TRANSLATION EQUIVALENTS IN BRITISH AND BULGARIAN PRINT MEDIA

Irina Stoyanova-Georgieva[*]

Abstract: *The paper deals with collocations of intensifiers and adjectives existing in original English and Bulgarian texts and aims to pair them as possible translation equivalents, thus helping to map the possible renditions of one and the same concept in two languages. It uses a non-parallel, comparable corpus of letters to the editor, published in British and Bulgarian newspapers and magazines, and applies a method for extracting bilingual expressions from such corpora.*

Key words: *translation equivalents, collocation, intensifier*

> *Interpreting means making a bet on the sense of a text, among other things. This sense that a translator must find – and preserve, or recreate – is not hidden in any pure language, neither a divine reine Sprache nor any Mentalese. It is just the outcome of an interpretative inference that can or cannot be shared by other readers.* (see Eco 2001: 16).

1. Introduction

For centuries, the concept of translation equivalence and especially the idea of absolute translation equivalence has been considered by many not only as feasible but also as an alluring and enticing idea, promising straightforward problem solutions. Others, on the contrary, criticised it as being 'an illusion of symmetry between languages which hardly exists' (see Mary Snell-Hornby 1988: 22). As a result, the notion can generally be conventionalised as the apple of discord between practicing translators and pure theorists of translation studies (see Shveitser 1993). Despite this, the role of translation equivalents in the process of translation, foreign language acquisition (see Shveitser 1993; Mcenery, Xiao 2007), and contrastive studies of languages (see James 1980; Ebeling 1998) is undoubtedly outstanding. The reason is that it provides us with a chance not only to communicate the meaning of lexicographical units, to demonstrate the understanding of a foreign language text, but also to find and discuss any similarities and differences between two languages.

---

[*] Assistant at Shumen University, Department of English Studies, Shumen, Bulgaria, e-mail: i.stoyanova-georgieva@shu.bg.

Considering the abovementioned roles of translation, the search for successful translation equivalents has been the subject of numerous studies but almost all of the research has been based on parallel corpora (see Kaji et al. 1992; Dagan, Church 1994; Smadja et al. 1996; Ebeling 1998; Perdek 2012) somehow implying that non-parallel corpora are not a reliable source for extracting such information.

The material analysed here presents a field which has not been discussed before and concerns the successful rendition of English collocations of intensifiers and adjectives, through the use of the same syntactic construction in Bulgarian. Although the translation equivalents studied here do not deal with idiomatic or enigmatic terminological units, their importance for the correct representation of the actual position of the communicator and attitude towards the matter in question is crucial. It is essential not only for registering the means of expression used in both languages in one and the same sphere of communication (media texts), but also for the more detailed understanding of the respective languages, the choice of means of expression, their combination, and use in certain areas of life.

Suggested by Firth (1957) the term collocation, 'collocations of a given word are statements of the habitual or customary places of that word' (see Firth 1968: 181), was recognized by Halliday (1966), Greenbaum (1974), Stubbs (1995), and Partington (1998). In the current study the term shall be used to stand for such pairs of adverbs and adjectives. From the point of view of the current study, collocations represent further predicament as they require the existence of the translated concept into the target language culture (see Newmark 1988: 146).

2. Data and Methods

For the purposes of the present study a comparable bilingual corpus, divided in two subcorpora, was employed (see Atkins, Rundell 2008: 479). The choice was primarily motivated by the material studied, which, unfortunately, cannot provide us with an opportunity to use parallel corpora due to the lack of such. The decision to use a comparable corpus was also backed by the belief that 'even non-parallel corpora, which are non-translations, include some phrases and compounds that have the same meanings and functions' (see Tanaka, Matsuo 1999: 109). Consequently, the current study can serve as a proof to their postulation. The two subcorpora contain texts, written in English, and published in British newspapers and magazines, and identical texts, written in Bulgarian, and published in Bulgarian newspapers and magazines.

The corpus consists of 1200 authentic letters to the editor extracted from twelve different sources (100 letters from each media). Half of the examined media were English newspapers and magazines (The Financial Times, The Guardian,

The Independent, The Daily Express, The Economist, and The BBC Focus), they formed the first subcorpus, and the rest were Bulgarian (24 Chasa, Dnevnik, Duma, Lichna Drama (Personal Drama; trans. author's), Tita and National Geographic) and formed the second subcorpus. The amount of the whole corpus is 345222 words, 252906 belonging to the Bulgarian subcorpus (63%) and 92316 to the English one (27%). This means that the Bulgarian subcorpus is twice as big as the English. The reason for this discrepancy is the fact that letters do not and cannot have a fixed length and it is virtually impossible to make a whole corpus of letters having the same length. That is why in cases of comparisons, percentage is used to eliminate any risks of compromised data.

