noun + noun)	EMA	11	
14	keyed up	phrasal verb (idiomatic)	EMA	1	
15	suffer from	grammatical collocation (verb + preposition)	EMA	301	
16	concerned about	grammatical collocation (adjective + preposition)	EMA	27	

PIL 2 (224 words) contained 15 phraseological text items. Table 3 shows the phraseological test items and their frequencies in the EMA TM, as well as their phraseological classifications.

Table 3. Phraseological items, classifications and TM frequencies for PIL 2

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

	Phraseological item	Phraseological classification	TM	Frequency in TM
1	centrally acting	lexical collocation: adverb + adjective	EMA	30
2	make up for	phrasal verb (idiomatic)	EMA	193
3	fight disease	lexical collocation: verb + noun	EMA	3
4	iron deficiency	lexical collocation: noun + noun	EMA	21
5	carry around	phrasal verb (idiomatic)	EMA	0
6	on edge	Idiom	EMA	1
7	associated with	grammatical collocation: adjective + preposition	EMA	3585
8	bleeding gums	lexical collocation: adjective + noun	EMA	21
9	in close contact with	complex preposition	EMA	14
10	act on	grammatical collocation: verb + preposition	EMA	11
11	far exceed	lexical collocation: adverb + verb	EMA	2
12	amenable to	grammatical collocation: adjective + preposition	EMA	3
13	ill effects	lexical collocation: adjective + noun	EMA	5
14	at risk for	complex preposition	EMA	62

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

15	in accordance with	complex preposition	EMA	639
----	--------------------	---------------------	-----	-----

The translations were corrected using error categories based on the MeLLANGE error typology (Kübler et al., 2016). The MeLLANGE error typology also allows for user-defined categories. This led to the use of additional error categories based on the annotation guidelines for the English-Dutch language pair (Daems & Macken, 2013) and the SCATE error typology (Tezcan et al., 2018). One of the error categories added to the MeLLANGE error typology for the purpose of this paper is *word sense disambiguation (WSD)* (Daems & Macken, 2013; Tezcan et al., 2018). This occurs when a "Dutch word is a possible translation of the word in the ST [source text] but not of the meaning the word has in this context" (Daems & Macken, 2013, p. 26). Tezcan et al. (2018) consider *WSD* a sub-category of *mistranslation*. However, in this paper we make a distinction between *WSD* (viz. a partial mistranslation, as the translation is not incorrect, but does not fit the context) and (*full*) *mistranslation* as two separate error categories. In some cases, a translation may be correct in a general or other context, but not in a specific context or in a specialized text such as a patient information leaflet

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

(PIL). Consider, for example, the Dutch translation ‘vatbaar’ in ... *de tumor niet vatbaar is voor* (English: not prone to) *een chirurgische ingreep*. Dutch *vatbaar voor* is a possible translation of *amenable to*, but does not fit the given context of removing a tumour (cf. section 4). Therefore, this type of error is a partial mistranslation and needs to be classified in a more fine-grained manner as WSD. It is not a full mistranslation, since Dutch *vatbaar voor* is a correct translation for *amenable to* in other contexts.

4. RESULTS

Time

The CAT tool users, working with TMs only or with a combination of TMs and external resources, translated both PILs faster than the non-CAT tool users, who used external resources only.

Table 4: average translation time per translation condition for PIL 1 and PIL 2

	CAT tool users - TM only	non-CAT tool users - external resources only
PIL 1	23 minutes	31 minutes
	CAT tool users - TMs + external resources	non-CAT tool users - external resources only
PIL 2	25 minutes	28 minutes

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

Phraseological error frequencies

In this experiment particular focus is on the translation of the pre-selected phraseological test items.

If we look at the phraseological error frequencies, we see that the most frequently occurring error type is *(full) mistranslation* (5 cases), followed by *hyponymy* (4 cases) and *omission* (3 cases). Taking the different translation conditions into account there is no difference between the number of phraseological errors of CAT versus non-CAT tool users. In both groups 4 phraseological errors were made, with *(full) mistranslation errors* mostly, viz. 2 in each group.

