FACTORS INFLUENCING ENGLISH MULTI-WORD UNITS KNOWLEDGE IN RUSSIAN SPEAKERS. PARTIAL REPLICATION OF REVIER (2009)

Tatiana Khlebopros

M.A. Thesis
Supervised by: Dr. Raquel Serrano

Applied Linguistics and Language Acquisition in Multilingual Contexts
M.A. Program
Department of English and German Philology
Faculty of Philology
University of Barcelona
July 2018
ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my thesis supervisor Dr. Raquel Serrano, a professor and a researcher at the University of Barcelona, Catalonia, Spain, for the continuous support of my study, for her knowledge and expertise, for her insightful and thought-provoking comments and for her patience. Her guidance helped me in all the time of designing and administering the research and writing of this thesis. I could not wish for having a better advisor and mentor for my MA study.

Besides my supervisor, I would like to thank all the professors teaching on the Applied Linguistics and Language Acquisition in Multilingual Contexts programme for sharing their knowledge and skills, for their most helpful remarks on my projects which laid a strong foundation to my professional development.

My sincere thanks also goes to Dr. Revier, a professor and a researcher at the University College Capital, Copenhagen, Denmark, whose scientific work inspired me to choose a topic for this thesis, and who kindly gave me the opportunity to use the CONTRIX-2 test developed by him, and provided me with all the materials necessary to better understand the properties of this remarkable research tool.

I thank my fellow group mates for creating a stimulating environment, for engaging discussions, for the exchange of the ideas and thoughtful questions and critique which helped in shaping this MA thesis project.

Last but not the least, I would like to thank my beloved mother Irina Fomina, a researcher at the Institute of Basic Biologic Problems RAS, assistant professor at the Moscow State University, Russia, for being with me throughout writing this thesis. Her willingness to provide me with any possible assistance, from spiritual support to practical advising in respect with organising my work, was the most essential for me to write this thesis.
Understanding and producing English multi-word units is an essential ability, which needs to be mastered for communicating successfully in English. Yet, it has been shown to be challenging for non-native speakers of English and is often neglected in teaching practice. Very little is known about the factors that may facilitate or hinder the acquisition of multi-word units. The present study investigates the role of semantic transparency and the effect of the learners’ L1 on multi-word expression knowledge. Additionally, the study examines the relationship between the learners’ vocabulary size and their ability to use multi-word expressions. L1 Russian learners of English (n = 29) performed three online tests to measure their overall proficiency in English, productive vocabulary size and their knowledge of multi-word expressions of three levels of semantic transparency. The target multi-word expressions used in the study were subjected to L1-L2 congruence analysis in order to establish the level of similarity between English expressions and their equivalents in Russian. The statistical analysis of the collected data revealed that non-transparent multi-word expressions posed more difficulties to the learners than transparent and semi-transparent. These results differ from findings in an earlier study on Danish high-school and university students, who showed gradually better knowledge of multi-word expressions with the increase of transparency across the three levels. Although L1-L2 congruence had been previously shown to have a facilitating effect on multi-word expressions for speakers of Danish and German, no effect of L1-L2 congruence was found for the Russian-English language pair. Finally, strong correlations were found between vocabulary size and multi-word expression knowledge regardless of their semantic transparency level.

Keywords: multi-word expression, multi-word unit, lexical knowledge, L1-L2 congruence, Russian speaking learners of English.
Table of content

ACKNOWLEDGEMENTS ........................................................................................................ii

ABSTRACT .......................................................................................................................... iii

Table of content .................................................................................................................. iv

List of symbols, abbreviations, and other conventions ......................................................... vi

List of figures ......................................................................................................................... viii

List of tables ......................................................................................................................... viii

1. Introduction ..................................................................................................................... 1

2. Literature review .............................................................................................................. 2

   2.1 Multi-word units ......................................................................................................... 2

      2.1.1 MWU semantic categories .................................................................................. 3

      2.1.2 MWU congruence ............................................................................................. 6

      2.1.3 MWU and single word lexical knowledge ......................................................... 7

   2.2 Research aims .............................................................................................................. 8

3. Methodology .................................................................................................................... 9

   3.1 Participants ................................................................................................................ 9

   3.2 Data collection instruments and procedure ............................................................. 10

      3.2.1 CONTRIX-2 test (Appendix 2) ....................................................................... 10

      3.2.2 Productive Vocabulary Levels Test (PVLT) (Appendix 3) ............................... 11

      3.2.3 Quick Oxford Placement Test (Appendix 4) .................................................... 12
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Cronbach’s alpha</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>BNC</td>
<td>British National Corpus</td>
</tr>
<tr>
<td>CONTRIX-2</td>
<td>Constituent matrix test, version 2</td>
</tr>
<tr>
<td>Eng.</td>
<td>In the English language</td>
</tr>
<tr>
<td>EFL</td>
<td>English as a foreign language</td>
</tr>
<tr>
<td>FL</td>
<td>Foreign language</td>
</tr>
<tr>
<td>ESL</td>
<td>English as a second language</td>
</tr>
<tr>
<td><em>italics</em></td>
<td>Linguistic words (e.g. ‘make tea’)</td>
</tr>
<tr>
<td>L1</td>
<td>First language</td>
</tr>
<tr>
<td>L2</td>
<td>Second language</td>
</tr>
<tr>
<td>$M$</td>
<td>Mean</td>
</tr>
<tr>
<td>$M_x$</td>
<td>Mean for the category X</td>
</tr>
<tr>
<td>MWU(s)</td>
<td>Multi-word unit(s)</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of cases (participants or items)</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of cases in a subset (participants or items)</td>
</tr>
<tr>
<td>$\eta^2$</td>
<td>eta squared, a measure of the strength of the relationship</td>
</tr>
<tr>
<td>NNSs</td>
<td>Non-native speakers</td>
</tr>
<tr>
<td>NSs</td>
<td>Native speakers</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>NT</td>
<td>Non-transparent multi-word unit</td>
</tr>
<tr>
<td>p</td>
<td>Probability value</td>
</tr>
<tr>
<td>PVLT</td>
<td>Productive Vocabulary Levels Test</td>
</tr>
<tr>
<td>QOPT</td>
<td>Quick Oxford Placement Test</td>
</tr>
<tr>
<td>r</td>
<td>correlation</td>
</tr>
<tr>
<td>RQ</td>
<td>Research question</td>
</tr>
<tr>
<td>Rus.</td>
<td>In the Russian language</td>
</tr>
<tr>
<td>SCT</td>
<td>Simple completion test</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Spa.</td>
<td>In the Spanish language</td>
</tr>
<tr>
<td>SPSS</td>
<td>Statistical Package for the Social Science</td>
</tr>
<tr>
<td>ST</td>
<td>semi-transparent multi-word unit</td>
</tr>
<tr>
<td>TT</td>
<td>transparent multi-word unit</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>computed value of chi-square test</td>
</tr>
</tbody>
</table>
List of figures

Figure 1. Example of the CONTRIX-2 item................................................................. 11

Figure 2. Sentence-cloze format of PVLТ online form.............................................. 11

Figure 3. Test-section scores for higher and lower proficiency groups...................... 17

List of tables

Table 1. Distribution of the items with varying levels of congruence across semantic categories................................................................................................................................. 14

Table 2. Mean scores on CONTRIX-2 for the whole test and per section ................... 16

Table 3. Pairwise comparison of mean scores per semantic category (aggregate sample) .... 18

Table 4. Pairwise comparison of mean scores per semantic category (lower proficiency) .... 18

Table 5. Pairwise comparison of mean scores per semantic category (higher proficiency) ... 19

Table 6. Independent sample T-tests for CONTRIX and section test scores for higher and lower proficiency groups................................................................................................................................. 20

Table 7. Descriptive statistics for normalised mean scores for non-, mono-, bi- and tri-congruent CONTRIX-2 items................................................................................................................................. 21

Table 8. Spearman’s correlations for mean scores on PVLТ, QOPT and CONTRIX-2 as a whole and per section................................................................................................................................. 22

Table 8. Correlations between participants’ age and their scores on PVLТ, QOPT and CONTRIX-2 ......................................................................................................................................................... 35

Table 9. Normality test for mean scores for the whole CONTRIX-2 test and per sections in two proficiency groups................................................................................................................................. 36

Table 10. Descriptive statistics for raw mean scores for non-, mono-, bi- and tri-congruent
Table 11. Normality test for mean scores for non-, mono, bi and tri-congruent CONTRIX-2 items......................................................................................................................... 37

Table 12. Friedman test ranks for non-, mono-, bi- and tri-congruent CONTRIX-2 items........ 37

Table 13. Friedman test statistics for non-, mono-, bi- and tri-congruent CONTRIX-2 items. 37

Table 14. Normality test for mean scores on PVLT, QOPT and CONTRIX-2 as a whole and per section .................................................................................................................................................. 38

Table 15. The PVLT mean scores for Danish and Russian test-takers ................................. 38
1. Introduction

Over the past decades, lexical knowledge has been receiving increasingly more research attention and has been recognised as one of the most important aspects in second language acquisition. The earlier research on vocabulary knowledge mostly referred to the knowledge of single-word lexical items. Vocabulary was shown to play an imperative role in L2 production (Adolphs & Schmitt, 2003; Laufer & Nation, 1995; Nation, 1990; Stæhr, 2008), reading comprehension (Hu & Nation, 2000; Laufer, 1992; Laufer & Ravenhorst-Kalovski, 2010; Nation, 2001, 2006; Pulido, 2007), listening comprehension (Stæhr, 2009; van Zeeland & Schmitt, 2013) and writing skills (Mcnamara, Crossley, & Mccarthy, 2010; Schoonen, Gelderen, & Hulstijn, 2011). Vocabulary size has also been found to impact grammatical competence (Bardovi-Harlig, 2002; Myles, Hooper, & Mitchell, 1998; Wilkins, 1972) and phonological control (Bundgaard-Nielsen, Best, & Tyler, 2011).

However, more recent works have also addressed a broadening range of multi-word lexical items (Barfield & Gyllstad, 2009; Gyllstad, 2007; Schmitt, 2004). In the relevant literature, multi-word items are referred to as multi-word units (MWUs) (Erman & Warren, 2000; Nation, 2006; Pulido, 2007), multi-word expressions (Bardovi-Harlig, 2002; Briggs, 2014), multi-word phrases (Miralpeix, 2008; Smith & Murphy, 2015), phrasal expressions, formulas (McLaughlin, 1995), formulaic sequences (Schmitt, 2004) etc.

The importance of MWUs should not be underestimated. According to several scholars, the language produced by proficient users consists of strings of multi-word units rather than strings of individual words. Biber, Johansson, Leech, Conrad, and Finegan (1999) estimated the percentage of MWUs in spoken and written English as 30% and 21% respectively. Erman and Warren (2000) argue for even higher numbers: 58.6% and 52.3% for spoken and written English respectively. However, MWUs are shown to cause difficulties for L2 learners (Martinez & Murphy, 2011) and are often acquired late (Laufer & Waldman, 2011).

The literature review introduces key definitions and relevant theories with respect to MWUs, and, based on empirical findings, presents factors that may play a role in MWU acquisition.
2. Literature review

2.1 Multi-word units

Some scholars state that MWUs might be seen as combinations of two or more individual words preferred by the native speakers of the language over other possible word combinations that could express the same meaning (e.g. Nesselhauf, 2004). Other definitions imply that the meaning of the whole unit must be impossible to derive from the individual meanings of the MWU components (Cain & Towse, 2008). Bardovi-Harlig and Stringer (2017, p. 62) distinguish MWU from acquisitional formulas (e.g. ‘How do you spell... ’), which are learned holistically at the initial stages of L2 learning but are not bound to particular social situations. However, Briggs (2014, p. 76) argues that MWU phenomena cover a wide variety of sub-categories such as frequent collocations (‘make the bed’), phrasal verbs (‘make up one’s mind’), idioms (‘make ends meet’), fixed or semi-fixed phrases (‘make some ...’), acquisitional formulas (‘my name is ...’), etc. Thus definitions given by Briggs and Nesselhauf include MWUs with and without analysable meanings. The only feature inherent to MWUs across the various definitions is that they consist of more than one individual word bound in fixed or semi-fixed formulas. Thus, in the present study, the term MWU is used to address all kinds of word combinations which are commonly used together.

