English-Russian-Finnish Cross-Language Comparison of Phrasal Verb Translation Equivalents

Olga Mudraya[1], Scott S. L. Piao[2], Laura Löfberg[3], Paul Rayson[2], Dawn Archer[4]

[1] Department of Linguistics and Modern English Language, Lancaster University – UK
[2]Computing Department, Lancaster University – UK
[3]Department of Translation Studies, University of Tampere – Finland
[4]Department of Humanities, University of Central Lancashire – UK

Keywords: MWE extraction, phrasal verbs, English, Russian, Finnish

1. Introduction

A phraseological expression in a language may have equivalent expressions in other languages with different morpho-syntactic structures and semantic properties. Our recent experience in the Benedict Project (EU IST-2001-34237), in which a Finnish semantic lexicon compatible to the Lancaster English semantic lexicon (Rayson *et al*., 2004) has been built, shows that there can exist complex cross-language relations between English phraseological expressions, or multiword expressions (MWE), and their translation equivalents in other languages. A deeper understanding of such relations between phraseological expressions across languages is important for various tasks such as language learning, translation, automatic bilingual/multilingual lexicon extraction, etc. This work forms part of two research projects which involve porting of the Lancaster English semantic tagger (Rayson *et al*., 2004) to Russian and Finnish languages, as well as the continuing improvement of the English tagger. Previous computational approaches to MWE have mainly focussed on English, and there has been little previous research on computational approaches to MWE in Russian (Sharoff, 2004) or Finnish.

In this paper, we will compare some frequently occurring English phrasal verbs with their translation equivalents in Russian and Finnish, both in terms of morpho-syntactic structures and semantic properties. According to Longman's Grammar of Spoken and Written English (Biber at al. 1999), phrasal verbs are multi-word units consisting of a verb followed by an adverbial particle with a spatial or locative meaning (e.g. *find out*, *run away*, *catch up*, etc.). When combined together, they represent single semantic units with extended meanings that cannot be derived from the individual meanings of the parts. Our study focuses on the comparison between a list of English phrasal verbs, which share frequently occurring morpho-syntactic structures in the form of part-of-speech (POS) patterns, and their equivalent expressions in Russian and Finnish. We further narrow our focus on the phrasal verbs in which other word(s) can be embedded, such as *scare {somebody/something} off*.

2. Procedures

Firstly, for the purpose of comparison, a frequency list of MWE POS patterns is extracted from the semantic MWE lexicon of the Lancaster semantic tagger, which contains

approximately 18,800 MWE template entries. From the pattern-frequency list, the most frequent phrasal verb patterns occurring ten or more times in the list are selected. As a result, eight phrasal verb POS patterns whose frequencies range from 103 to 10 are extracted, as listed in Table 1 below.

Table 1 Frequent phrasal verb POS patterns extracted from USAS MWE lexicon[1]

Phrasal Verb POS Pattern	Frequency
VV* {Np/P*/R*} RP	103
VV* {R*} RP	34
VVD {Np/P*/R*} RP	29
VVD {R*} RP	15
VV* {R*} RL	13
VVD {R*} II NN1	11
VV* {Np/P*/R*} RL	10
VVD {R*} RL	10

where POS tags are members of the CLAWS C7 tagset[2]

VV* = any form of a lexical verb
VVD = past tense of a lexical verb (e.g. gave, worked)
Np = noun phrase
P* = any pronoun
R* = any adverb
RP = prepositional adverb, particle (e.g. about, in)
RL = locative adverb (e.g. alongside, forward)
II = general preposition
NN1 = singular common noun (e.g. book, girl)

Secondly, the English phrasal verbs sharing each of the patterns are retrieved, as shown in the following sample:

Pattern: VV* {Np/P*/R*} RP
break_VV* {Np/P*/R*} in_RP
find_VV* {Np/P*/R*} out_RP
lace_VV* {Np/P*/R*} up_RP
pay_VV* {Np/P*/R*} up_RP
wake_VV* {Np/P*/R*} up_RP
wipe_VV* {Np/P*/R*} out_RP