Table 5 contains the 4 phraseological errors of the non-CAT tool users.

Table 5. Phraseological errors in PIL 1 (non-CAT tool users)

Source term	Target translation	Error type
bowel movement	Darmfuncties	Hyperonymy
flare up	plotseling opkomen	Mistranslation
flare up	gewaar worden	Mistranslation
concerned about	-	omission

The translation of *bowel movement* as *darmfuncties*

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

in the phrase ‘... *avoid medication which slows down or stops bowel movement*’ is regarded as *hyperonymy*, because the translation is too general. A possible correct Dutch translation is *darmbeweging* (literally ‘movement of the bowel’). In the sentence ‘*In case of side effects or if your symptoms flare up, seek medical help immediately.*’ the verb *flare up* indicates that the symptoms are already there and worsen. When *flare up* is translated as *plotseling opkomen* in Dutch, this implies the opposite: the symptoms are not there yet and suddenly they appear, hence a (*full*) *mistranslation error*.

In Dutch *gewaar worden* means *getting aware of something*, which is unrelated to *flare up* in this sentence. Therefore, this is also a (*full*) *mistranslation error*.

An *omission* occurs in the following Dutch translation of the phrase ‘... *you might ask a relative or close friend whether they are concerned about changes in your behavior.*’ as no translation is provided for *concerned about*: ... *vraag een familielid of vriend of er veranderingen in uw gedrag op te merken zijn* (... *ask a relative or close*

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

friend whether there are changes in your behaviour).

Table 6 contains the 4 phraseological errors of the CAT tool users.

Table 6. Phraseological errors in PIL 1 (CAT tool users)

Source term	Target translation	Error type
draw blood	bloed opnemen	Incorrect collocation
flare up	de kop opsteken	mistranslation
concerned about	bewust zijn van	mistranslation
prone to	meer vatbaar voor	addition

Bloed opnemen as a translation for the English *draw blood* is an incorrect collocation in Dutch. The correct collocation is *bloed afnemen*.

Like the non-CAT tool users' translation *plotseling opkomen* above, the CAT tool users' translation *de kop opsteken* for *flare up* in the sentence '*In case of side effects or if your symptoms flare up, seek medical help immediately*' refers to symptoms which only just manifested themselves for the first time (inchoative). The inchoative aspect is incorrect here, however, as the symptoms are already there and worsen when *flare up* is used.

Dutch *bewust zijn van* is a mistranslation in the phrase '*... you might ask a relative or close friend*

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

whether they are concerned about changes in your behavior' as in English this (back)translates as ... *you might ask a relative or close friend whether they are aware of changes in your behaviour.*

In the Dutch translation of *If you are prone to diarrhea*, ... an *addition* occurs: *Indien u meer vatbaar bent voor diarree*, ... (English: ... *if you are more prone to diarrhea* ...).

When comparing the phraseological errors in PIL 2 we see that only 1 error (*hyponymy*) was made by a CAT tool user, as opposed to 8 errors in the non-CAT tool user group. The most common error type for this group in PIL 2 is also *hyponymy* (3 cases), viz. the translation provided is too specific.

Table 7 contains the 8 phraseological errors of the non-CAT tool users.

Table 7. Phraseological errors in PIL 2 (non-CAT tool users)

Source term	Target translation	Error type
fight disease	die ziekte te bestrijden	hyponymy
fight disease	tegen de ziekte te vechten	hyponymy
on edge	Slap	mistranslation
far exceed	Overtreffen	omission
far exceed	overstijgen	omission
not amenable to	niet <u>meer</u> ... kan worden	addition
not amenable to	niet vatbaar voor	WSD

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

ill effects	misselijkheid	hyponymy
-------------	---------------	----------

In the English text the collocation *fight disease* has a general meaning. It should be translated as ... *de lymfocyten: cellen die het lichaam helpen ziektes te bestrijden*, with *disease* translated as the plural form *ziektes* in Dutch to indicate generality or indefinite reference. The Dutch translations *die ziekte te bestrijden* and *tegen de ziekte te vechten* imply definite reference (one specific disease), which is incorrect in this sentence.