Empirical studies demonstrated that MWU knowledge promotes speech fluency (Wray, 2002), and language processing speed (Gibbs, Bogdanovich, Sykes, & Barr, 1997; Pawley & Syder, 1983), contributes to L2 grammatical development (Bardovi-Harlig, 2002), and are used as building blocks in creative production (Myles et al., 1998).

Various opinions exist with respect to storing and processing MWUs. According to Cutting and Bock (1997), MWU use simultaneously activates the concept related to the MWU itself and the concepts related to all the individual words in it in a lexical storage. Other scholars argue that as far as MWU refers to a singular meaning, MWU acts and is processed similarly to an individual word (Martinez & Schmitt, 2012; Weinert, 1995; Wray, 2002). Norbert Schmitt (2004) suggests that semantically opaque or semi-opaque items, where the meaning of an MWU cannot be derived from the meanings of its parts, might only be learned and retrieved as a whole and not as a set of individual items. This claim is to some extent supported by Nesselhauf’s (2004) study, where she analysed 318 English essays from
German sub-corpus of the International Corpus of Learner English (Granger, 1993). From this corpus she extracted 2082 collocations and divided them into 1) more semantically restricted (i.e. the verb could be used with a very limited number of nouns, e.g. ‘run a risk’) and 2) less semantically restricted (i.e. the verb could be used with a large yet still limited number of nouns, e.g. ‘make tea’). The analyses of acceptability of produced collocations performed by four experts revealed that less restricted collocations presented more difficulties to the learners than more restricted, i.e. more errors were made in the former (36% as compared to 23%). Nesselhauf concluded that less restricted collocations were learned and processed as inseparable units, therefore fewer errors occurred, as compared to less restricted collocations, which were processed as combinations of their components. However, no statistical analysis was performed on the collected data to establish significance of the findings. Moreover, in another study administrated by Huang (2001), the more restricted collocations were shown to trigger more incorrect or blank responses than less restricted or free word combinations. Therefore, the question remains open as to whether or not opacity has an effect on how MWUs are processed: as a whole or by components.

Three main factors are believed to influence MWU knowledge of non-native speakers. The most frequently addressed by researchers is semantic properties of the MWUs, i.e. how literate is the meaning of its components (Howarth, 1998; Revier, 2009). The second factor is the influence of learners' first language/s (L1), with a facilitating effect of positive transfer between languages which have semantically similar multi-word expressions (e.g. Spa. ‘romper el hielo’ and English ‘break the ice’) (Biskup, 1992; Nesselhauf, 2004). Finally, MWU knowledge is related to knowledge of single word lexical items (Huang, 2001; Martinez & Murphy, 2011).

The role of semantic category, L1 and vocabulary size will be addressed in the sub-sections below.

### 2.1.1 MWU semantic categories

Semantic transparency of MWUs is assumed to vary from full transparency to full opacity, with MWUs falling all along the range on this continuum (Boers & Lindstromberg, 2009; Gyllstad, 2007; Howarth, 1996, 1998; Nesselhauf, 2004; Revier, 2009). The earlier taxonomies are four-level models, with:
1) free combinations or open collocations, where the meaning of the MWU is composed of the literal meanings of its elements, e.g. ‘blow a trumpet’,

2) restricted collocations, which use one component in a specialized context, e.g. ‘blow a fuse’,

3) figurative idioms, with a metaphorical meaning that can still be derived from the literal meanings of its components, e.g. ‘blow your own trumpet’,

4) pure idioms, with a meaning unpredictable from the meanings of its components, e.g. ‘blow the gaff’. (Cowie, 1998; Howarth, 1998).

The three-level solution offered by Revier (2009) is more relevant for L2 English learners. His taxonomy lays on the assumption that when L2 learners first encounter a new MWU they attempt to decipher its meaning by analysing decontextualized meanings of its individual components. Since the core meanings of the individual words are most likely to be known by the learners, Revier’s classification is based on first dictionary entries only, with:

1) transparent (TT) MWUs, where the verb and the object noun are both used in their literal meanings, first dictionary entry meaning (e.g. ‘make tea’),

2) semi-transparent (ST) MWUs, where the noun is used in its literal meaning, while the verb is not (e.g. ‘make a complaint’),

3) non-transparent (NT) MWUs, where neither the verb nor the noun is used in its literal meaning (e.g. ‘run the show’) or the two components form a unitary meaning which could not be derived from their literal senses (e.g. ‘make the grade’),


Empirical evidence suggests a facilitating effect of semantic transparency on MWU knowledge. Huang (2001) investigated the effect of semantic properties on MWU knowledge on 60 Taiwanese university students with non-English majors. In her study she used 40 English word combinations taken from Longman Active American Idioms (Booker, 1994). The eliciting method was a sentence completion test where target items were embedded in the stimuli sentences with the part of speech of the missing word given in parenthesis.

The target items were divided into four groups according to their transparency based on Howarth’s (1998) classification. The examples of the test items are as follows:

1) free combinations (n = 11), “Those boys and girls don’t ________ organic juice.
They prefer something special, like pineapple or punch”. (Fill in a verb),

2) restricted collocations \((n = 25)\), “They also provide \___________\ drinks at the party for those who don’t drink alcohol”. (Fill in an adjective),

3) figurative idioms \((n = 34)\), “A lazy person always gives the excuse that working is not his cup of \___________\”. (Fill in a noun about food),

4) pure idioms \((n = 47)\), “The Browns bought a very cheap house, but later they spent a lot of money repairing it. We all think that they bought a \___________\ in a poke”. (Fill in a noun about animal).

The scores were found to be the highest \((M = 49.2\) out of 60) for the ‘easiest’ category (free combinations), and the lowest \((M = 4)\) for the ‘most difficult’ category (pure idiom), but there was no significant difference revealed between the two categories in the middle. Nevertheless, scores on restricted collocations \((M = 8.2)\) and figurative idioms \((M = 8.6)\) differed from the two others. Although the study shed some light on the effect of semantic transparency on English MWU knowledge among L2 English learners, several issues must be taken into account. First, the instrument used in the study was not standardised and no information on its validity and reliability was provided. Second, each of the four semantic categories included various semantic types (e.g. adjective-noun, verb-noun, noun-verb, etc.), thus there were very few MWUs in each sub-group. For instance, there was only one noun-noun figurative idiom. Third, the items varied in terms of frequency and complexity, which might have had an effect on the participants’ performance along with the effect of transparency. The potential effect of L1 interference was not addressed since the sample represented only one first language and no analysis of parallel MWUs between Chinese and English was performed.

In a more recent study, Revier (2009) investigated the effects of semantic transparency and first language influence by testing 132 Danish high school and university students for MWU knowledge, using his own instrument CONTRIX-2 previously submitted to validity and reliability analyses on a different sample. He targeted only verb+noun MWUs and used his own model of semantic transparency introduced above.

In respect with semantic transparency, it was found that knowledge of transparent items exceeded that of semi-transparent, and knowledge of semi-transparent MWUs was greater than that of non-transparent. In respect with MWU knowledge development, he analyses of
mean-scores for three proficiency levels as measured by the Productive Vocabulary Levels Test (Laufer & Nation, 1999) it showed that more transparent items are acquired earlier than less transparent items.

2.1.2 MWU congruence

The congruence status of an MWU reflects the possibility of positive or negative transfer from the learner’s L1 into L2 while understanding or producing MWUs. As mentioned above, learners tend to translate the unfamiliar L2 word combinations word-by-word. Similarly, they rely on literate translation from L1 to produce a new L2 MWU (Huang, 2001; Revier, 2009). Thus L1-L2 congruence, e.g. the existence of (partially) parallel fixed expressions, is likely to have a facilitating effect on MWU knowledge in L2.

Biskup (1992) investigated the effect of L1-L2 congruence on MWU knowledge by comparing the performance of two groups of students majoring in English, whose L1s were Polish (n = 34) and German (n = 28). The participants translated into English 23 MWUs from their respective L1s. No explanation is given on how the items were selected. Nor did Biskup mention what elicitation technique was implied, but it is possible that the items were given in isolation to be translated as a whole without prompted components. The given translations were analysed by three native speakers of English and taken as unacceptable (e.g. ‘drive a bookshop’), incorrect (e.g. ‘keep a bookshop’), correct alternative (e.g. ‘manage a bookshop’) and fully equivalent (e.g. ‘run a bookshop’). Polish students outscored Germans in production of full equivalents ($M = 5.21$ vs. $M = 3.82$). German speakers produced more correct and incorrect alternatives. The analyses of L1 influence revealed that 46% of errors made by Polish participants were due to L1 transfer, whereas for the Germans it was only 21%. Biskup concluded that the differences in performance between the groups might have been due to different teaching approaches with focus on accuracy in Poland as compared to focus on fluency in Germany. Thus Polish learners would produce more accurate translations while Germans would provide a greater number of acceptable alternatives for the unfamiliar items. However, it seems reasonable to assume an inherent effect of L1 on the participants’ performance.

Two studies attempted to find what factor has a stronger effect on MWU knowledge: semantic transparency or congruence. In the analysis performed by Nesselhauf (2004)
described in Section 2.1, the congruence status of English MWUs for L1 German was established as follows: German equivalents of English MWUs were translated literally back into English and then these translations were compared to target English MWUs. For the congruent items, only 27% of produced word combinations were deviant, whereas for non-congruent items the errors constituted 50%, with the strongest source of errors presented by non-congruence in articles and prepositions. As mentioned above the number of errors per semantic category was only 23% and 36% for more and less restricted MWUs respectively. Therefore, Nesselhauf concluded that non-congruence was more likely to induce errors in MWU production than opacity.

Revier (2009) also found a facilitating effect of L1-L2 congruence, however, his results showed it was smaller than the effect of transparency. One possible source of the disparity in Nesselhauf’s and Revier’s findings is the different approaches to categorising MWUs. While Nesselhauf relied on combinatory restrictions (how many nouns can be used with a given verb and vice versa), Revier looked at whether the components were used in their core meanings. Additionally, the congruence analysis was performed differently; Nesselhauf relied on her own judgment, whereas Revier used two Danish-English bilingual dictionaries to establish the most probable word-by-word L1-L2 translation. It is also possible that effect of congruence varies across languages.

2.1.3 MWU and single word lexical knowledge

At the moment there is no clear understanding of whether MWU knowledge develops in parallel with individual word knowledge amongst adult L2 learners (Barfield & Gyllstad, 2009). Martinez and Murphy (2011) tested 101 Brazilian adult learners of English by giving them short texts composed of the 2,000 most frequent English words organised in two parts; with words used as individual items (Test 1) and with the same words involved in the MWUs (Test 2). Half of the participants took Test 1 first; the other half took Test 2 first. The test was piloted on a group of English native speakers (n = 8) who scored 28 and 27.75 on comprehension questions for the two parts, which indicates that the level of difficulty of the two parts was similar for Test 1 and Test 2. The non-native participants in the study though scored significantly lower on the MWU part (t1(100) = 24.10, p < 0.001 vs. t2(100) = 14.76, p < 0.01) demonstrating that high idiomaticity of the text may hinder reading comprehension for L2 readers. This might be taken as evidence that a broad vocabulary does not necessarily
coincide with a good MWU knowledge. However, other studies showed that a greater vocabulary size is associated with a better MWU knowledge. In the Revier’s study discussed above, the participants’ scores on Productive Vocabulary Levels test showed a strong positive correlation with the scores on the MWU knowledge test \([r = .88, p < .01]\). Unfortunately, correlation analysis was not performed for the mean scores per semantic category. It might be anticipated that participants’ vocabulary knowledge would have had a stronger correlation with their performance on more transparent MWUs and vice versa. It would have aligned with Huang (2001), who came to the conclusion that knowledge of restricted MWUs (or all MWUs except free word combinations) does not parallel vocabulary knowledge.

### 2.2 Research aims

There is some evidence that semantic transparency has a facilitating effect on MWU use by L2 English speakers, however, the number of categories and categorisation criteria are not fully established. The only model that was tested successfully, i.e. statistically significant differences were found between all the categories, is the one offered by Revier (2009). This model is also the most relevant for L2 learners since the categorisation assumes knowledge of the core meanings of individual words composing MWUs. However, the model was only tested on speakers of L1 Danish learning English at high school or university aged 16-22. To further investigate the model, the present study focuses on another L1 (Russian) and age (only adults) and compare the results with those in Revier. Additionally, the participants’ proficiency is measured to address its influence on MWU knowledge.