In order to receive reliable and trustworthy results all of the collocations, extracted from the original corpus, were translated by a group of informants, native Bulgarian speakers, with a very good command of the English language. The age of the informants ranged between 19 and 35, their language expertise was between B2-C2, according to the Common European Framework of Reference for Languages, in addition to this each of the participants had to translate not more than 30 pairs. Our primary aim was to allow the informants to provide best rendition of the collocation without running the risk of exhaustion, and the secondary was to ensure that each of the collocations was translated by at least 3 different people.

The initial corpus was analysed by Tanaka and Matsuo's method for collecting word sequences from each subcorpus by using syntactic patterns of translations as a clue for translation equivalence. The method then searched for equivalents among the collected expressions (see Tanaka, Matsuo 1999: 110). The abovementioned method was adopted because of its reliability in the evaluation of the candidates for translation equivalents and the fact that it was initially suggested for English and Japanese, the correlation of which bears some resemblance with that of English and Bulgarian. Japanese and Bulgarian both share the feature *different alphabet*. From a genealogical perspective, Bulgarian is even closer to English as it only belongs to a different branch of the same Indo-European language family, while Japanese belongs to a completely different – the Japonic language family.

The syntactic structure examined in the current study is an intensifier modifying adjective[*], which is a phenomenon widely-used in both languages, despite their differences. In order to be able to examine the cases of translation equivalence, the study was further limited by the choice of twenty English intensifiers and their Bulgarian counterparts (see Table 1). They cover all five paradigms of

---

[*] For the purposes of the study we also consider collocations with participles functioning as adjectives.

degree adverbs, offered by Paradis (1997: 27), as a method of differentiating the adverbs.

No	Category	Subcategory	English adverb	Bulgarian adverb
1.	Reinforcers	Maximizers	absolutely	абсолютно (absoljutno)
2.			completely	съвършено (savarsheno)
3.			extremely	изключително (izklyuchitelno)
4.			totally	напълно (napalno)
5.		Boosters	very	много (mnogo)
6.			highly	извънредно (izvunredno)
7.			really	наистина (naistina)
8.			pretty	твърде (tvarde)
9.			so	така (taka)
10.			too	прекалено (prekaleno)
11.	Attenuators	Moderators	almost	почти (pochti)
12.			quite	съвсем (savsem)
13.			rather	доста (dosta)
14.		Approximators	fairly	сравнително (sravnitelno)
15.			relatively	относително (otnositelno)
16.			hardly	едва (edva)
17.			nearly	приблизително (priblizitelno)
18.		Diminishers	a bit	мъничко (manichko)
19.			a little	малко (malko)
20.			somewhat	до известна степен (do izvestna stepen)

Table 1 Intensifiers, according to function, after Paradis's classification

The entries in Table 1 were compiled on the basis of the Cambridge Advanced Learner's Dictionary (2008), Merriam-Webster Dictionary, and PONS Business English-Bulgarian/Bulgarian-English Dictionary (2002). They, however, represent the meaning of the words as separate entities, and not as part of a certain context or in collocation with other lexical items.

In this case, only translation equivalents that respond to the previously established syntactic pattern of the adverb modifying adjective collocation are discussed, thus cases where the construction can be related through different

part-of-speech sequences are omitted. Consequently, this makes it possible to examine the similarities and dissimilarities between both languages, namely their ability to express one and the same notion through one and the same syntactic construction, despite their predetermined remoteness. In other words, as Kenning renders it "the mapping of correspondences between languages" not only sheds "light on the commonalities and differences between language pairs", but improves "the accuracy of descriptions of individual languages" (see Kenning 2010: 493).

According to Tanaka and Matsuo's model, in most cases of translation equivalence, the adjectives and adverbs building the collocation are also related to their counterparts in the target language. This feature was also used in order to estimate the correspondence of the expressions and to find the pairs with the highest probability chance. The study also adopted their word correspondence types: lexical, similarity, and co-occurrence correspondence (see Tanaka, Matsuo 1999: 111).