The Dutch translation *slap* (English: *weak*) is the opposite of *on edge*, classifying this as a (*full*) *mistranslation error*.

An omission occurs in the Dutch translation of *far exceeds* in *This medicine will only be prescribed to you if the benefits far exceed the risks taken ...* as the adverb *far* (Dutch: *ver*) remains untranslated in Dutch: *Dit medicijn zal alleen voorgeschreven worden als de voordelen de risico's _overtreffen ...* and *Dit geneesmiddel zal alleen aan u worden voorgeschreven indien de voordelen ervan de risico's _overstijgen ...* .

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

For ... *the tumour is not amenable to a surgical procedure* an *addition* and *word sense disambiguation* occurred among the non-CAT tool users. With regard to the addition the Dutch translation ... *de tumor niet meer operatief behandeld kan worden* ... should be understood as ... *the tumour is not amenable anymore to a surgical procedure*.

WSD occurred in the Dutch translation ... *de tumor niet vatbaar is voor* (English: *not prone to*) *een chirurgische ingreep*. *Vatbaar voor* is a possible translation of *amenable to*, but does not fit the given context of removing a tumour.

The translation of *ill effects* in '*Stop taking this medicine immediately if you experience ill effects*' as *misselijkheid* (English: *nausea*) is again a hyponymy, because *misselijkheid* is too specific. The correct Dutch translation of *ill effects* is *bijwerkingen* (cf. *side effects*).

Table 8 contains the 1 phraseological error of the CAT tool users.

Table 8. Phraseological errors in PIL 2 (non-CAT tool users)

Source term	Target translation	Error type

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

fight disease	<u>de</u> ziekte te bestrijden	hyponymy
------------------	--------------------------------	----------

Fight disease was again translated incorrectly in a definite sense, relating to one specific disease (cf. the non-CAT tool users' translations of PIL 2).

Use of external resources

At times when the participants were allowed to use external resources (non-CAT tool users for both PIL translations, CAT tool users for the translation of PIL 2), they all used the general bilingual *Van Dale* English-Dutch dictionary. Livbjerg & Mees (2003) already established that particularly in general translation, students overuse dictionaries, focussing too much on lexical units instead of other factors such as context (p. 123). The second most used aid for both non-CAT tool and CAT tool users was *Linguee* (www.linguee.com) (79%), followed by *Taalvlinder* (www.taalvlinder.com), which contains various specialized glossaries (29%), and the search engine *Google* (29%).

We also found that CAT tool users looked up less information on subject-specific websites if they had access to them than non-CAT tool users.

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

TM use

For the translation of the pre-selected phraseological test items in PIL 1 and PIL 2 the CAT tool users (6) mostly used the translations provided by the TMs. For 14 out of 16 of the phraseological items in PIL 1 the translation most frequently used by the CAT tool users was present in the TMs. The first exception is *draw blood*, which was not in the TM in the particular sense intended in the PIL. The second exception was the translation of *breast discharge*, which was included in the TM as *afscheiding van melk uit de borst* (English: *discharge of milk from the breast*). 2 out of 6 CAT tool users kept this translation. However, *breast discharge* does not necessarily imply the discharge of milk. 4 out of 6 CAT tool users overruled the TM changing the TM translation provided into the shorter translation *afscheiding uit de borst* (English: *discharge from the breast*) (3 CAT tool users) or *borstafscheiding* (English: *breast discharge*) (1 CAT tool user), which does not necessarily imply the discharge of milk. 4 out of 6 CAT tool users overruled the TM changing the TM translation provided into *afscheiding uit de borst* (English: *discharge from the breast*) (3 CAT tool users) or *borstafscheiding* (English: *breast discharge*) (1

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

CAT tool user). For the translation of *keyed up* 3 out of 6 CAT tool users followed the TM, and used *gevoel van spanning* (English: *feeling of tension*). The other 3 CAT tool users overruled the TM translation by using the shorter translation *gespannen* (English: *tense*).