It was shown that L1 plays a role in the acquisition of English MWUs, however, the strength of this effect is not yet established. In the studies discussed above the L1s were German and Danish, which belong to the Germanic group as well as English, thus to further investigate the effect of congruence, the present study looks at a more typologically different L1 (Russian) following the steps realised by Revier. The congruence effect for Russian is then compared to that of Danish, found by Revier.

Regarding a relationship between vocabulary size and MWU knowledge, no conclusive evidence was obtained of how single word knowledge may predict learners’ knowledge of MWUs across semantic categories. Therefore, the present study attempts to compare the
predictive power of single lexical item knowledge on the knowledge of transparent, semi-transparent and non-transparent MWUs.

Thus the research questions are the following:

RQ 1 Is semantic categorization in terms of transparency (i.e., transparent, semi-transparent and non-transparent) an important factor in the acquisition of English MWUs by Russian learners? If so, can evidence be found that MWU knowledge (in terms of semantic categories of transparency) develops at stages with overall language proficiency growth?

RQ 2 Is level of L1-L2 congruence an important factor in the acquisition of English MWUs by Russian learners? If so, what factor has a stronger effect on MWU knowledge, semantic category or L1-L2 congruence?

RQ 3 Is vocabulary knowledge related to MWU knowledge? If so, is vocabulary equally related to the knowledge of the MWUs from different semantic categories (TT, ST, NT)? What is a better predictor of MWU knowledge – vocabulary knowledge or overall proficiency level?

3. Methodology

In light of this study being a partial replication, the study design reflects the one of Revier’s (2009) study, taking a quantitative approach and adopting a quasi-experimental cross-sectional research design.

3.1 Participants

This study used convenience and snowball sampling to access potential participants whose first language is Russian and who speak English as a foreign language. Thirty-six people volunteered to participate and were sent the forms described in Section 3.2. There were 31 complete responses. Two cases were taken out of the analysis because the participants had identified themselves as Russian-Ukrainian bilinguals and it was decided that a strong knowledge of a language other than Russian may interfere with the analysis of MWU L1-L2 congruence. Seventeen participants also indicated speaking another foreign language, but
their proficiency in these languages was B1 and below. Boers and Lindstromberg (2009) report a great number of studies where formulaic language is shown to be an attribute of advanced L2 users (p.83). Thus a weaker knowledge of other foreign languages is not expected to influence English MWUs knowledge.

The final sample included 29 (5 male and 24 female) monolingual Russian speakers with different levels of proficiency in English and no strong knowledge of any other language/s. The age of the participants ranged from 21 to 67 ($M = 34.7$). Although it is a rather wide range, the preliminary analysis revealed no statistically significant correlation between the participants’ age and their scores on the three tests (see Table 9 Appendix 1), thus the group was analysed as age-homogenous.

3.2 Data collection instruments and procedure

3.2.1 CONTRIX-2 test (Appendix 2)

The CONTRIX-2 (Revier, 2009) is a diagnostic test designed to assess semi-productive\(^1\) knowledge of whole English MWUs as opposed to tests assessing the knowledge of collocation parts where testees only supply one component. It contains 45 verb-noun MWUs extracted from the British National Corpus (BCN). Based on Revier’s (2009) taxonomy, the items are divided into three semantic categories: transparent (e.g. break a bone), semi-transparent (e.g. break the law), and non-transparent (e.g. break the ice), 15 items per category. The items are presented as stimuli sentences with 3x3 matrixes of possible MWU parts each, as in the example below (Figure 1). Each matrix contains more than one meaningful MWU, but only one combination is correct. The 3x3 answer matrix reduces test sensitivity to guessing which was an issue in some earlier measurements using multiple choice, accept-reject or gap fill elicitation techniques (Bonk, 2000; Gyllstad, 2007; Mohammad, 2012; Schmitt, 1998).

The test was validated in several studies on speakers of L1 Danish and was shown to return reliable scores (Cronbach’s $\alpha = .86$). It is also the only test developed so far that measures

\(^1\) The elicitation technique of CONTRIX-2 requires the testees to produce the whole MWU without any components supplied, however, the test cannot be considered purely productive since the test-takers select three parts of the target item out of three options each component.
separately the knowledge of transparent, semi-transparent and non-transparent MWUs.

The scoring is dichotomous, i.e. the testee is awarded 1 if the answer is correct and 0 if it is partially or not at all correct.

Figure 1. Example of the CONTRIX-2 item

<table>
<thead>
<tr>
<th>Mary can never _____________. If you tell her something interesting, she immediately repeats it to everyone she knows.</th>
<th>save</th>
<th>a/an</th>
<th>secret</th>
</tr>
</thead>
<tbody>
<tr>
<td>keep the idea</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hold -- though</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2.2 Productive Vocabulary Levels Test (PVLT) (Appendix 3)

The PVLT was integrated into the study to investigate the relationship between knowledge of single words and MWUs. The PVLT measures productive vocabulary size by giving stimuli sentences and the first letters of the missing target word (see Figure 2 for an illustration of the format). The test was developed and subjected to validation by Laufer and Nation (1999), and is now available online (http://www.lextutor.ca/tests/levels/productive/). The test is applicable to learners at different levels of proficiency including advanced speakers (Laufer & Nation, 1999). As in Revier’s study, here the modified version is used. It consists of the items from the original PVLT general module (Test A) and does not include the academic module (Test B). The stimuli are grouped into four sections, 18 items per section. The stimuli in the sections elicit target words drawn from a clearly defined frequency level (2000, 3000, 5000 and 10000 most frequent English words).

Figure 2. Sentence-cloze format of PVLT online form

There are a doz__________ eggs in the basket.
The scoring system is adopted from Revier’s study, who awarded a score for each correct answer regardless of spelling.

### 3.2.3 Quick Oxford Placement Test (Appendix 4)

The participants took the Quick Oxford Placement Test (QOPT), which measures grammatical, pragmatic and vocabulary knowledge in 60 multiple choice questions (Allan, 1992). The QOPT was chosen because it had been shown to return reliable scores (Harrison, 1994). The scoring system is dichotomous, as with CONTRIX-2. The total scores correspond to proficiency levels as follows: 0-17 beginner, 18-29 elementary, 30-39 lower intermediate, 40-47 upper intermediate, 48-54 advanced and 55-60 very advanced.

In Revier’s study, the participants were initially divided into three proficiency groups according to their educational level (10th grade, 12th grade, and 2nd-year university). However, no correlation was found between the length of study and CONTRIX-2 scores. Therefore, Revier regrouped the participants according to their proficiency as measured by PVLT. For the present study, the participants were assigned to lower and higher proficiency groups based on their performance on the QOPT. The two groups were compared to explore the patterns of MWU knowledge development as a whole and with respect to semantic categories, namely transparent, semi-transparent and non-transparent.

### 3.2.4 L1-L2 congruence analysis

To investigate semantic similarity between the English MWUs and their counterparts in the participants’ L1 Russian, the CONTRIX-2 items were analysed for L1-L2 congruence, following, where possible, the technique used by Revier (2009).

First, the test items (N = 45) were translated into Russian by three proficient users of English and Russian. For the cases with multiple possible translations (items 12, 27, 33, 36, 39) it was assumed that Russian does not have stable equivalents and such items were assigned a non-congruent status. For instance, item 12 ‘run the show’ was translated as ‘править балом’ (‘pravit’ balom’ Eng. ‘rule ball’), ‘командовать парадом’ (‘komandovat’ paradom’ Eng.

---

2 Here and farther in the text, the Russian-English translation of the MWUs are given word-by-word.
'command parade’) and ‘руководить процессом’ (‘rukovodit’ protsessom’ Eng. ‘guide process’).

The other Russian word combinations were given as translations to the target English MWUs by at least two of the three experts. For example, item 4 ‘raise a finger’ was translated by all three translators as ‘пощевелить пальцем’ (‘poshevelit’ paltsem’ Eng. ‘move finger’). Item 38 ‘make tea’ was translated as ‘делать чай’ (‘delat’ chai’ Eng. ‘make/do tea’), where the Russian verb ‘делать’ can be equally translated into English as ‘make’ and ‘do’, by two translators and as ‘заваривать чай’ (‘zavarivat’ chai’ Eng. ‘brew tea’) by one translator. Here the translation ‘делать чай’ was taken as the Russian equivalent of the English collocation ‘make tea’.

The Russian equivalents of the English MWUs were subjected to further analysis in order to define the level of congruence. Each Russian word combination was split into its separate components: the verb, the article³, and the noun. Then, each component was translated into English by means of the Concise Oxford Russian English Dictionary to determine the most probable out-of-context translation of a given word. The first dictionary entry was used as it is more likely to be known as a decontextualized L2 translation of an L1 word (Sinclair, 1991). The assumption above was tested by prompted elicitation of the English translations of Russian words from Russian L1 speakers and it was found that the first dictionary entries were the most frequently given.

The pair of Russian-English words was considered congruent if the English translation offered by the dictionary matches the word in the original English MWU. For example, for the English MWU ‘get a taxi’ the established Russian equivalent is ‘взять такси’ (‘vzyat’ taksi’), where the verb ‘взять’ corresponds to ‘take’ and the noun ‘такси’ corresponds to English ‘taxi’. In some cases, a Russian word might be equally translated as two English words; here the pair was taken as congruent if one of the translations matched the original component of the English MWU. For example, the item 38 ‘make tea’ is ‘делать чай’ (‘delat’ chai’) in Russian, where the verb ‘делать’ can be translated either as ‘do’ or as ‘make’. Thus the verb pair делать-make was further treated as congruent.

---

³ As articles are absent in the Russian language, it was decided to code them as zero article in all cases.
If a Russian verb was translated into English as a phrasal verb that includes a target verb, the pair was considered non-congruent. For instance, for item 37 ‘give orders’ the Russian equivalent is ‘отдавать приказы’ (‘otdavat’ prikazi’), the verb ‘отдавать’ can be translated as ‘give away’. The pair was considered non-congruent due to the semantic difference between ‘give’ and ‘give away’ in English.

As the result of such analysis, the MWUs were divided into four groups: tri-congruent items where all three components are congruent, bi-congruent items where two out of three components are congruent, mono-congruent items with only one congruent component and non-congruent items (see all the items with translations and Roman alphabet equivalents in Appendix 5).

Examples of items (sources of non-congruence are underlined):

- tri-congruent (30) give money/ давать деньги,
- bi-congruent (10) draw a map/ нарисовать _ карту,
- mono-congruent items (3) break the ice/ растопить _ лед,
- non-congruent (9) make the grade/ добиться успеха.

Table 1 shows the distribution of the items with varying levels of congruence across semantic categories.

<table>
<thead>
<tr>
<th></th>
<th>Total number</th>
<th>Transparent items</th>
<th>Semi-transparent items</th>
<th>Non-transparent items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tri-congruent</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bi-congruent</td>
<td>11</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Mono-congruent</td>
<td>18</td>
<td>2</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Non-congruent</td>
<td>12</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

The MWU L1-L2 pairs should also match in terms of noun number (e.g. price vs. prices), but it was not of relevance here, as the noun number is fixed in the test instrument (i.e., the three noun options given for each item in the CONTRIX-2 have the same form in relation to
noun number).

3.2.5 Background questionnaire (Appendix 6)

The background questionnaire was developed and included in the study to sample monolingual Russian speakers and to collect the information on age and gender of the participants. Additionally, the questionnaire included questions about the participants’ language learning experience, but these data were not included in the analysis.

3.2.6 Study administration

The test battery was administered online in the time chosen autonomously by the participants. The online survey platform used to collect the data was www.qualtrics.com. As the data were collected online the order in which the tests were taken might have varied among the participants. However, it was recommended to take QOPT first, followed by PVLT and CONTRIX-2, and finally to fill in the questionnaire. There was no time limit set for any of the tests. The participants were warned to use their own knowledge only without consulting any sources, such as dictionaries, translators, grammar books, etc. The collected data were submitted to statistical analyses using SPSS.