The corpus was then analysed, the strings of adverbs and adjectives were extracted and a list of the candidates for translation equivalents was compiled. Then based on the informants' data the possible translation equivalents were mapped. The results were divided according to the three criteria for correspondence, mentioned above. In order to achieve maximum quality, as all the informants were native Bulgarian speakers, all the translations of the previously extracted collocations were checked in Google Books Ngram Viewer and the British National Corpus (BNC) for any inconsistencies, while the Bulgarian translations were checked against the results in the Bulgarian National Corpus (BulNC).

Our hypothesis was that a considerable amount of translation equivalents can be extracted from the relatively small corpus of collocations, which will prove that despite the differences between both languages, similar means of expression conveying the attitude of the speaker towards the discussed matter are used.

3. Results

The collocations extracted from the corpus were 450: 153 English and 297 Bulgarian. The first phase of the research was aimed at the most obvious translation equivalents in the list of extracted collocations. As the English collocations in the corpus were almost 50% less than the Bulgarian ones, the translation of all English collocations was matched against the Bulgarian collocations. In case a match was encountered, the adverbs and adjectives, constituting the collocation, were also examined, according to the translations provided in PONS Business English-Bulgarian/Bulgarian-English Dictionary (2002). As a result, 43 equivalent pairs were found. According to Tanaka and Matsuo's word correspondence model, these pairs have the highest possible

correspondence rank – lexical (see Tanaka, Matsuo 1999: 112). This means that 28.10% of all the word pairs answering the syntactic model for the twenty predefined intensifiers can be considered direct lexical translation equivalents.

Among the translation equivalents with highest lexical correspondence rank were 2 word sequences containing the intensifier absolutely: absolutely correct and absolutely right, which three different informants decoded as абсолютно верен /absolyutno veren/, абсолютно правилно /absolyutno pravilno/, напълно прав /napulno prav/ and абсолютно верен /absolyutno veren/, абсолютно точно /absolyutno tochno/, напълно верен /napulno veren/, respectively. Because вярно /vyarno/ - верен /veren/ are forms of one and the same lexeme with верен acting as the lemma, we assume that абсолютно вярно /absolyutno vyarno/, which is one of the collocations extracted from the Bulgarian corpus, can be considered a translation equivalent for the abovementioned English word sequences. In order to receive maximal quality of the probable translation equivalents, the Bulgarian pair of words абсолютно вярно /absolyutno vyarno/ was also submitted for translation by the informants and the outcome was 2 different renditions: absolutely right and absolutely correct, which totally proves our previous decision was right, thus turning absolutely right and absolutely correct into synonyms. There were 2 examples of translation equivalents with the adverb extremely, where extremely skilled was translated as изключително умел /izklyuchitelno umel/, which corresponds directly to the same collocation in the Bulgarian subcorpus and its translation into Bulgarian also matches the initial English collocation.

The corpus also contained 2 examples of collocations with completely, which can also be translated by абсолютно, despite the fact that the latter has absolutely and not completely as lexical correspondent in Table 1. Both adverbs belong to the subcategory of maximizers, express almost the same grade, and are frequently categorised as synonyms: completely right – абсолютно вярно / absolyutno vyarno/ and completely unacceptable – абсолютно неприемливо / absolyutno nepriemlivo/.

The list contains 20 word sequences with very operating as intensifier modifying adjectives (see Table 2). The abovementioned procedure was followed for all of them. The outcome was a complete match, for all the translations in both directions. A closer examination of the collocations in which the intensifier very is found not only proves that very is the most prolific among all the intensifiers, explored in the study, both in the English and the Bulgarian subcorpus, but that it has the biggest number of lexical translation equivalents as well.

English		Bulgarian	
Source Language Collocation	**Translation of the Collocation**	**Source Language Collocation**	**Translation of the Collocation**
very good	много добър /mnogo dobar/	много добра/и/ър /mnogo dobar/	very good
very wide	много широк /mnogo shirok/, много обемен /mnogo obemen/	много широк /mnogo shirok/	very wide
very high	много висок /mnogo visok/	много висок/о /mnogo visok/	very high, very tall
very tall	много висок /mnogo visok/		
very different	много различен /mnogo razlichen/	много различни /mnogo razlichni/	very different
very dangerous	много опасен/но /mnogo opasen/	много опасни /mnogo opasni/	very dangerous
very bad	много лош/о /mnogo losh/	много лошо /mnogo losh/	very bad
very comfortable	много удобен /mnogo udoben/	много удобен /mnogo udoben/	very comfortable
very careful	много внимателен /mnogo vnimatelen/	много внимателен /mnogo vnimatelen/	very careful
very small	много малък /mnogo malak/	много малка/ък /mnogo malak/	really small
very valuable	много ценен/но /mnogo tsenen/	много ценни/а /mnogo tsenni/	very precious
very hard	много труден/но /mnogo truden/	много трудно /mnogo truden/	very hard
very simple	много прост/о /mnogo prost/	много прост /mnogo prost/	very simple
very expensive	много скъп /mnogo skap/	много скъпи /mnogo skapi/	very expensive
very wealthy	Много богат /mnogo bogat/, много заможен /mnogo zamozhen/	много богати /mnogo bogati/	very wealthy
very interesting	много интересен /mnogo interesen/	много интересен /mnogo interesen/	very interesting
very large	много голям /mnogo golyam/	много голям /mnogo golyam/	very big/ very large

