



Exploring the Mental Lexicon of the Multilingual: Vocabulary Size, Cognate Recognition and Lexical Access in the L1, L2 and L3

Csaba Z Szabo ^{*,a}

^a *The Open University, Milton Keynes, UK*

Abstract

Recent empirical findings in the field of Multilingualism have shown that the mental lexicon of a language learner does not consist of separate entities, but rather of an intertwined system where languages can interact with each other (e.g. Cenoz, 2013; Szubko-Sitarek, 2015). Accordingly, multilingual language learners have been considered differently to second language learners in a growing number of studies, however studies on the variation in learners' vocabulary size both in the L2 and L3 and the effect of cognates on the target languages have been relatively scarce. This paper, therefore, investigates the impact of prior lexical knowledge on additional language learning in the case of Hungarian native speakers, who use Romanian (a Romance language) as a second language (L2) and learn English as an L3. The study employs an adapted version of the widely used Vocabulary Size Test (Nation & Beglar, 2007), the Romanian Vocabulary Size Test (based on the Romanian Frequency List; Szabo, 2015) and a Hungarian test (based on a Hungarian frequency list; Varadi, 2002) in order to measure vocabulary sizes, cognate knowledge and response times in these languages. The findings, complemented by a self-rating language background questionnaire, indicate a strong link between Romanian and English lexical proficiency.

© 2016 EJAL & the Authors. Published by *Eurasian Journal of Applied Linguistics (EJAL)*. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Vocabulary size; cognates; lexical access; response time, multilingualism; mental lexicon; VST; L1, L2 and L3

1. Introduction

The ubiquity and significance of vocabulary in language acquisition has led to a growing number of studies exploring the relationship between lexical knowledge and language competence, especially in English (e.g. Laufer and Goldstein, 2004; Milton, 2009). The reason for this is that although language ability might comprise various components (Bachmann and Palmer, 1996), a large cohort of researchers now agree that lexical knowledge is a core element of language performance and therefore, 'lexical knowledge is conceptually the most important component of variability in language ability' (Treffers-Daller, 2011, p. 149). Milton (2013), in his review of several studies, extrapolates that there is a strong link between vocabulary size and performance in a foreign language, including all the four key skills of reading, writing, speaking and listening. This sizable relationship between lexical dexterity and language performance has been addressed in the case of native speakers by Treffers-

* Csaba Z Szabo. Tel.: +447552624435
E-mail address: csaba.szabo@open.ac.uk
<http://dx.doi.org/>.....

Daller and Milton (2013). They contend that '[s]tudents with larger vocabularies tend to score higher in their assignments and exams and to obtain higher degree classifications than those with smaller vocabularies' (p. 166). Measuring vocabulary size in English, thus, became the central foci in the past decades for not just researchers, but teachers as well, and this led to the development of a number of vocabulary size tests, albeit mainly receptive.

Based on the number of studies and validation data, the most notable examples of such receptive tests are the Vocabulary Levels Test (VLT, Nation, 1983; Schmitt, Schmitt and Clapham, 2001), the Vocabulary Size Test (VST, Nation, 2006; Nation & Beglar, 2007) and the so called yes/no tests, such as Meara and Milton's X-Lex (2003). These ostensibly uncomplicated tests have been employed in SLA to explore learners' vocabulary knowledge to allow for, *inter alia*, monitoring of progress and allocating students to different proficiency groups in the case of placement tests. The problem with these tests, however, is that although they measure lexical knowledge to some extent, the results cannot necessarily be compared across cultures, learning contexts and individual language backgrounds, which in turn makes it difficult to standardize these tests. As Cobb rightly argues, 'vocabulary tests can focus on either language or the learner' (2000, p. 300) and in most research to date, the focus has been mainly on the language, disregarding other factors that can influence lexical knowledge. Recently, it has been argued, however, that due to the effect of prior lexical knowledge on additional language learning, the conceptualization and operationalization of a multifaceted lexical knowledge posits several challenges and perhaps this is the reason why measuring vocabulary size in multiple languages and exploring the relationship and interconnection between different languages is still an understudied area.