For 13 out of the 15 phraseological items in PIL 2 the translation most frequently used by the CAT tool users was present in the TMs. The first exception was *on edge*. The translation provided in the TM was *gevoel van irritatie* (English: *feeling of irritation*), used by 1 CAT tool user. 3 CAT tool users translated *on edge* as *geïrriteerd* (English: *irritated*). Other shorter translations were *gespannen* (English: *tense*) (1 CAT tool user) and *gevoel van spanning* (English: *feeling of tension*) (1 CAT tool user). However, irritation and tension are two different notions. The second exception is the translation of *far exceed* in *This medicine will only be prescribed to you if the benefits far exceed the risks taken ...* In the TM *far exceed* is translated as *ver overschrijden* (literal Dutch translation). Only 1 CAT tool user kept this translation. The others translated *far exceed* as *veel*

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

groter zijn dan (English: *be much larger than*) (3 CAT tool users) - cf. ... *als de voordelen veel groter zijn dan de risico's* ... (English: *if the benefits are much larger than the risks* ...) - and *ver overstijgen*. In Dutch *overstijgen* is a synonym of *overschrijden* (English: *exceed*).

5. DISCUSSION

With regard to the time needed for the translations the experiment described for this paper confirms that TM use speeds up the translation process. Provided that the TM is of good quality, it can be a useful aid if translation deadlines need to be met. If external resources have to be consulted in addition to a TM, the translation obviously takes longer than with a TM only. It is clear that translation takes longer when more resources are available for consultation. The experiment described here confirms that the time invested in consulting external sources needs to be weighed against the quality of TM: if a translator can fully rely on (quality) TMs consulting any additional resources may not weigh up against the potential quality benefits it might offer. This experiment confirms the findings from previous research stating that TMs

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

increase translation speed (cf. Verplaetse & Heylen, 2015).

In the experiment as a whole (*full*) *mistranslation* is the most common error type. It mostly occurs in the PIL 1 translations, which may imply that these were considered more difficult by the students. TMs, nor external resources have been useful to remedy the mistranslation issues in these PILs. External resources generated overly specific translations (cf. *hyponymy*) at times (e.g. *ill effects* translated as *misselijkheid* (English: *nausea*)). This occurred less often when TMs were used. This shows that in this experiment TM support helps to provide more accurate information, which is particularly important for PILs, as these are read by a broad group of readers who need accurate information and instructions.

In this experiment the use of TMs in combination with external resources yields fewer phraseological errors compared to using a TM only or external resources only. Thus, even though the combination of TMs and external resources resulted in longer translation times, it does lead to an increase in (phraseological) text quality in this experiment, viz.

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

fewer errors. Therefore, for translation tasks with more flexible deadlines for which unambiguous interpretation is key, this may imply that the support of additional resources apart from TMs is advisable. Bilingual dictionaries (*Van Dale*) and corpus-based resources (*Linguee*) were the preferred translation aids for both groups (CAT tool and non-CAT tool users) in the experiment. But even with access to external resources, CAT tool users seemed to rely more on the TMs at their disposal, as they looked up less information on subject-specific websites. However, as TMs of poor quality will obviously also negatively impact the overall translation quality, the use of additional resources next to TMs is a valuable asset in translation practice.

The high TM usage rate in the experiment confirms the reliance of CAT tool users on TMs. Only in a few cases did the CAT tool users overrule the translations of the phraseological items provided by the TMs, opting for shorter or more idiomatic translations, e.g. *keyed up*, translated as *gespannen (tense)* instead of the longer term *gevoel van spanning (feeling of tension)* provided by the TM. An example of an idiomatic translation is ... *als de voordelen veel groter zijn dan de risico's* ... (less

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

idiomatic in English: ... *if the benefits are much larger than the risks ...*) as opposed to ... *als de voordelen de risico's ver overschrijden ...* (English: *far exceed*), which is a more literal translation from the source language.

6. CONCLUSIONS

When strict deadlines are at play, TMs may be of great use to speed up the translation process. However, the quality of the TM content must be taken into account. The number of translators who contributed to the TM (including their individual translation competences) and agreement to the use of particular terms play a role in determining the TM quality. Translation students as well as professional translators need to be aware of these aspects when judging whether a TM suggestion is suitable for use in their actual translations. Therefore, it is important that in translation education a critical attitude is developed towards TM input.