4. Results

The results are presented with reference to the three research questions.

4.1 Research question 1

Is semantic categorization in terms of transparency (i.e., transparent, semi-transparent and non-transparent) an important factor in the acquisition of English MWUs by Russian learners? If so, can evidence be found that MWU knowledge (in terms of semantic categories of transparency) develops at stages with overall language proficiency growth?

To obtain evidence that semantic category plays a role in the acquisition of English MWUs, the comparison should be made between the three mean test-section scores on CONTRIX-2 (see Section 3.2.1 for the description) for the participants as an aggregate. Then, to investigate the role of proficiency, the mean test-section scores should be analysed at
different levels of proficiency. To address the latter, the participants were divided into two groups according to their scores on the QOPT test as based on median split. The alternative approach of splitting the participants into two groups based on levels according to QOPT was considered inconvenient since it would have resulted in putting into one group the participants of two-four different levels or unequal group sizes. Fourteen participants were assigned lower proficiency with the score range 29-46 \((M = 39)\) and 15 participants were assigned higher proficiency with the score range 47-58 \((M = 51)\). According to the original QOPT scoring, the level in the lower proficiency group (LPG) approximately corresponds to intermediate and upper intermediate level (scores 30-47) and the higher proficiency group (HPG) represents advanced and very advanced level (scores 48-60). The cross-sectional data were analysed to understand how MWU knowledge might have developed over time.

The mean section scores in Table 2 show that transparent items seem to present less challenge to the L2 English speaker than semi-transparent, and semi-transparent than non-transparent, i.e. MWU knowledge increases as a function of an increase in semantic transparency. This tendency can be observed in the two groups of proficiency \((M_{TT} = 6.14, M_{ST} = 5.57, M_{NT} = 3.86\) for LPG and \(M_{TT} = 9.27, M_{ST} = 8.53, M_{NT} = 6.07\) for HPG) and in the aggregate group \((M_{TT} = 7.76, M_{ST} = 7.10, M_{NT} = 5.00)\). The difference between groups on each section is even more noticeable than on the test as a whole.

**Table 2. Mean scores on CONTRIX-2 for the whole test and per section**

<table>
<thead>
<tr>
<th></th>
<th>Lower proficiency</th>
<th>Higher proficiency</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Transparent ((n = 15))</td>
<td>6.14</td>
<td>2.905</td>
<td>9.27</td>
</tr>
<tr>
<td>Semi-transparent ((n = 15))</td>
<td>5.57</td>
<td>3.106</td>
<td>8.53</td>
</tr>
<tr>
<td>Non-transparent ((n = 15))</td>
<td>3.86</td>
<td>1.875</td>
<td>6.07</td>
</tr>
<tr>
<td>Whole test ((N = 45))</td>
<td>15.57</td>
<td>6.548</td>
<td>24.53</td>
</tr>
</tbody>
</table>

The difference between the mean scores on TT and ST categories is smaller for the testees of higher proficiency. As in Revier’s study, knowledge of transparent, semi-transparent and non-transparent items appear to develop in parallel, although with a slightly slower progress for the latter (Figure 3).
A between-within subjects ANOVA was conducted to explore the effects of semantic categories (TT, ST and NT) and proficiency level (higher and lower) on MWU knowledge as measured by CONTRIX-2. Preliminary analyses by means of the Shapiro-Wilk test was performed to ensure there was no violation of the assumptions of normality ($p > .05$) (see Table 10 Appendix 1). The main effect of semantic category was statistically significant [Wilks’ Lambda = .430, $F(2, 26) = 17.214, p < .0005$] and according to Cohen’s (1988) classification, the size effect was very large [$\eta^2 = .570$]. The effect of proficiency was also statistically significant [$F(1,27) = 11.238, p < .005$] with a very large size effect [$\eta^2 = .294$]. There was no statistically significant interaction effect between semantic category and proficiency level [Wilks’ Lambda = .963, $p > .05$]. Therefore, the participants’ knowledge of MWUs is supported to some extent by semantic transparency and this effect changes across proficiency levels.

Pairwise comparisons for the aggregate population (Table 3) revealed no significant difference between transparent and semi-transparent MWU knowledge [$t(28) = 1.299, p > .05$], yet there was a significant difference between semi-transparent and non-transparent MWU knowledge [$t(28) = 4.017, p < .0005$] and between transparent and non-transparent MWU knowledge [$t(28) = 471, p < .000$].
**Table 3. Pairwise comparison of mean scores per semantic category (aggregate sample)**

<table>
<thead>
<tr>
<th>(I) Semantic category</th>
<th>(J) Semantic category</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>TT</td>
<td>ST</td>
<td>.652</td>
<td>.514</td>
<td>.645</td>
<td>-.659 - 1.964</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NT</td>
<td>2.743*</td>
<td>.471</td>
<td>.000</td>
<td>1.541 - 3.944</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST</td>
<td>TT</td>
<td>-.652</td>
<td>.514</td>
<td>.645</td>
<td>-1.964 - .659</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NT</td>
<td>2.090*</td>
<td>.529</td>
<td>.001</td>
<td>.741 - 3.440</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NT</td>
<td>TT</td>
<td>-2.743*</td>
<td>.471</td>
<td>.000</td>
<td>-3.944 - 1.541</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ST</td>
<td>-2.090*</td>
<td>.529</td>
<td>.001</td>
<td>-3.440 - .741</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on estimated marginal means

*. The mean difference is significant at

b. Adjustment for multiple comparisons: Bonferroni.

Similarly, pairwise comparisons for each of the two proficiency level do not reveal statistically significant differences between mean section scores for TT and ST items, whereas the NT mean scores were different from the others (Tables 4 and 5).

**Table 4. Pairwise comparison of mean scores per semantic category (lower proficiency)**

<table>
<thead>
<tr>
<th>(I) Semantic category</th>
<th>(J) Semantic category</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>.571</td>
<td>.581</td>
<td>.343</td>
<td>-.684 - 1.827</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.286*</td>
<td>.822</td>
<td>.016</td>
<td>.511 - 4.061</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-.571</td>
<td>.581</td>
<td>.343</td>
<td>-1.827 - .684</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.714*</td>
<td>.744</td>
<td>.038</td>
<td>.106 - 3.323</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-2.286*</td>
<td>.822</td>
<td>.016</td>
<td>-4.061 - -.511</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-1.714*</td>
<td>.744</td>
<td>.038</td>
<td>-3.323 - -.106</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on estimated marginal means

*. The mean difference is significant at

a. PROF_RANK = 1 (FILTER) = Selected

c. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).
Table 5. Pairwise comparison of mean scores per semantic category (higher proficiency)

<table>
<thead>
<tr>
<th>(I) Semantic category</th>
<th>(J) Semantic category</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Bound</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>.733</td>
<td>.831</td>
<td>.392</td>
<td>-1.048</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3.200*</td>
<td>.490</td>
<td>.000</td>
<td>2.149</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-.733</td>
<td>.831</td>
<td>.392</td>
<td>-2.515</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2.467*</td>
<td>.749</td>
<td>.005</td>
<td>.861</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-3.200*</td>
<td>.490</td>
<td>.000</td>
<td>-4.251</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-2.467*</td>
<td>.749</td>
<td>.005</td>
<td>-4.073</td>
</tr>
</tbody>
</table>

Based on estimated marginal means

* The mean difference is significant at the

a. PROF RANK = 2 (FILTER) = Selected

c. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

With respect to MWU knowledge development, CONTROX-2 mean scores were compared for the two proficiency levels by means of independent sample T-tests. More proficient English speakers significantly outperformed their less advanced counterparts on CONTRIX-2 as a whole and on each section ($p < .05$) (Table 6).
Table 6. Independent sample T-tests for CONTRIX and section test scores for higher and lower proficiency groups

<table>
<thead>
<tr>
<th></th>
<th>Levene’s Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
<th></th>
<th></th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
<td>t</td>
<td>df</td>
<td>Sig. (2-tailed)</td>
<td>Mean Diff.</td>
</tr>
<tr>
<td>CONTRIX-TOTAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EVA</td>
<td>.024</td>
<td>.878</td>
<td>-3.598</td>
<td>27</td>
<td>.001</td>
<td>-8.962</td>
</tr>
<tr>
<td>CONTRIX-TT</td>
<td></td>
<td></td>
<td>-2.839</td>
<td>27</td>
<td>.008</td>
<td>-3.124</td>
</tr>
<tr>
<td>EVA</td>
<td>.127</td>
<td>.724</td>
<td>-2.772</td>
<td>27</td>
<td>.010</td>
<td>-2.962</td>
</tr>
<tr>
<td>EVNA</td>
<td></td>
<td></td>
<td>-2.757</td>
<td>25.632</td>
<td>.011</td>
<td>-2.962</td>
</tr>
<tr>
<td>EVNA</td>
<td></td>
<td></td>
<td>-2.653</td>
<td>25.541</td>
<td>.014</td>
<td>-2.210</td>
</tr>
</tbody>
</table>

4.2 Research question 2

Is level of L1-L2 congruence an important factor in the acquisition of English MWUs by Russian learners? If so, what factor has a stronger effect on MWU knowledge, semantic category or L1-L2 congruence?

To establish whether there is there a facilitating effect of semantic similarity of MWU between Russian and English a preliminary L1-L2 congruence analysis was administered to the CONTRIX-2 items (N = 45) (see Section 3.2.4). The four groups of items: tri-congruent, bi-congruent, mono-congruent and non-congruent (n = 4, 11, 18 and 12 respectively) were subjected to statistical analysis. (see descriptive statistics for the raw mean scores for each group in Table 11 Appendix 1). Due to the unequal sizes of the four groups, the percentage

4 Equal variances assumed
5 Equal variances not assumed
of mean scores was used in the further analysis instead of the raw mean scores (Table 7).

Table 7. Descriptive statistics for normalised mean scores for non-, mono-, bi- and tri-congruent CONTRIX-2 items.

<table>
<thead>
<tr>
<th>Description</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTRIX-0-corg</td>
<td>29</td>
<td>8</td>
<td>83</td>
<td>42.53</td>
<td>22.310</td>
</tr>
<tr>
<td>CONTRIX-1-corg</td>
<td>29</td>
<td>17</td>
<td>78</td>
<td>42.15</td>
<td>16.502</td>
</tr>
<tr>
<td>CONTRIX-2-corg</td>
<td>29</td>
<td>9</td>
<td>100</td>
<td>51.72</td>
<td>22.414</td>
</tr>
<tr>
<td>CONTRIX-3-corg</td>
<td>29</td>
<td>0</td>
<td>100</td>
<td>45.69</td>
<td>27.604</td>
</tr>
</tbody>
</table>

Valid N (listwise) 29

Preliminary analysis by means of the Shapiro-Wilk test revealed a violation of normality assumption (see Table 12 Appendix 1). Therefore, to explore the effect of congruence level on MWU knowledge the Friedman test was administered (see Table 13 Appendix 1) and no statistically significant differences between the four groups were found ($p > 0.5$) (see Table 14 Appendix 1). Consequently, the expectation that MWU knowledge increases as a function of an increase in L1-L2 congruence has not been confirmed. Given the non-significant role of congruence in English MWU acquisition, it appears unnecessary to compare the effect of semantic category and the effect of congruence.

4.3 Research question 3

Is vocabulary knowledge related to MWU knowledge? If so, is vocabulary equally related to the knowledge of the MWUs from different semantic categories (TT, ST, NT)? What is a better predictor of MWU knowledge – vocabulary knowledge or overall proficiency level?