Table 2 List of some of the collocations with very and their translation equivalents (lexical correspondence)

The results contain ten examples of lexical correspondence between word sequences containing the adverbs too and прекалено /prekaleno/ and твърде / tvarde/, respectively:

too expensive – прекалено скъп /prekaleno skap/

too long – твърде дълъг /tvarde dalag/

The first phase of the study uncovered seven more translation equivalents pairs: *highly skilled – изключително умели* /izklyuchitelno umeli/; *quite different* which can be connected with *съвсем различен* /savsem razlichen/*, съвсем други* /savsem drugi/*, and съвсем отделни* /savsem otdelni/*,* all three are considered synonyms in Bulgarian; *really silly* was transferred as *доста глупав* /dosta glupav/, which corresponds directly to *доста глуповата* /dosta glupovata/, extracted from the Bulgarian subcorpus, and rendered as *rather silly*; *nearly empty* and its translation equivalents *почти празен,* and *so powerful – така мощни* /taka moshtni/.

The choice of pairs in this phase of the study was primarily governed by the twenty adverbs, previously selected for the study, and their translations, given in Table 1, and the outcome was that almost 30% of all the extracted English collocations had their lexical match found in the Bulgarian subcorpus.

The second phase of the experiment covered the **similarity correspondence,** 'when the meaning of one word is similar to another' (see Tanaka and Matsuo 1999: 112). In order to provide an accurate classification of the word sequences, the translations of the English words were checked in the BulNet and the Google Books Ngram Viewer, which gave information on the word similarity.

Based on this, further twelve translation equivalents were established: 5 containing the intensifier *too*, 3 for *quite,* and 2 for *rather*, while *very* and *extremely* each had only one. The main dissimilarities between the members of each pair were the slight differences in the subcategories of the intensifiers, modifying the adjectives in the collocations. In most of the cases, the Bulgarian intensifier belonged to a subcategory which had a stronger or weaker reinforcing effect. Such is the case with *extremely*, *quite, rather,* and some cases of *too* where the Bulgarian adverb expressed a grade a bit lower than the one communicated by the English intensifier: *extremely difficult – прекалено сложни* /prekaleno slozhni/, *много сложна* /mnogo slozhna/ or a bit higher than the expected – *quite high – много висок* /mnogo visok/. Other cases featured the use of synonymous adjectives: *very informative – много поучителен* /mnogo pouchitelen/.

Taking into consideration the fact that the differences from a semantic perspective are not significant, we believe that these twelve examples can be added up to the total number of translation equivalents listed above. Thus, the total number of translation equivalents extracted from the corpus is 55, which represents 36% of all of the English word sequences.

The last phase of the research was an analysis of the word sequences belonging to the **co-occurrence correspondence**, 'translations of expressions containing word A often include another word B' (op. cit. 1999: 113). Here, only adjectives that were found at least three or more times in the corpora were considered so that a parallel can be made between the results from the BNC and the BulNC. *Successful, difficult, right, small, high,* and *simple* were classified as the most productive adjectives, but all of them were modified by three different adverbs. Unfortunately, for most of them the Bulgarian subcorpus presented only one combination and it usually did not fully comply with the semantics of the English adverbs, apart from *absolutely right* (218 times in the BNC) – *абсолютно вярно* /absolyutno vyarno/ (148 times in the BulNC), and *very simple* (229 times in the BNC) – *много прост* /mnogo prost/ (239 in the BulNC).

4. Conclusion

Based on the aforementioned results the following observations can be made:

1. Despite its size and limitations, the corpus of the study proves that non-parallel corpora can also be regarded as a reliable source for extraction of translation equivalents. The statement is clearly advocated by the fact that more than one third of all combinations of adverbs modifying adjectives from the English subcorpus found their counterpart in the Bulgarian section.