Pattern: VV* {Np/P*/R*} RL
bring_VV* {Np/P*/R*} together_RL
hold_VV* {Np/P*/R*} together_RL
hide_VV* {Np/P*/R*} away_RL
keep_VV* {Np/P*/R*} away_RL
leave_VV* {Np/P*/R*} behind_RL
scrape_VV* {Np/P*/R*} together_RL
take_VV* {Np/P*/R*} apart_RL

Pattern: VV* {R*} RP
die_VV* {R*} down_RP

Pattern: VV* {R*} RL
move_VV* {R*} ahead_RL

[1] One may notice that the tag *VVD* in this table is redundant as the tag *VV** covers it. This redundancy occurs due to the irregular English verbs, such as "buy", "drink", etc. For such verbs, *VV** cannot cover the tense inflectional variants them.
[2] For a full list, see http://www.comp.lancs.ac.uk/ucrel/claws7tags.html

Next, these and some similar (in terms of the general morpho-syntactic structural pattern) English phrasal verbs are compared with their equivalent expressions in Russian and Finnish, using dictionaries and corpora resources. The comparison is carried out both in terms of morpho-syntactic structures and semantic features. By doing so, we attempt to generalise differences in syntactic structure and typical cases of semantic diversities.

3. Cross-Language Comparison

Our study reveals some interesting cross-language structural divergences between the three languages under consideration. One of the marked divergences is that neither Russian nor Finnish has phrasal verbs. However, both languages employ a variety of grammatical mechanisms to convey the meanings expressed by English phrasal verbs.

3.1. English-Russian

We have observed that the Russian translation equivalents of the English phrasal verbs following the *VV** {*Np/P*/R**} *RP/RL* pattern tend to show a distinctly different structure from English morpho-syntactic structure. In Russian, the *RP/RL* elements in this type of English phrasal verbs, are generally expressed by means of verbal prefixes. For instance, English phrasal verbs with *RP* elements, which on their own would be function words such as *in*, *on*, *out*, *up*, *down*, etc., are usually translated into Russian as single verbs with an additional prefix, as in the following examples:

die down - <u>за</u>мирать
find out - <u>вы</u>яснить
lace up - <u>за</u>шнуровать
pay up - <u>о</u>платить
wake up – <u>раз</u>будить
wipe out - <u>вы</u>тирать

The *RP/RL* elements in English phrasal verbs denote the general spatial direction of the action or express its qualitative or quantitative characteristics, such as beginning, duration, completion, intensity, etc. Likewise, Russian verbal prefixes (*вы-*, *за-*, *на-*, *о/об-*, *раз-*, etc.) can indicate various qualities of actions and states, thus closely resembling the semantic functions of the *RP/RL* elements in English phrasal verbs. To illustrate, the Russian equivalent for *wipe out* is literally 'outwipe' (*<u>вы</u>тирать*).

On the other hand, English phrasal verbs with *RL* elements that independently may act as content words, such as *forward*, *ahead*, *behind*, *apart*, *together*, etc., are often expressed as verb + adverb in Russian, e.g. *bring together - сводить вместе, leave behind - оставлять позади*. However, this should be regarded more as a tendency rather than a rule, as some phrasal verbs in the second group can be translated by means of prefixation (i.e. *take apart – <u>раз</u>бирать; scrape together - <u>на</u>скрести*) whilst some others allow both means. For example, *move ahead* can be translated into Russian as either *<u>про</u>двигаться* (verb with prefix) or *двигаться вперёд* (MWE).

In addition, there are cases when an English phrasal verb of *VV** {*Np/P*/R**} *RP/RL* morpho-syntactic structure with multiple senses can be translated into different Russian verbs for different meanings:

die down - замирать, затихать, увядать, угасать
hold together - сплачивать(ся); держаться; держаться вместе; выдерживать
wipe out - вытирать, утирать; смывать, уничтожать

Furthermore, if an English phrasal verb is highly idiomatic, i.e. its meaning is unpredictable from the sum of its constituents' meanings, it will be rendered in Russian either with an idiomatic expression that may have a very different literal breakout or with a lengthy explanation. Yatskovich (1999) argues that it is almost impossible to create a consistent rigid system of lexical correspondences between English adverbial particles and Russian prefixes without encountering numerous debatable problems.