As a TM is not a panacea, translators should also not hesitate to consult additional resources as a benchmark for the TM content, particularly when deadlines are less strict. In case of questionable TM

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

content, this can provide them with more reliable, concise or idiomatic translation solutions. Referring to additional, (parallel) corpus-based resources may enhance target text quality and general idiomaticity, notably on the phraseological level. In this experiment we saw that dictionaries and corpus-based resources (e.g. *Linguee*) rank among the most popular resources among our translation students. Students' translation look-up and searching strategies could be improved by introducing them to additional resources, such as corpus compilation and querying software (e.g. *AntConc*).

7. LIMITATIONS AND FURTHER RESEARCH

Only a limited number of students (14) was available for this experiment. A higher number of students may yield different results. In addition, the texts as well as the TMs used for the experiment were domain-specific. Other text types and/or TMs from other domains may impact phraseological translation quality in different ways.

The theoretical phraseological framework described was specifically designed for this experiment. Different classifications based on other theoretical frameworks might result in different experimental

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

outcomes.

Similar experiments with a higher number of students as well as professional translators can be envisaged for the future. In-depth research with regard to the quality of existing domain-specific TMs and corpora may also uncover areas of improvement for efficient TM/corpus use. Next to the impact of the integration of parallel corpora as TMs in CAT tools the integration of other types of corpora, e.g. monolingual original corpora (texts produced by native speakers) could also be researched (cf. Verplaetse & Lambrechts, forthcoming).

ACKNOWLEDGEMENT

The research for this paper was conducted under project fund STG/15/040/IOF for the project entitled *Impact of the use of corpora and CAT tools on translation quality and efficiency for translation of specialized texts*, granted by KU Leuven Internal Funds.

REFERENCES

- Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.
- Alsulayyi, M. N. (2015). The use of grammatical collocations by advanced Saudi EFL learners in the UK and KSA. *International Journal of English Linguistics*, 5,1, 32-43.
- Austermühl, F. (2006). Training translators to localize. In A. Pym, A. Perekstenko & B. Starink (Eds.), *Translation Technology and its Teaching*, Tarragona: *Intercultural Studies Group* (pp. 69-81). Tarragona: Intercultural Studies Group.
- Bahns, J. (1993). Lexical collocations: a contrastive view. *ELT journal*, 47, 1, 56-63.
- Benson, M., Benson, E., & Ilson R. (1986). *The BBI Combinatory Dictionary of English. A Guide to Word combinations*. Amsterdam: John Benjamin Publishing Company.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Grammar of spoken and written English*. Harlow: Longman.
- Bowker, L. (1998). Using specialized monolingual native-language corpora as a translation resource: a pilot study. *Meta: Journal des traducteurs / Meta: Translators' Journal*, 43, 4, 631-651.

- Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.
- Bowker, L. (2005). Productivity vs quality. A pilot study on the impact of translation memory systems. *Localisation Focus*, 4, 13-20.
- Colson, J. P. (2008). Cross-linguistic phraseological studies: An Overview. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 191-202). Amsterdam/Philadelphia: Meunier, John Benjamins Publishing Company.
- Cowie, A.P. (1994). Phraseology. In R.E. Asher (Ed.), *The Encyclopedia of Language and Linguistics* (pp. 3168-3171). Oxford: Oxford University Press.
- Daems, J. & Macken, L. (2013). *Annotation Guidelines for English-Dutch Translation Quality Assessment, version 1.0*. LT3 Technical Report-LT3 13.02. Retrieved 27 October, 2017 from <https://www.lt3.ugent.be/media/uploads/publications/2013/Technical%20Report%20TQA%20Annotation.pdf>
- Faghih, E., & Mehdizadeh, M. (2013). A contrastive analysis of patterns of grammatical collocations between the English “Animal Farm” and its Azeri Turkish translation. *Theory and Practice in*

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

Language Studies, 3, 9, 1603-1609.

Frankenberg-Garcia, A. (2015). Training translators to use corpora hands-on: challenges and reactions by a group of thirteen students at a UK university. *Corpora*, 10, 3, 351-380.