2. Secondly, considering the genealogical distance between the two languages, it is especially interesting that such a considerable number of English collocations, could be represented by the same syntactic construction and corresponding vocabulary, evidently describing one and the same concept of the reality. This positively states that the same concepts are used in the media texts in both countries, and they are manifested by the same lexical means and syntactic structures.

3. The detailed analysis of the excerpted material reveals that very (used 60 times) and много /mnogo/ (149 times) are the most frequent intensifiers in both languages. The fact that the English subcorpus comprises one third of the whole corpus, while the size of Bulgarian is more than two-thirds of it also explains the discrepancy between the number of intensifiers and collocations in the two subcorpora. However, the proportion of degree adverbs in both subcorpora is almost the same 0.19, for the English, and 0.15 for the Bulgarian. So further to the above considerations, the paper also demonstrates that, both languages make use of more or less the same amount and type of intensifiers and they collocate with more or less the same adjectives, the function and semantic categories of which may be the subject of further studies.

References:

Atanasov et al. 2002: Atanasov A., Baychev V., Voykov V., Georgiev V., Chileva G., Angelova E., M. Stoycheva. PONS Business English-Bulgarian / Bulgarian-English Dictionary. Stuttgart: PONS.

British National Corpus <http://corpus.byu.edu/bnc/> [Accessed 30 June 2017]

Bulgarian National Corpus <http://dcl.bas.bg/bulnc/> [Accessed 30 June 2017]

Common European Framework of Reference for Languages <https://www.coe.int/t/dg4/linguistic/source/framework_en.pdf> [Accessed 30 June 2017]

Dagan, Church 1994: Dagan I., K. Church. Termight: Identifying and Translating Technical Terminology. – Fourth Conference on Applied Natural Language Processing: ANLP-94. 34-40.

Ebeling 1998: Ebeling J. Contrastive Linguistics, Translation, and Parallel Corpora. – Meta. 434. 602-615

Google Books Ngram Viewer <https://books.google.com/ngrams> [Accessed 30 June 2017]

James 1980: James C. Contrastive Analysis. London: Longman.

Kaji, Kida, Morimoto 1992: Kaji H., Kida Y., Y. Morimoto. Learning Translation Templates from Bilingual Text. – 14th International Conference on Computational Linguistics, Nantes. 672-678.

Kenning 2010: Kenning M. M. What are Parallel and Comparable Corpora and How Can We Use Them? – In McCarthy M., A. O'Keeffe (eds.). The Routledge Handbook of Corpus Linguistics. Abingdon: Routledge. 487-501.

McEnery, Xiao 2007: McEnery A. M., R. Z. Xiao. Parallel and Comparable Corpora: What Are They Up To? – In: James G., G. Anderman (eds.). Incorporating Corpora: Translation and the Linguist. (Translating Europe). Clevedon, UK: Multilingual Matters.

Merriam-Webster Dictionary <https://www.merriam-webster.com/dictionary/> [Accessed 30 June 2017]

Newmark 1988: Newmark P. A Textbook of Translation. Hertfordshire: Prentice Hall.

Paradis 1997: Paradis C. Degree Modifiers of Adjectives in Spoken British English. Lund: Lund University Press.

Perdek 2012: Perdek M. Lexicographic Potential of Corpus Equivalents: The Case of English Phrasal Verbs and Their Polish Equivalents. University of Oslo, Department of Linguistics and Scandinavian Studies. 376-388.

Shveitser 1993: Shveitser A. Equivalence and Adequacy. – Translation as Social Action: Russian and Bulgarian Perspectives (Translation Studies). London, New York: Routledge. 47-56.

Smadja, McKeown, Hatzivassiloglou 1996: Smadja F., McKeown K. R., V. Hatzivassiloglou. Translating Collocations for Bilingual Lexicons: A Statistical Approach. – Computational Linguistics. 22(1), 1-38.

Snell-Hornby 1988: Snell-Hornby M. Translation Studies. An Integrated Approach. Amsterdam, Philadelphia: John Benjamins.

Tanaka, Matsuo 1999: Tanaka T., Matsuo Y. Extraction of Translation Equivalents from Non-Parallel Corpora. – TMI 99: Proceedings of 8th International Conference on Theoretical and Methodological Issues in Machine Translation. Chester. 109-119.

Walter et al. 2008: Walter E. et al. Cambridge Advanced Learner's Dictionary. Cambridge: Cambridge University Press.