There is a growing amount of evidence that the difference between monolinguals, bilinguals and multilinguals can be fundamental from a language acquisition perspective. For instance Bialystok, Craik and Luk (2008) found that bilinguals from a variety of language backgrounds can even outperform their monolingual counterparts on letter fluency and word naming tasks due to their executive control. Barac and Bialystok (2011) tested different bilingual groups and a monolingual group on verbal and non-verbal tasks. Their conclusion indicates that on the executive control tasks bilinguals outperformed monolinguals, regardless of language background, however on the receptive and productive vocabulary tasks only Spanish-English bilinguals showed an advantage due to the similarity between the languages. Employing a different research paradigm, Molnar (2008) found that Hungarian-Romanian bilinguals outperformed the Hungarian monolingual group on the VLT, however the Romanian monolingual group obtained significantly higher scores than both other groups due to their high proficiency in the typologically closer language. This illustrates Treffers-Daller's (2011) claim about the importance of obtaining detailed information of participants' lexicon, and as this paper attempts to demonstrate, since background languages have a fundamental impact on the mental lexicon of multilingual language learners, prior to establishing the relationship

between vocabulary size and other competences, the understanding of vocabulary size in the case of multiple languages and the interaction between all these languages is highly pressing.

The present paper therefore sets out to explore vocabulary knowledge, with a focus on the learner and their language background. In particular, it investigates the connection between vocabulary sizes in different background languages in the case of Hungarian (L1) native speakers living in Romania (L2) and studying English (L3) at university level. Similarly to Molnar (2010) and Szabo (forthcoming) a widely used receptive, meaning recognition vocabulary test was employed to measure lexical knowledge at different levels and cognate recognition in not just the target language (L3), but the typologically closer language (L2) as well. By contrast to these two studies however, the present study opted for the VST, as its format also allowed for an online measure that records response time. Furthermore, there are only a handful of studies that explore lexical access or reaction time on yes/no vocabulary size tests (e.g. Miralpeix and Meara, 2014; Pellicer-Sánchez and Schmitt, 2012) or on multiple-choice tests (e.g. Harrington, 2006; Laufer and Nation, 2001; Tanabe, 2016), thus it became of interest whether vocabulary size is also a good indication of lexical access at different proficiency levels. Another point of interest was related to the cognate facilitation effect. While García (1991) found that learners could only take advantage of their cognate knowledge after explicit instruction on cognates, Hall (2002), Molnar (2010), and Szabo (forthcoming) found that learners might take advantage of lexical similarities between languages intuitively. Since we can corroborate from several studies that cognates have a facilitative effect as they are processed faster and more easily than control words (e.g. Cobb, 2000; Helms-Park and Perhan, 2016), and this concurs with Elgort's assertion that cognates are recognized, translated and acquired faster and more accurately by bilinguals' (2013, p. 255), accounting for them in lexical tests seems essential.

2. Literature Review

2.1. Lexical knowledge and multilingualism

Research in SLA has always been influenced by theories and findings in first (or child) language acquisition. However, recent proclivities of language research demonstrate that in many cases, especially in bilingual communities and immersion schools, a multilingual rather than a monolingual perspective is of paramount importance for an in-depth understanding of language acquisition. Many researchers have argued that the influence of and transfer from the L1 is non-negotiable, in particular if the L1 and the target language share similarities (e.g. Ard and Homburg, 1983; Gass and Selinker, 2001). Recently, it became apparent however that this one-dimensional approach may mask some of the quintessential interactions between other known languages. The reason for this is that while transfer in L2 acquisition comes from the L1, in L3 acquisition a so called interlanguage transfer may take place (De Angelis and Selinker, 2001), meaning that the target language can be influenced

by all previously known languages. This interlanguage or cross-linguistic transfer is typically from the language that is typologically more proximate to the target language.

Other evidence also suggests that cross linguistic transfer to English occurred from Spanish, rather than Basque as that can be regarded as the typologically closer language, regardless if it was an L1 or L2 (Cenoz, 2001). In the Canadian context, Chen et al. (2012) compared Spanish and Chinese speakers and corroborated that due to the cognate knowledge that Spanish speakers could rely on, they outperformed their Chinese counterparts.