As there was a violation of the assumption of normality (see Table 15 Appendix 1), Spearman’s correlation coefficient was used to investigate the relationship between vocabulary knowledge and overall proficiency with participants’ performance on the test sections and the whole test (Table 8).
Table 8. Spearman’s correlations for mean scores on PVLT, QOPT and CONTRIX-2 as a whole and per section

<table>
<thead>
<tr>
<th>Correlations</th>
<th>PVLT- TOTAL</th>
<th>CONTRIX- TOTAL</th>
<th>CONTRIX- TT</th>
<th>CONTRIX- ST</th>
<th>CONTRIX- NT</th>
<th>OPT- TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s rho</td>
<td>1.000</td>
<td>.760**</td>
<td>.648**</td>
<td>.565**</td>
<td>.673**</td>
<td>.713**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.000</td>
<td>.000</td>
<td>.001</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX- TOTAL</td>
<td>.760**</td>
<td>1.000</td>
<td>.901**</td>
<td>.772**</td>
<td>.847**</td>
<td>.685**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX- TT</td>
<td>.648**</td>
<td>.901**</td>
<td>1.000</td>
<td>.658**</td>
<td>.668**</td>
<td>.591**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.001</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX- ST</td>
<td>.565**</td>
<td>.772**</td>
<td>.658**</td>
<td>1.000</td>
<td>.552**</td>
<td>.576**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.001</td>
<td>.000</td>
<td>.000</td>
<td>.002</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX- NT</td>
<td>.673**</td>
<td>.847**</td>
<td>.668**</td>
<td>.552**</td>
<td>1.000</td>
<td>.571**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.002</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>OPT- TOTAL</td>
<td>.713**</td>
<td>.685**</td>
<td>.591**</td>
<td>.576**</td>
<td>.571**</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).

According to Cohen (1998), there was a strong positive correlation between vocabulary and overall MWU knowledge \( r = -0.760, p < 0.0005 \), with greater vocabulary size associated with better MWU knowledge. Similarly, there was a strong positive correlation between proficiency and MWU knowledge \( r = -0.685, p < 0.0005 \).
In respect with semantic categories, vocabulary knowledge was found to have positive strong correlation with the knowledge of MWUs in all three semantic categories \[r_{TT} = .648, p < .0005, \quad r_{ST} = .565, p < .005, \quad \text{and} \quad r_{NT} = .673, p < .0005 \text{ respectively}\]. Similarly, overall proficiency correlated positively and strongly with all three test section scores \[r_{TT} = .591, p < .0005, \quad r_{ST} = .576, p < .005, \quad \text{and} \quad r_{NT} = .571, p < .0005 \text{ respectively}\].

5. Discussion

The aim of the present study was to deepen the understanding of the factors influencing English MWU knowledge in non-native speakers of English.

5.1 Role of MWU transparency

The first question addressed the role of semantic transparency. As in Revier’s (2009) study, MWU knowledge was shown to increase with level of proficiency both for the whole test and for its three sections, with less remarkable progress on the non-transparent MWU test section. The test-takers’ knowledge of non-transparent items was significantly lower than that of transparent and semi-transparent items regardless of proficiency level and vocabulary size. Therefore, although MWU competence generally increases with proficiency level growth, opacity posits a greater difficulty to the learners. This suggests that learners approach opaque items differently from more transparent items. As suggested by N. Schmitt (2004) and Nesselhauf (2004), while the meanings of more transparent items can be reconstructed from the meanings of the MWU components, non-transparent MWU are most likely to be learned as inseparable units. Thus, for TT and ST MWUs learners may successfully ‘guess’ the target expression by simply putting together individual words. However, due to limited linguistic proficiency and vocabulary knowledge in terms of breadth and depth, learners below a certain level face more difficulties to interpret MWUs by extending the literal meanings of their individual components, which may result in incorrect MWU use. From a pedagogical point of view, it means that non-transparent MWUs should be taught as independent lexical items through sufficient exposure to the MWU as a whole.

The present analysis did not reveal significant differences between participants’ knowledge of the transparent and semi-transparent MWUs, which differs from the results obtained by
Revier (2009), who found significant differences for all three pairs of categories for his population as an aggregate. One way to explain this difference is by looking at the proficiency level of Russian (present study) and Danish (Revier’s study) participants. In Revier’s study there was one group who performed similarly to the participants of the present study (TT=ST<NT). His most proficient test-takers (2nd-grade university students) scored significantly lower on the NT section while showing no difference in mean scores for TT and ST categories. Revier hypothesised that at some level the gap between TT and ST knowledge closes. The only measurement allowing comparison of participants’ proficiency in the present study and the one of Revier is PVLT. The mean scores obtained by the Russian group as a whole are lower than those of the most advanced Danish group, yet greater than those of the medium proficiency group (12th grade) (see Table 16 Appendix 1). If Revier’s hypothesis is correct, the Russian testees might have surpassed the critical point in their language development and mastered TT and ST MWUs at comparable levels and this turning point must lay somewhere between intermediate and advanced proficiency level. It is also possible that at some level (more advanced than in the present and Revier’s studies), the gap between ST and NT section scores disappears as well.

One more factor that could have influenced the results directly or indirectly is age. The participants in the Revier’s study who showed significant difference for the three test-section mean scores were high school students of the age 16-18, whereas the group who scored similarly to the Russian test-takers ($M_{TT} = M_{ST} < M_{NT}$) was aged 22. Although the present analyses revealed no effect of age from 21 to 63 on the participants’ test scores, it might be the case that the use of English and overall language-related experience is dissimilar for teenagers and adults, which has an effect on the pattern of MWU knowledge development in terms of semantic categories. That could account for the differences in performance between Danish students below 18 and Danish and Russian learners above 21.

In any case, a possible pedagogical implication is to introduce teaching non-transparent MWU into L2 English curriculum for the levels above intermediate, as their acquisition seems to be most difficult for the learners and is not certain to occur naturally.

5.2 Role of MWU L1-L2 congruence

The second research question addressed the role of L1-L2 congruence in English MWU
acquisition. Analysis of the CONTRIX-2 items resulted in the formation of four groups by level of congruence (4 non-, 11 mono-, 18 bi-, 12 tri-congruent MWUs). Unlike previous research performed with speakers of German or Danish (Nesselhauf, 2004; Revier, 2009), the present study found no effect of congruence for the Russian-English language pair, which can possibly be explained by a greater typological distance between Russian and English, as compared to German-English or Danish-English language pairs. Specifically, unlike English, German and Danish, Russian is a non-article language and according to Nesselhauf (2004) of all the MWU components, it is non-congruent non-lexical items that trigger the greatest amount of errors in MWU production (over 50%). Therefore, Russian speakers might be at disadvantage in terms of the potential facilitating effect of L1-L2 congruence as compared to Germans and Danish. From a pedagogical perspective, it means that Russian learners of English are unlikely to benefit from word-by-word L1-L2 translation strategy while learning English MWUs, and need more teaching attention with respect to word combinations. The same is possibly true for L1 speakers of any language typologically distant from English.

5.3 Role of vocabulary knowledge

The third research question focused on the relationship between single lexical item knowledge and knowledge of MWUs with different levels of transparency. The results of the analysis suggest that vocabulary size is as good a predictor of overall MWU knowledge as general proficiency level, which falls in line with Revier’s (2009) findings. As the more transparent items are assumed to be decipherable from the meanings of their elements, it was anticipated that vocabulary size would correlate more strongly with the knowledge of TT items, followed by ST items and finally by NT items. However, participants’ vocabulary size was found to correlate comparably with all three test-section scores. Thus, although NT word expressions are believed to be undecipherable from the meanings of their components, larger vocabulary seems to be a predictor of better NT MWU knowledge.

It is assumed that the less transparent MWU are learned and processed as a whole (Schmitt, 2004) rather than from the meanings of their individual components. Therefore, the nature of the correlation between vocabulary size and knowledge of MWU with different level of transparency might be different. For the more transparent items the learners would rely on vocabulary knowledge as a source of translation and thus exploit their knowledge of single lexical items to deal with MWUs. For the less transparent items, the probability of them
being known would be as big as the probability of any single lexical item of similar frequency to be known, thus they are treated by learners as low-frequent single lexical items, which are usually acquired by learners at the intermediate level and above.

Alternatively, the strong correlation between vocabulary knowledge and MWU knowledge shown by the test-takers in the present study might be explained by the properties of PVLT. According to the observations made by several researchers (Revier, 2009; Schmitt, 2010) its stimuli sentences, in many cases, contain partially supplied MWUs with one missing element which should be provided by the testees. For example, item 19 is ‘successful career’ (missing part is underlined), item 24 is ‘to get on one’s nerves’ or item 35 ‘stable relationship’ are MWUs of different transparency levels. Thus, the test measures single lexical item and MWU knowledge. To further explore the relationship between vocabulary and MWU knowledge, another vocabulary test targeting single words only should be used.

### 5.4 Limitations and suggestions for future research

The study is not exempt from several limitations. Firstly, the participants were recruited via online advertising, which resulted in a rather small sample ($N = 29$). Secondly, given the imparity of male ($n = 5$) and female ($n = 24$) participants, the effect of gender was left unexplored. Thirdly, since the measurements were administered via online forms, it was impossible to control that the participants did not consult any resources, such as dictionaries, translators or textbooks, and there was no time limit set for any measurement. Thus, the results may not accurately represent the participants’ linguistic proficiency, vocabulary size, and MWU knowledge. Fourthly, 17 participants reported some knowledge of language/s other than Russian and English; their responses were taken into the analysis since their self-reported proficiency in these languages was B1 and below, however, this information was not verified by any objective measurements. Fifthly, the scoring system for PVLT adopted from Revier’s study takes an answer as correct regardless of spelling, however it does not allow for grammar mistakes. Thus, for instance, omitted s-ending for the verbs in third person singular resulted in 0 score, yet it can be argued that the word was known and missing s-ending is, in fact, unlikely to cause breaks in communication. If the scoring system had taken such deviations as acceptable, the mean scores on the vocabulary test might have differed and it might have changed the results obtained regarding the relationship between vocabulary and MWU knowledge. These issues might be addressed in the future studies by
testing a bigger and more gender-balanced group of Russian speakers, who will take the tests under control of the researcher and with a reconsidered scoring system for vocabulary knowledge.

Among more specific issues, it is worth noting, that given the small number of participants and a rather short list of tested MWUs, the absence of L1-L2 congruence effect on MWU knowledge revealed in the analysis might be inconclusive. According to Nesselhauf (2004), the possibility of positive transfer from the L1 is not always exploited by the learners and vice versa, in some cases non-congruence does not impede the testee from producing a correct MWU. Moreover, there were very few CONTRIX-2 items at each congruence level (e.g. only 4 tri-congruent MWUs). Thus, the findings with respect to RQ 2 should be taken with caution. More studies should be performed with Russian speakers on a greater set of MWU items to obtain more generalizable results.

Furthermore, as mentioned in Section 5.3, the PVLT might not be the most objective measure for vocabulary size as referred to single word knowledge, since some items are in fact MWUs with missing components to be supplied by the testees. Thus future studies can use another test, specifically targeting the knowledge of individual lexical items.

Finally, it was shown that there might be two turning points in MWU acquisition; one where the learners’ knowledge of semi-transparent MWUs level with that of transparent MWUs, and the other, where the non-transparent MWU knowledge levels that of transparent and semi-transparent. Participants with various levels of proficiency should be tested in order to identify these two turning points in MWU knowledge development.

6. Conclusion

The present study contributed to a deeper understanding of the factors influencing English MWU knowledge in L2 English learners, specifically in the case of speakers of L1 Russian. It was shown that MWU knowledge develops in stages with linguistic proficiency growth with a slightly slower increase of non-transparent MWU knowledge. The results also showed that semantic transparency plays a facilitating role in MWU acquisition and that there is a turning point where the L2 English learners master semi-transparent MWUs at a level that is
comparable to that of transparent MWUs. It might be expected that another turning point exists at a more advanced level, where the knowledge of non-transparent MWUs reaches that of transparent and semi-transparent MWUs. Although no studies, including the present, have yet brought any conclusive evidence to support the latter assumption, these findings open a new direction for future research.

The study is the first to have analysed MWU congruence between L1 Russian and L2 English for the CONTRIX-2 items. Based on this analysis, it was shown that L1-L2 congruence has no effect on MWU knowledge for Russian learners of EFL/ESL. After comparing the present findings with other relevant works, where the effect of L1-l2 congruence was found, it was concluded that congruence is likely to play a role for languages typologically close to English, and is likely to have little or no effect for the typologically different languages. Alternatively, in order to assess different degrees of congruence, another congruence categorisation should be adopted for typologically different L1s. The assumption that L1-L2 congruence effect on English MWU knowledge may vary across languages should be taken into consideration in future investigations.