Frérot, C. (2009). Designing a corpus-based translation course for translation teaching and translator training. *International Journal of Translation*, 21, 1-2, 133-148.

Granger, S., & Paquot, M. (2008). *Disentangling the phraseological web*. Retrieved October 28, 2016 from https://dial.uclouvain.be/downloader/downloader.php?pid=boreal:75809&datastream=PDF_01

Gries, S. T. (2008). Phraseology and linguistic theory: A brief survey. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 3-25). Amsterdam: John Benjamins.

Kübler, N. (2011). Working with corpora for translation teaching in a French-speaking setting. In A. Frankenberg-Garcia, L. Flowerdew & G. Aston (Eds.), *New Trends in Corpora and Language*

Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.

Learning (pp. 62-80). London: Continuum.

Kübler, N., Mestivier, A., Pecman, M., & Zimina, M. (2016). Exploitation quantitative de corpus de traductions annotés selon la typologie d'erreurs pour améliorer les méthodes d'enseignement de la traduction spécialisée. *Actes des 13èmes Journées internationales d'analyse statistique des données textuelles (JADT 2016)*. Nice, France, 7-10 June 2016 (pp. 731-741).

Livbjerg, I. & Mees, I. M. (2003): Patterns of dictionary use in non-domain-specific translation. In F. Alves (ed.), *Triangulating Translation. Perspectives in Process Oriented Research* (pp.123-136). Amsterdam: John Benjamins, 123-136

Mel'čuk, I. (2012). Phraseology in the language, in the dictionary, and in the computer. *Yearbook of Phraseology*, 3, 1, 31-56.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins.

Nofal, K. H. (2012). Collocations in English and Arabic: A comparative study. *English Language and literature studies*, 2, 3, 75-93.

- Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.
- Olteanu, A. R. (2012). A Holistic Approach to Phrasal Verbs. Retrieved March 29, 2017 from <https://pdfs.semanticscholar.org/1d74/91eaaade58cf35fb244addffca906568d953.pdf>
- Steinberger, R., Eisele A., Klocek, S., Pilos S., & Schlüter, P. (2012). DGT-TM: A freely Available Translation Memory in 22 Languages. In *Proceedings of the 8th international conference on Language Resources and Evaluation (LREC 2012)*. Istanbul, Turkey, 21-27 May 2012.
- Tezcan, A., Hoste, V., & Macken, L. (2018). SCATE Taxonomy and Corpus of Machine Translation Errors. In G. Corpas Pastor, I. Durán-Muñoz (Eds.), *Trends in e-tools and resources for translators and interpreters* (pp. 219-248). Leiden: Brill/Rodopi.
- Tiedemann, J. (2009). News from OPUS - A Collection of Multilingual Parallel Corpora with Tools and Interfaces. In N. Nicolov, K. Bontcheva, G. Angelova & R. Mitkov (Eds.), *Recent Advances in Natural Language Processing (vol V)* (pp. 237-248). Amsterdam/Philadelphia: John Benjamins.

- Verplaetse, H. and Lambrechts, A. (2019). Impact of parallel corpora as translation memories on phraseological translation quality in student translations of specialized medical texts. *Current Trends in Translation Teaching and Learning E*, 6, 232 - 268.
- Verplaetse, H., Heylen, K. (2015). Parallel Corpora for Medical Translation Training: An Analysis of Impact on Student Performance. *CULT Conference*. Alicante, Spain, 26-29 May 2015.
- Verplaetse, H. & Lambrechts, A. (forthcoming). Dictionary, corpus and CAT tool use in legal and business translation: a comparative pilot study. *The Papers of the Linguistic Society of Belgium*
- Wu, W. S. (1996). Lexical collocations: One way to make passive vocabulary active. In *Proceedings of the 11th Conference on English Teaching and Learning in the Republic of China* (pp. 461-480). Taipei: The Crane Publishing Co.
- Zerkina, N., & Kostina, N. (2015). English Phraseology in Teaching: Interrelation of Theory and Practice. *Procedia-Social and Behavioral Sciences*, 199, 143-148