Using a stepwise multiple regression analysis, Willis and Ohashi (2012) tested 69 Japanese students on the first seven levels of the VST and controlled for cognateness, word length and frequency to find the most parsimonious model. They, found that cognates are easier to learn and retain than non-cognates, and cognates of similar frequency and phoneme length have a higher item facility ratio (proportion of hits to items on the test), which is in contrast with Milton and Daller (2007), who only found a frequency effect. Furthermore, Willis and Ohashi's (2012) results also confirm previous studies showing that the more frequent and shorter words are more easily recognized. According to their report, cognateness accounted for 54% variation in the model, while frequency and word length in phonemes explained 34% and 11% respectively. In a recent replication of this study, Reynolds et al. (2015) tested 330 Taiwanese English majors on all levels of the VST and found that polysemy contributed the most to the model (60%), followed closely by frequency (59%) and noun as variables (26%). The reason for the lack of a significant cognate facilitation effect is due to the fact that the amount of cognates in the study was only 2.4% out of the 140 items in comparison to the 29% out of 70 reported by Willis and Ohashi (2012). The results from these two studies confirm that given the proportion and distribution of cognates in the test, word difficulty can largely influence the results on the test, based on prior lexical knowledge of the learners.

2.2. Lexical knowledge and explicit instruction on cognates

As the above shows, in the case in which one of the known languages is related to the target language and thus the number of cognates on the test increases, the results can be fundamentally different. The question here is whether this result can also be influenced by a brief explicit instruction on cognates. Molnar (2010) used the VLT (53% Romanian cognates) to measure L3 vocabulary knowledge in the case of Hungarian - Romanian bilinguals. Her study explored whether raising awareness of structural similarities between Romanian and English cognates has an influence on the vocabulary scores in English. Furthermore, she used teacher ratings of students' Romanian knowledge to compare English vocabulary scores and Romanian knowledge due to the lack of a Romanian lexical test. The findings, nevertheless, suggest that there is a sizable correlation between Romanian proficiency and English lexical knowledge. Overall scores indicate that there is a statistically significant correlation

between the L2 and L3 (Pearson's $r = .420$ and $r = .525$ for cognate scores, $p < .001$; sig. 2-tailed). This can be interpreted as students with a high proficiency in their L2 will be more prone to achieve higher scores in their L3, which supports previous research emphasizing that all previously known languages will interact with the target language and this interaction will be most noticeable in the area of vocabulary (De Angelis, 2007). This interaction between languages, however, seems to be automatic as it appears from Molnar (2010) that the instruction on cognates had no effect on L3 vocabulary scores.

In a more controlled study, Szabo (forthcoming) using the same VLT and the Romanian Vocabulary Levels Test (Szabo, 2015), also found a close relationship between L2 Romanian and L3 English vocabulary scores. Vocabulary scores in the two languages showed robust correlation between overall lexical proficiency $r = .792$ and cognate knowledge $r = .861$ (Pearson's correlation, $p < .001$). Moreover, similarly to Molnar's study (2010), this study also confirms that despite the instruction on cognates that the study group received, the difference in participants' scores on cognates was not significant ($F(1, 38) = .033$, $p = .86$, *eta squared* = .001). The author concludes that students at higher proficiency levels use their cognate knowledge intrinsically and no explicit instruction is required.

The above studies support the connectionist views, which, similarly to Meara (1996), consider that the mental lexicon is comparable to a massive network of simple processing units, and within this network new words are acquired through similarity detection and exploitation or pattern-matching and assimilation. Hall (2002), based on the Parasitic Model of vocabulary development that relies upon this network building, argues that due to the form-meaning connections that are either lexical or conceptual, cognate recognition in the case of bi- or multilinguals is an automatic process. Conversely, García (1991) asserts that cognates are only recognized after explicit instruction. Research conducted by Helms-Park and Perhan (2016) also indicates that explicit instruction on cognates in the case of Ukrainian and English facilitated vocabulary learning and also allowed attaining higher vocabulary scores. The reason for the discrepancy between the two cohorts may be found in research by Nagy et al. (1993), who conclude that cognate recognition might be age, and thus by extension, also proficiency dependent.