Finally, the present study was the first to look at the relationship between vocabulary size and knowledge of MWU of different degrees of semantic transparency. It was found that the correlations are equally strong regardless of transparency level, thus vocabulary size can be used as a predictor of more and less transparent MWU knowledge in research and pedagogical practice.
References


Martinez, R., & Schmitt, N. (2012). A Phrasal Expressions List 1,2, 1, (April), 299–320. h


## Appendix 1. Descriptive statistics

### Table 8. Correlations between participants’ age and their scores on PVLT, QOPT and CONTRIX-2

<table>
<thead>
<tr>
<th></th>
<th>PVLT-TOTAL</th>
<th>CONTRIX-TOTAL</th>
<th>QOPT-TOTAL</th>
<th>AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s rho</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.260</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX-TOTAL Correlation Coefficient</td>
<td>.760**</td>
<td>1.000</td>
<td>.685**</td>
<td>.109</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.</td>
<td>.000</td>
<td>.573</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>QOPT-TOTAL Correlation Coefficient</td>
<td>.713**</td>
<td>.685**</td>
<td>1.000</td>
<td>.003</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.</td>
<td>.986</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>AGE Correlation Coefficient</td>
<td>.216</td>
<td>.109</td>
<td>.003</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.260</td>
<td>.573</td>
<td>.986</td>
<td>.</td>
</tr>
<tr>
<td>N</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
Table 9. Normality test for mean scores for the whole CONTRIX-2 test and per sections in two proficiency groups

<table>
<thead>
<tr>
<th>Tests of Normality</th>
<th>Proficiency level</th>
<th>Kolmogorov-Smirnov</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>df</td>
<td>Sig.</td>
</tr>
<tr>
<td>CONTRIX-TOTAL Lower</td>
<td>.160</td>
<td>14</td>
<td>.200*</td>
</tr>
<tr>
<td>CONTRIX-TOTAL Higher</td>
<td>.131</td>
<td>15</td>
<td>.200*</td>
</tr>
<tr>
<td>CONTRIX-TT Lower</td>
<td>.224</td>
<td>14</td>
<td>.054</td>
</tr>
<tr>
<td>CONTRIX-TT Higher</td>
<td>.131</td>
<td>15</td>
<td>.200*</td>
</tr>
<tr>
<td>CONTRIX-ST Lower</td>
<td>.144</td>
<td>14</td>
<td>.200*</td>
</tr>
<tr>
<td>CONTRIX-ST Higher</td>
<td>.244</td>
<td>15</td>
<td>.017</td>
</tr>
<tr>
<td>CONTRIX-NT Lower</td>
<td>.112</td>
<td>14</td>
<td>.200*</td>
</tr>
<tr>
<td>CONTRIX-NT Higher</td>
<td>.127</td>
<td>15</td>
<td>.200*</td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 10. Descriptive statistics for raw mean scores for non-, mono-, bi- and tri-congruent CONTRIX-2 items.

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTRIX-0-cong</td>
<td>29</td>
<td>1</td>
<td>10</td>
<td>5.10</td>
<td>2.677</td>
</tr>
<tr>
<td>CONTRIX-1-cong</td>
<td>29</td>
<td>3</td>
<td>14</td>
<td>7.59</td>
<td>2.970</td>
</tr>
<tr>
<td>CONTRIX-2-cong</td>
<td>29</td>
<td>1</td>
<td>11</td>
<td>5.69</td>
<td>2.466</td>
</tr>
<tr>
<td>CONTRIX-3-cong</td>
<td>29</td>
<td>0</td>
<td>4</td>
<td>1.83</td>
<td>1.104</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 11. Normality test for mean scores for non-, mono, bi and tri-congruent CONTRIX-2 items.

**Tests of Normality**

<table>
<thead>
<tr>
<th>CONTRIX-0-cong, %</th>
<th>Kolmogorov-Smirnov&lt;sup&gt;a&lt;/sup&gt; Statistic</th>
<th>df</th>
<th>Sig.</th>
<th>Shapiro-Wilk Statistic</th>
<th>df</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.108</td>
<td>29</td>
<td>.200*</td>
<td>.949</td>
<td>29</td>
<td>.172</td>
</tr>
<tr>
<td>CONTRIX-1-cong, %</td>
<td>.187</td>
<td>29</td>
<td>.011</td>
<td>.943</td>
<td>29</td>
<td>.120</td>
</tr>
<tr>
<td>CONTRIX-2-cong, %</td>
<td>.162</td>
<td>29</td>
<td>.050</td>
<td>.961</td>
<td>29</td>
<td>.341</td>
</tr>
<tr>
<td>CONTRIX-3-cong, %</td>
<td>.183</td>
<td>29</td>
<td>.014</td>
<td>.909</td>
<td>29</td>
<td>.016</td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.

<sup>a</sup>. Lilliefors Significance Correction

### Table 12. Friedman test ranks for non-, mono-, bi- and tri-congruent CONTRIX-2 items.

**Ranks**

<table>
<thead>
<tr>
<th>CONTRIX %</th>
<th>Mean Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.29</td>
</tr>
<tr>
<td>1</td>
<td>2.24</td>
</tr>
<tr>
<td>2</td>
<td>3.05</td>
</tr>
<tr>
<td>3</td>
<td>2.41</td>
</tr>
</tbody>
</table>

### Table 13. Friedman test statistics for non-, mono-, bi- and tri-congruent CONTRIX-2 items.

**Test Statistics<sup>a</sup>**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>29</td>
</tr>
<tr>
<td>Chi-Square</td>
<td>7.463</td>
</tr>
<tr>
<td>df</td>
<td>3</td>
</tr>
<tr>
<td>Asymp. Sig.</td>
<td>.059</td>
</tr>
</tbody>
</table>

<sup>a</sup>. Friedman Test
Table 14. Normality test for mean scores on PVLT, QOPT and CONTRIX-2 as a whole and per section

Tests of Normality

<table>
<thead>
<tr>
<th></th>
<th>Kolmogorov-Smirnov(^a)</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>df</td>
</tr>
<tr>
<td>PVLT-TOTAL</td>
<td>.139</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX-TOTAL</td>
<td>.092</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX-TT</td>
<td>.177</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX-ST</td>
<td>.197</td>
<td>29</td>
</tr>
<tr>
<td>CONTRIX-NT</td>
<td>.139</td>
<td>29</td>
</tr>
<tr>
<td>OPT-TOTAL</td>
<td>.153</td>
<td>29</td>
</tr>
</tbody>
</table>

* This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 15. The PVLT mean scores for Danish and Russian test-takers

<table>
<thead>
<tr>
<th>PVLT scores</th>
<th>In Revier (2009)</th>
<th>In the present study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10(^{th}) grade high school</td>
<td>12(^{th}) grade high school</td>
</tr>
<tr>
<td>2000</td>
<td>12.3</td>
<td>14.7</td>
</tr>
<tr>
<td>3000</td>
<td>6.9</td>
<td>10.2</td>
</tr>
<tr>
<td>5000</td>
<td>3.7</td>
<td>7.1</td>
</tr>
<tr>
<td>10000</td>
<td>2.8</td>
<td>4.4</td>
</tr>
<tr>
<td>total</td>
<td>25.8</td>
<td>36.4</td>
</tr>
</tbody>
</table>
**Appendix 2. CONTRIX-2 (Revier, 2009)**

Instructions:

Welcome to the CONTRIX-2 test. It measures your knowledge of the whole English collocations. You will be presented with 45 questions. Please, do not use any dictionaries or translating resources. Answer as many questions as possible as shown in the example below:

Example

| 0. Your landlord has the right to decide whether or not you can keep a/an animals hold the creatures tend -- pets. |
|---|---|---|---|

Response

| 0. Your landlord has the right to decide whether or not you can keep a/an animals hold the creatures tend -- pets. |
|---|---|---|---|

Example

| 00. The quickest way to win a friend’s trust is to show that you are able to tell a/an joke take the secret keep -- truth. |
|---|---|---|---|

Response

| 00. The quickest way to win a friend’s trust is to show that you are able to tell a/an joke take the secret keep -- truth. |
|---|---|---|---|
1. If you win the contest, you’ll __________ stating when and where the prize can be picked up.

<table>
<thead>
<tr>
<th>take</th>
<th>a/an</th>
<th>message order report</th>
</tr>
</thead>
<tbody>
<tr>
<td>obtain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>get</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. I understand what you’re saying, but your argument doesn’t seem to __________.

   | stand | a/an | air                 |
   | clear |      | ground              |
   | hold  | --   | water               |

3. Jane had prepared some fun interactive games to __________ at the start of the party.

   | make   | a/an | atmosphere          |
   | break  |      | mood                |
   | clear  | --   | ice                 |

4. Sarah claims that her husband refuses to __________ when it comes to cleaning.

   | do     | a/an | effort               |
   | give   |      | finger               |
   | raise  | --   | share                |

5. We’re moving next Saturday. Would you by __________ have time to drop by and __________.

   | raise  | a/an | part                |
   | play   |      | hand                |
   | give   | --   | effort              |

6. Finally having found the murder weapon, the police are convinced that they will soon __________.

   | catch  | a/an | offender |
   | obtain |      | killer   |
   | detect | --   | attacker |
7. In the evening my mother was always the first to **fix** and to turn on the light.
   | fix | a/an | curtains |
   | draw | the | covers |
   | drag | -- | screens |

8. If you are subjected to discrimination at work, you should **raise** with your union representative.
   | raise | a/an | question |
   | pose | the | topic |
   | mention | -- | matter |

9. Joseph wanted to get into medical school, but he failed to **meet**.
   | meet | a/an | bill |
   | cut | the | grade |
   | make | -- | mark |

10. Declining profits often force companies to **cut**.
    | curb | a/an | jobs |
    | cut | the | occupations |
    | scratch | -- | careers |

11. You should avoid becoming too tired when skiing. Otherwise you could easily lose control or even worse **tear**.
    | tear | a/an | knee |
    | break | the | hand |
    | smash | -- | leg |

12. I can’t be there to direct the meeting next week. Would you be willing to step in for me and **guide**?
    | guide | a/an | matter |
    | lead | the | show |
    | run | -- | scene |

13. I have to take a train to Copenhagen, but I’m afraid I just don’t have enough money to **pay**.
    | pay | a/an | rate |
    | foot | the | fare |
    | meet | -- | price |
14. In our country, we ______ to choose the new government.

- make
- take
- hold

votes
elections
surveys

15. The children will have to ______ to decide who gets to go first.

- pick
- flip
- draw

lots
coins
dice

16. The surest way to ______ is to come in contact with someone who is sneezing or coughing.

- take
- catch
- draw

cold
fever
chill

17. Jill didn’t think it was wise to ______ in the argument between her two best friends.

- take
- join
- make

sides
choices
forces

18. Michael is a safety control technician. His job is to ______ on electrical equipment to ensure that they fulfil safety requirements.

- make
- take
- run

tests
samples
exams

19. Can you ______ so I can find your house?

- write
- draw
- plot

graph
chart
map
20. When I saw my sister’s purse lying on the table, I decided to **take** a/an **money** that she had owed me for ages.

21. When you paint your apartment, you should resist the temptation to **take** a/an **money** by not cleaning the walls thoroughly before you begin.

22. In order to get rid of the dampness in our bathroom, we had to **chop** a/an **hole** in the wall and **cut** the **nick** in order to install a fan.

23. We don’t have time to stand here and wait any longer for the bus. If we are to arrive on time, we’ll have to **run** a/an **taxi**.

24. As a child, Heidi always looked forward to April Fools’ Day when she and her classmates were allowed to **stage** a/an **drills** on each other.

25. The father was delighted when he got to **bear** a/an **baby** just after it was born.
26. She sat completely alone for hours. There was only the tick of the clock to [stop] [a/an] [stillness].

27. There’s no direct connection between Moscow and Chelyabinsk. You’ll have to [change] [a/an] [trains] in Ekaterinburg.

28. She suspected that she’d be made to [pay] [a/an] [can] for her supervisor’s mistakes.

29. What makes a hurricane extremely dangerous is that it can suddenly [turn] [a/an] [path] and hit an entirely different place than expected.