By way of illustration, Yatskovich shows that in the English sentence "He liked to *break in* his assistants slowly", neither the context nor the RP element hints at the real meaning of the phrasal verb *break in*, which according to the latest edition of the Longman Phrasal Verbs Dictionary (2000:36) means *to make someone get used to doing a job or activity, especially by letting them do it a little and then gradually making them do it more*. An earlier Russian edition of this dictionary (1997) treats this phrasal verb with a Russian idiomatic expression *вводить (кого-л.) в курс (новой работы и т.п.)*, literally meaning *to introduce (sb) to (a new job, etc.)*.

3.2. English-Finnish

Our study reveals that Finnish also employs various means to convey the meaning expressed by English phrasal verbs. Similar to the case for Russian, many English phrasal verbs can be translated into single verbs in Finnish, as shown in the following examples:

wake up – herättää; herätä
die down – lakata; vaieta; hiipua; tyyntyä
wipe out – hävittää
move ahead – edetä
take apart – purkaa; arvostella
scrape together – haalia

In some other cases, Finnish also uses phrases or idiomatic complement combinations {Verb+noun/adjective/adverb} which contain two or more words:

find out – saada selville
hold together – pysyä koossa; pitää yhtä
switch off – kytkeä pois päältä

There are also cases in which both the above practices are used:

hide away – piilottaa, panna piiloon
bring together – eheyttää; yhdistää, saattaa yhteen
leave behind – jättää; hylätä; luopua; jättää (jälkeensä), pudottaa kannoiltaan

However, often these longer expressions tend to be more colloquial than the single-word expressions. For instance, the phrase *panna piiloon* (= *hide away*) has a more formal equivalent of *piilottaa*.

Moreover, there are also some cases in which English phrasal verbs and their Finnish equivalent expressions share similar morpho-syntactic structures. Some of these expressions, such as *write down* - *kirjoittaa ylös (*literally means *write up*), are the result of the influence from other languages, mostly from Swedish. Often, there exist more native Finnish equivalent expressions for them, e.g. *kirjoittaa muistiin* as a substitute for *kirjoittaa ylös*.

Finally, the Finnish translations of some English phrasal verbs contain fixed collocations in which the object elements are indispensable, such as *tie up – sitoa kengännauhat*.

4. Conclusion

In this paper, we presented our comparative study of English phrasal verbs and their equivalent expressions in the Russian and Finnish languages. Our study not only reveals marked differences between English phrasal verbs and their equivalents in the two other languages, but also discovers some general corresponding structural patterns between them. This study is one part of a larger research effort to test the applicability of an existing framework for English semantic tagging to new languages, Finnish and Russian in this case. Phrasal verb templates form a significant part of the MWE resource for the English semantic tagger. This study has shown that such patterns usually have single-word translation equivalents in Russian and Finnish. Two conclusions can be drawn from the computational perspective: (i) a full morphological analysis pre-processing phase is vital for Finnish and Russian semantic taggers and (ii) the balance between single-word and MWE lexicons changes significantly across languages. Despite the limited depth and scale of our study, it helps us gain a deeper insight into the morpho-syntactic structural relations between equivalent expressions across the three languages, which can benefit various research areas including language teaching, contrastive linguistics and multilingual lexicon extraction.

Acknowledgements

This work is supported by the EU funded Benedict Project (IST-2001-34237) and the UK-EPSRC funded ASSIST Project (EP/C004574/1).

References

Английские фразовые глаголы. Англо-русский словарь(1997). Moscow: Russkiy Yazyk Publishers.

Biber D., S. Johansson, G. Leech, S. Conrad & E. Finegan (1999). *Longman grammar of spoken and written English*. Harlow: Pearson Education Limited.

Longman phrasal verbs dictionary (2000). Harlow: Pearson Education Limited.

Rayson P., D. Archer, S. L. Piao & T. McEnery (2004) The UCREL semantic analysis system. In *Proc. of the workshop on Beyond Named Entity Recognition Semantic Labelling for NLP Tasks in association with LREC 2004 (25th May 2004, Lisbon, Portugal)*, 7-12.

Sharoff, S. (2004). What is at stake: a case study of Russian expressions starting with a preposition. *Second ACL Workshop on Multiword Expressions: Integrating Processing*, July 2004, 17-23.

Yatskovich I. (1999) Some ways of translating English phrasal verbs into Russian. *Translation Journal* 3 (3), http://accurapid.com/journal/09russ.htm.