30. When they are asked to help the victims of disasters, many people are willing to [grant] [a/an] [capital].

31. In a large city like Copenhagen, a small business can easily [replace] [a/an] [hands] two to three times a year.
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>32. In some countries, only the police are authorized to bring a/an shotgun to carry the rifle hold -- gun.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33. If you continue to show up late for work, you’re going to take a/an sack get the dust hit -- road.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. At French universities, all English students are required to take a/an course perform the lesson get -- study.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. Some medicines are known to carry a/an risk contain the chance bring -- worry.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. Despite being married, Tom continues to play a/an contest run the game lead -- field.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37. Military officers have the authority to take a/an rules give the requests make -- orders.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise</td>
<td>Sentence</td>
<td>Blank(s)</td>
<td>Option(s)</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>-----------</td>
</tr>
<tr>
<td>38.</td>
<td>The next time you decide to [ ] remember to strain off the leaves before you serve it.</td>
<td>to strain off the leaves</td>
<td>boil, cook, make</td>
</tr>
<tr>
<td>39.</td>
<td>Today, in gym class, Steve and I were asked to [ ] to find out which of us was faster.</td>
<td>to find out which of us was faster</td>
<td>hold, run, stage</td>
</tr>
<tr>
<td>40.</td>
<td>If the teenagers next door don’t turn down the music, I’m going to [ ] to the police.</td>
<td>to the police</td>
<td>pay, lay, make</td>
</tr>
<tr>
<td>41.</td>
<td>Jill encourages her children to [ ] as a way to learn numbers.</td>
<td>as a way to learn numbers</td>
<td>play, do, deal</td>
</tr>
<tr>
<td>42.</td>
<td>In the summer, our children like to [ ] to their grandparents who live on a farm.</td>
<td>to their grandparents</td>
<td>pay, go, spend</td>
</tr>
<tr>
<td>43.</td>
<td>If you continue to neglect your homework, you’ll [ ] when exam time comes.</td>
<td>when exam time comes</td>
<td>pay, bear, carry</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>44. When it’s hot and dry, it doesn’t take much for a straw roof to draw a/an flame.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45. Car manufacturers plan to hoist a/an fees to offset the rising cost of labour and raw materials.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3. PVLT, (Laufer & Nation, 1999)

This test measures English productive vocabulary size. Please answer 72 questions below based on your own knowledge only. Do not consult any dictionaries or translating resources. For each sentence put the correct word in the box as shown in the example #0:

Please note that this survey will be best displayed on a laptop or desktop computer. Some features may be less compatible for use on a mobile device.

Example

0. The plants receive water through their ro____________

Response
0. The plants receive water through their ro____________

Test A 2000 level

I’m glad we had this opp____________ to talk

There are a doz__________ eggs in the basket.

Every working person must pay income t__________ .

The pirates buried the trea__________ on a desert island.
Her beauty and ch__________ had a powerful effect on men.

La__________ of rain led to a shortage of water in the city.

He takes cr__________ and sugar in his coffee.

The rich man died and left all his we__________ to his son.

Pup__________ must hand in their papers by the end of the week.

This sweater is too tight. It needs to be stret__________

Ann intro _________ her boyfriend to her mother.

Teenagers often adm__________ and worship pop singers.

If you blow up that balloon any more it will bu__________
In order to be accepted into the university, he had to impr__________ his grades.

The telegram was deli__________ two hours after it had been sent.

The differences were so sl__________ that they went unnoticed.

The dress you’re wearing is lov__________

He wasn’t very popu__________ when he was a teenager, but he has many friends now.

Test A 2000-3000 levels

He has a successful car__________ as a lawyer.

The thieves threw ac__________ in his face and made him blind.

To improve the country’s economy, the government decided on economic ref__________

She wore a beautiful green go__________ to the ball.
The government tried to protect the country’s industry by reducing the imp_________ of cheap goods.

The children’s games were amusing at first, but finally got on the parents’ ner________.

The lawyer gave some wise coun_________ to his client.

Many people in England mow the la_________ of their houses on Sunday morning.

The farmer sells the eggs that his he_________ lays.

Sudden noises at night sca_________ me a lot.

France was proc_________ a republic in the 18th century.

Many people are inj_________ in road accidents every year.

Suddenly he was thru_________ into the dark.
Heperc__________a light at the end of the tunnel.

Children are not independent. They are att__________to their parents.

She showed off her sle__________figure in a long narrow dress.

She has been changing partners often because she cannot have a sta__________relationship with one person.

You must wear a bathing suit on a public beach. You’re not allowed to bath na__________ .

Test A 3000-5000 levels

Soldiers usually swear an oa__________of loyalty to their country.

The voter placed the ball__________in the box.

They keep their valuables in a vau__________at the bank.
A bird perched at the window led__________.

The kitten is playing with a ball of ya__________.

The thieves have forced an ent__________ into the building.

The small hill was really a burial mou__________.

We decided to celebrate New Year’s E__________ together.

The soldier was asked to choose between infantry and cav__________.

This is a complex problem that is difficult to compr__________.

The angry crowd sho__________ the prisoner as he was leaving the court.

Don’t pay attention to this rude remark. Just ig__________ it.
The management held a secret meeting. The issues discussed were not disc________ to the workers.

We could hear the sergeant bel________ commands to the troops.

The boss got angry with the secretary and it took a lot of tact to soo__________ him.

We do not have adeq _________ information to make the decisions.

She is not a child, but a mat__________ woman. She can make her own decisions.

The prisoner was put in soli__________ confinement.

**Test A 5000 = 10000 levels**

The baby is wet. Her dia__________ needs changing.

The prisoner was released on par__________.

Second year university students in the US are called soph__________.
Her favourite flowers were or ________.

The insect causes damage to plants by its toxic sec__________.

The evacu__________ of the building saved many lives.

For many people, wealth is a prospect of unimaginable felic__________.

She found herself in a pred__________ without any hope for a solution.

The deac__________ helped with the care of the poor of the parish.

The hurricane whi__________ along the coast.

Some coal was still smol__________ among the ashes.

The dead bodies were mutil__________ beyond recognition.
She was sitting on a balcony and bas__________ in the sun.

For years waves of invaders pill__________ towns along the coast.

The rescue attempt could not proceed quickly it was imp__________ by bad weather

I wouldn’t hire him. He is unmotivated and indo__________ .

Computers have made typewriters old-fashioned and obs__________ .

Watch out for his wil__________ tricks.
Appendix 4. QOPT (Allan, 1992)

This test measures level of proficiency in English. Please answer 60 questions below based on your own knowledge only. Do not consult any dictionaries or translating resources.

Please note that this survey will be best displayed on a laptop or desktop computer. Some features may be less compatible for use on a mobile device.

For the questions 1-5 choose one answer only.

#1

**Please leave your room key at Reception.**

Where can you see this notice? Mark one answer.

- In a shop
- In a hotel
- In a taxi

#2

**Foreign money changed here**

Where can you see this notice? Mark one answer.

- In a library
- In a bank
- In a police station
#3

**AFTERNOON SHOW BEGINS AT 2PM**

Where can you see this notice? Mark one answer.

- [ ] Outside a theatre
- [ ] Outside a supermarket
- [ ] Outside a restaurant

#4

**CLOSED FOR HOLIDAYS**

Lessons start again on the 8th January

Where can you see this notice? Mark one answer.

- [ ] At a travel agent's
- [ ] At a music school
- [ ] At a restaurant

#5

**Price per night:**

- £10 a tent
- £5 a person
Where can you see this notice? Mark one answer.

- At a cinema
- At a hotel
- At a camp-site

In this section you must choose the word which best fits each space in the text below. Choose one option in the dropdown lists 6-10.

Scotland

Scotland is the north part of the island of Great Britain. The Atlantic Ocean is on the west and the North Sea on the east. Some people (6) ................. Scotland speak a different language called Gaelic.

There are (7) ................. five million people in Scotland, and Edinburgh is (8) ................. most famous city.

Scotland has many mountains; the highest one is called ‘Ben Nevis’. In the south of Scotland, there are a lot of sheep. A long time ago, there (9) ................. many forests, but now there are only a (10) ................. .

Scotland is only a small country, but it is quite beautiful.

#6

- on
- in
- at

#7

- about
- between
- among
In this section you must choose the word which best fits each space in the text below. Choose one option in the dropdown lists 11-15.

Alice Guy Blaché was the first female film director. She first became involved in cinema whilst working for the Gaumont Film Company in the late 1890s. This was a period of great change in the cinema and Alice was the first to use many new inventions, **(11)** ................. sound and colour.

In 1907 Alice **(12)** ................. to New York where she started her own film company. She was **(13)** ................. successful, but, when Hollywood became the centre of the film world, the best days of the independent New York film companies were **(14)** ................. . When Alice died in 1968, hardly anybody **(15)** ................. her name.
#11
- bringing
- including
- containing
- supporting

#12
- moved
- ran
- entered
- transported

#13
- next
- once
- immediately
- recently

#14
- after
- down
- behind
- over

#15
- remembered
- realised
- reminded
- repeated
In this section you must choose the word which best fits each space in the text below. Choose one option in the dropdown lists 16-20.

---

**UFOs – do they exist?**

UFO is short for ‘unidentified flying object’. UFOs are popularly known as flying saucers, *(16)************* that is often the *(17)************* they are reported to be. The *(18)************* "flying saucers" were seen in 1947 by an American pilot, but experts who studied his claim decided it had been a trick of the light.

Even people experienced at watching the sky, *(19)************* as pilots, report seeing UFOs. In 1978 a pilot reported a collection of UFOs off the coast of New Zealand. A television *(20)************* went up with the pilot and filmed the UFOs. Scientists studying this phenomenon later discovered that in this case they were simply lights on boats out fishing.

---

**#16**

- because
- therefore
- although
- so

**#17**

- look
- shape
- size
- type
In this section you must choose the word or phrase which best completes each sentence. Choose one option in the dropdown lists 21-40.

#21

The teacher encouraged her students ................. to an English pen-friend.

- should write
- write
- wrote
- to write
#22
They spent a lot of time ................ at the pictures in the museum.

- looking
- for looking
- to look
- to looking

#23
Shirley enjoys science lessons, but all her experiments seem to ................ wrong.

- turn
- come
- end
- go

#24
....................... from Michael, all the group arrived on time.

- Except
- Other
- Besides
- Apart

#25
She .................... her neighbour’s children for the broken window.

- accused
- complained
- blamed
- denied
#26

As I had missed the history lesson, my friend went .................. the homework with me.

- by
- after
- over
- on

#27

Whether she’s a good actress or not is a .................. of opinion.

- matter
- subject
- point
- case

#28

The decorated roof of the ancient palace was ................ up by four thin columns.

- built
- carried
- held
- supported

#29

Would it .................. you if we came on Thursday?

- agree
- suit
- like
- fit
This form .................. be handed in until the end of the week.

- doesn't need
- doesn't have
- needn't
- hasn't got

If you make a mistake when you are writing, just .................. it out with your pen.

- cross
- clear
- do
- wipe

Although our opinions on many things .................. , we’re good friends.

- differ
- oppose
- disagree
- divide

This product must be eaten .................. two days of purchase.

- by
- before
- within
- under
The newspaper report contained important information.

- many
- another
- an
- a lot of

Have you considered moving to London?

- move
- to move
- to be moving
- moving

It can be a good idea for people who lead an active life to increase their intake of vitamins.

- upturn
- input
- upkeep
- intake

I thought there was a piece of jealousy in his reaction to my good fortune.

- piece
- part
- shadow
- touch
Why didn’t you .................. that you were feeling ill?

○ advise
○ mention
○ remark
○ tell

James was not sure exactly where his best interests ..................... .

○ stood
○ rested
○ lay
○ centred

He’s still getting .................. the shock of losing his job.

○ across
○ by
○ over
○ through

In this section you must choose the word or phrase which best fits each space in the text below. Choose one option in the dropdown lists 41-45.
The tallest buildings - SKYSCRAPERS

Nowadays, skyscrapers can be found in most major cities of the world. A building which was many (41) ………………. high was first called a skyscraper in the United States at the end of the 19th century, and New York has perhaps the (42) ………………. skyscraper of them all, the Empire State Building. The (43) ………………. beneath the streets of New York is rock, (44) ………………. enough to take the heaviest load without sinking, and is therefore well-suited to bearing the (45) ………………. of tall buildings.

#41

○ stages
○ steps
○ storeys
○ levels

#42

○ first-rate
○ top-class
○ well-built
○ best-known

#43

○ dirt
○ field
○ ground
○ soil

In this section you must choose the word or phrase which best fits each space in the text below. Choose one option in the dropdown lists 46-50.
SCRABBLE

Scrabble is the world’s most popular word game. For its origins, we have to go back to the 1930s in the USA, when Alfred Butts, an architect, found himself out of (46) ................. . He decided that there was a (47) ................. for a board game based on words and (48) ................. to design one. Eventually he made a (49) ................. from it, in spite of the fact that his original (50) ................. was only three cents a game.

#46
- earning
- work
- income
- job

#47
- market
- purchase
- commerce
- sale

#48
- took up
- set out
- made for
- got round
In this section you must choose the word or phrase which best completes each sentence. Choose one option in the dropdown lists 51-60.

Roger’s manager ______________ to make him stay late if he hadn’t finished the work.

insisted  warned  threatened  announced

By the time he has finished his week’s work, John has hardly ______________ energy left for the weekend.

any  much  no  same
As the game ................. to a close, disappointed spectators started to leave.

- led
- neared
- approached
- drew

I don’t remember ................. the front door when I left home this morning.

- to lock
- locking
- locked
- to have locked

I .................. to other people borrowing my books: they always forget to return them.

- disagree
- avoid
- dislike
- object

Andrew’s attempts to get into the swimming team have not .................. with much success.

- associated
- concluded
- joined
- met
Although Harry had obviously read the newspaper article carefully, he didn’t seem to have the main point.

- grasped
- clutched
- clasped
- gripped

A lot of the views put forward in the documentary were open to enquiry.

- enquiry
- query
- question
- wonder

The new college deals for the needs of students with a variety of learning backgrounds.

- deals
- supplies
- furnishes
- caters

I find the times of English meals very strange – I’m not used to having dinner at 6pm.

- to have
- to having
- having
- have

<table>
<thead>
<tr>
<th>№</th>
<th>English</th>
<th>SC</th>
<th>Russian</th>
<th>Roman alphabet equivalent</th>
<th>Verb</th>
<th>Art.</th>
<th>Noun</th>
<th>Verb CS</th>
<th>Art. CS</th>
<th>Noun CS</th>
<th>Total CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Get a message</td>
<td>TT</td>
<td>Получить сообщение</td>
<td>Poluchit’ soobsche nie</td>
<td>Get</td>
<td></td>
<td>Message</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Hold water</td>
<td>NT</td>
<td>Выдерживать критику</td>
<td>Viderzhiva t’ kritiku</td>
<td>bear, hold</td>
<td></td>
<td>criticism</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Break the ice</td>
<td>NT</td>
<td>Растопить лед</td>
<td>Rastopit’ liod</td>
<td>melt</td>
<td></td>
<td>ice</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Raise a finger</td>
<td>NT</td>
<td>Пошевелить пальцем</td>
<td>Poshevelit’ ‘pal’tsem</td>
<td>move</td>
<td></td>
<td>finger</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Give a hand</td>
<td>NT</td>
<td>Протянуть руку помощи</td>
<td>Protianuti ruku pomoschi</td>
<td>stretch out</td>
<td></td>
<td>Hand of help</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Catch the killer</td>
<td>TT</td>
<td>Поймать убийцу</td>
<td>Poimat’ ubijtsu</td>
<td>catch</td>
<td></td>
<td>killer</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>Draw the curtains</td>
<td>ST</td>
<td>Задернуть шторы</td>
<td>Zadernut’ shtori</td>
<td>Pull, draw</td>
<td></td>
<td>curtains</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>Raise the matter</td>
<td>ST</td>
<td>Поднять вопрос</td>
<td>Podniat’ vopros</td>
<td>raise</td>
<td></td>
<td>question</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Make the grade</td>
<td>NT</td>
<td>Добиться успеха</td>
<td>Dobit’sia uspeha</td>
<td>Get, achieve</td>
<td></td>
<td>success</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>Cut jobs</td>
<td>ST</td>
<td>Сокращать рабочие места</td>
<td>Sokrashe t’ rabochi’i e mesta</td>
<td>reduce</td>
<td></td>
<td>work-places</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>Break a leg</td>
<td>TT</td>
<td>Сломать ногу</td>
<td>Slomat’ nogu</td>
<td>Break</td>
<td></td>
<td>leg</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>Run the show</td>
<td>NT</td>
<td>Disparity of translations</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>Pay the fare</td>
<td>TT</td>
<td>Platit’ za pro’iezd</td>
<td>Platit’ za pro’iezd</td>
<td>Pay for</td>
<td></td>
<td>trip</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>Hold elections</td>
<td>ST</td>
<td>Проводить выборы</td>
<td>Provodit’ vibori</td>
<td>Carry out, conduct</td>
<td></td>
<td>elections</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>Draw lots</td>
<td>NT</td>
<td>Тянуть жребий</td>
<td>Tianut’ zhrebi</td>
<td>pull</td>
<td></td>
<td>lot</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>Catch a cold</td>
<td>NT</td>
<td>Подхватить простуду</td>
<td>Podхватит’ простуду</td>
<td>hook up</td>
<td></td>
<td>cold</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>Take sides</td>
<td>NT</td>
<td>Занимать сторону</td>
<td>Zanimati storonu</td>
<td>occupy</td>
<td></td>
<td>side</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

6 SC stands for semantic category, namely transparent (TT), semi-transparent (ST) or non-transparent (NT).
7 CS stands for congruence status.
<table>
<thead>
<tr>
<th>№</th>
<th>English</th>
<th>SC</th>
<th>Russian</th>
<th>Roman alphabet equivalent</th>
<th>Verb</th>
<th>Art.</th>
<th>Noun</th>
<th>Verb CS</th>
<th>Art. CS</th>
<th>Noun CS</th>
<th>Total CS</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>Run tests</td>
<td>ST</td>
<td>Тестируетъ</td>
<td>Testirovat’</td>
<td>test</td>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>19</td>
<td>Draw a map</td>
<td>TT</td>
<td>Рисовать карту</td>
<td>Risovat’ cartu</td>
<td>draw</td>
<td>map</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>Take the money</td>
<td>TT</td>
<td>Взять деньги</td>
<td>Vziat’ denigi</td>
<td>take</td>
<td>money</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>Cut corners</td>
<td>NT</td>
<td>Срезать углы</td>
<td>Srezat’ ugli</td>
<td>Cut off</td>
<td>Angle, corner</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Cut a hole</td>
<td>TT</td>
<td>Делать отверстие</td>
<td>Delat’ otverst’ie</td>
<td>Make, do</td>
<td>hole</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Get a taxi</td>
<td>ST</td>
<td>Взять такси</td>
<td>Vziat’ taksi</td>
<td>take</td>
<td>taxi</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Play tricks</td>
<td>ST</td>
<td>Разыгрывать</td>
<td>Razigrivat’</td>
<td>Play</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Hold the baby</td>
<td>TT</td>
<td>Держать ребенка</td>
<td>Derzhat’ rebionka</td>
<td>hold</td>
<td>baby</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Break the silence</td>
<td>ST</td>
<td>Нарушать тишину</td>
<td>Narushat’ tishinu</td>
<td>violate</td>
<td>silence</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Change trains</td>
<td>ST</td>
<td>Disparity of translations</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Carry a gun</td>
<td>NT</td>
<td>Носить оружие</td>
<td>Nosit’ oruzhi’ie</td>
<td>carry</td>
<td>Weapon, arm</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Change direction</td>
<td>TT</td>
<td>Менять направлени е</td>
<td>Meniat’ napravleni ‘ie</td>
<td>change</td>
<td>directio n</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Give money</td>
<td>TT</td>
<td>Дать деньги</td>
<td>Dat’ den’gi</td>
<td>give</td>
<td>money</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Change hands</td>
<td>NT</td>
<td>Переходить из рук в руки</td>
<td>Perehodit’ is ruk v ruki</td>
<td>pass</td>
<td>Fro m</td>
<td>hands to hands</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>32</td>
<td>Carry the can</td>
<td>TT</td>
<td>Платить за ошибки</td>
<td>Plaitit’ za oshibki</td>
<td>pay</td>
<td>for</td>
<td>Mistakes, errors</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>33</td>
<td>Get the sack</td>
<td>NT</td>
<td>Disparity of translations</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Take a course</td>
<td>ST</td>
<td>Пройти курс</td>
<td>Proiti kurs</td>
<td>pass</td>
<td>course</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Carry the risk</td>
<td>ST</td>
<td>Нести риски</td>
<td>Nesti riski</td>
<td>carry</td>
<td>risks</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Play the field</td>
<td>NT</td>
<td>Disparity of translations</td>
<td>-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Give orders</td>
<td>ST</td>
<td>Отдавать приказы</td>
<td>Otdavat’ prikazi</td>
<td>Give away</td>
<td>orders</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Make tea</td>
<td>TT</td>
<td>Делать чай</td>
<td>Delat’ chai</td>
<td>Make, do</td>
<td>tea</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>№</td>
<td>English</td>
<td>SC</td>
<td>Russian</td>
<td>Roman alphabet equivalent</td>
<td>Verb</td>
<td>Art.</td>
<td>Noun</td>
<td>Verb CS</td>
<td>Art. CS</td>
<td>Noun CS</td>
<td>Total CS</td>
</tr>
<tr>
<td>----</td>
<td>----------------</td>
<td>-----</td>
<td>---------</td>
<td>---------------------------</td>
<td>------</td>
<td>------</td>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>39</td>
<td>Run a race</td>
<td>TT</td>
<td>Disparity here but no RUN offered</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>Make a complaint</td>
<td>ST</td>
<td>Подать жалобу</td>
<td>Podat’ zhalobu</td>
<td>hand in</td>
<td>complain</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>41</td>
<td>Play cards</td>
<td>TT</td>
<td>Играть в карты</td>
<td>Igrat’ v karti</td>
<td>play</td>
<td>in</td>
<td>cards</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>42</td>
<td>Pay a visit</td>
<td>ST</td>
<td>Навещать</td>
<td>Naveschat’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>43</td>
<td>Pay the penalty</td>
<td>NT</td>
<td>Платить по счетам</td>
<td>Platit’ po schetam</td>
<td>pay</td>
<td>by</td>
<td>bills</td>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>Catch fire</td>
<td>ST</td>
<td>Воспламениться</td>
<td>Vosplame niat’sia</td>
<td>Flame up</td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>Raise prices</td>
<td>TT</td>
<td>Поднимать цены</td>
<td>Podnimat’ tseni</td>
<td>raise</td>
<td></td>
<td>prices</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix 6. Background questionnaire

This questionnaire is designed to collect general information about the participants and their experience with learning foreign languages. Please answer all the questions below.

Your age

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
</tr>
</tbody>
</table>

Your gender

- Male
- Female

Your first language (the one you speak from birth)

- Russian
- Other

Do you speak any other language as your first language?

- Yes
- No

Display This Question:

If Do you speak any other language as your first language? = Yes

What is the other language you speak from birth?

________________________________________________________________
How long have you been learning/using English (in years)? If you have been learning/using English for longer than 60 years, put 60.

![Bar Chart with options 0, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60]

What is your level of proficiency in English?

- Elementary
- Intermediate
- Advanced
- Native-like

Have you ever lived abroad?

- Yes
- No

Display This Question:
If Have you ever lived abroad? = Yes

How long have you lived abroad (in months)? If you spent abroad over than 120 months, put 120.

![Bar Chart with options 0, 12, 24, 36, 48, 60, 72, 84, 96, 108, 120]
Display This Question:
If Have you ever lived abroad? = Yes

Have you ever lived in an English speaking country?

○ Yes
○ No

Display This Question:
If Have you ever lived in an English speaking country? = Yes

How long have you lived in an English speaking country (in months)? If you spent there over 120 months, put 120.

0 12 24 36 48 60 72 84 96 108 120

Do you speak any other foreign language⁸?

○ Yes
○ No

———

⁸ As mentioned in the Section 3.1 the participants who had answer this question positively were later asked to evaluate their proficiency level for each foreign language they speak by means of CEFR descriptors.