2.3. *Lexical knowledge and lexical access*

Empirical studies that systematically investigate lexical access or reaction time (RT) on vocabulary size tests are rather limited (e.g. Harrington, 2006; Laufer and Nation, 2001; Miralpeix and Meara, 2014; Pellicer-Sánchez and Schmitt, 2012; Tanabe, 2016). These studies extrapolated, however, two important factors regarding English, one being the relationship between speed of lexical access and increasing vocabulary size (*size effect*) and the closely related *frequency effect*, according to which lower frequency items will require longer response times due to the increased difficulty of the word. Harrington (2006), for example, found that speed of lexical

access systematically decreased with increasing proficiency (size effect). Similarly, Pellicer-Sánchez and Schmitt (2012), also using the yes/no format, found that participants not only gave faster responses to higher frequency words (frequency effect), but also to words they knew better. Also employing the yes/no format, Miralpeix and Meara (2014) found that the group with the larger vocabulary score provided significantly faster responses (size effect). Their findings did not confer evidence for the frequency effect however, which lead the authors to suggest that there might be a vocabulary size threshold beyond which accessibility becomes a determinant factor. In an earlier study, Laufer and Nation (2001) measured lexical-semantic access and came to the conclusion that the speed of lexical access becomes prominent once a certain vocabulary level is reached. That is, in order to demonstrate significant increase in fluency (RT) at the 3k level, participants needed a vocabulary size around 5k. Furthermore, their study also indicates that in the case of non-native speakers lower frequency vocabulary items will have a longer response time as well. In conjunction with this, Tanabe (2016) found that students with a larger vocabulary size also had a quicker lexical access to meanings than students with a small vocabulary.

The importance of measuring vocabulary size with a temporal measure lies in the fact that speed is essential for the development of abilities such as speaking and reading fluency and, as confirmed by Tanabe, ‘fast-accessible lexical knowledge facilitates reading performance’ (2016, p. 135).

The aforementioned reveals some of the interactions between vocabulary size and lexical access, the effect of cognate instructions and the importance of exploring multiple languages to fully understand the mental lexicon. This also confirms Treffers-Daller et al. in that ‘one cannot discover universal principles or processes of Second Language Acquisition if the focus is on one language (English) only’ (2008, p. 271). Therefore, taking into account the influence of prior lexical knowledge in vocabulary studies seems to be justified. Particularly so in the case of Romanian and English, where the proportion of cognates accounts for over 50% on different lexical tests and many have argued that the interpretation of vocabulary scores on yes/no and form-meaning recognition tests poses some challenges in the case of speakers of Romance languages (e.g. Nation, 1983; Schmitt et al., 2001; Milton, 2009; Miralpeix and Meara, 2014). It is suggested hereby though that cognates are naturally part of real language and therefore, instead of eliminating them from lexical tests, different ways should be explored of accounting for cognate knowledge in the case of Romance languages.

3. The current research

3.1. Research Questions

In order to explore the multilingual mental lexicon of adult learners, the following research questions will be addressed:

- RQ1: Is there a relationship between overall lexical knowledge and cognate knowledge in the L2 (Romanian) and L3 (English)?
- RQ2: Does explicit instruction on cognates have a sizable effect on vocabulary scores for the study group compared to the control group?
- RQ3: Is vocabulary size a significant predictor of lexical access in the case of the three languages? Does lexical access systematically increase as word frequency decreases (frequency effect)? Do learners with a larger vocabulary size respond faster on form-meaning recognition tests than learners with smaller vocabulary sizes (size effect)?
- RQ4: To what extent does the questionnaire support the results on the English and Romanian lexical tests?

3.2. *Participants*

In order to test the effect of explicit instruction of structural relationships between a Romance L2 and English L3, and compare vocabulary sizes, we tested the lexical knowledge of 54 Hungarian native speakers, who live in Romania and therefore are exposed to Romanian as a second language, and study for a Bachelor's Degree in English Language and Literature (1st year $n = 23$, 2nd year $n = 17$, 3rd year $n = 14$). All participants considered themselves on the questionnaire native-speaker Hungarians and only three of them consider themselves at a native level in Romanian, while all of them started learning Romanian at or under the age of 7 ($N = 54$, $min = 0$, $max = 8$, $SD = 2.07$). Regarding English, the average starting age was 9 ($N = 54$, $min = 2$, $max = 14$, $SD = 4.4$). Furthermore, the background questionnaire reveals that 76% of participants never learnt French and those who did, started on average at the age of 15 ($min = 11$, $max = 21$) and very few learnt Spanish (four participants) and Italian (two participants) respectively as an L4. Just over 50% of students learnt German as an L4, some of them as beginner and some at advanced level. However, as the current study focuses on the effect of bilingualism and the relationship between an advanced L2 Romance language and English L3, the students' German knowledge was not tested.

3.3. *The instruments*

3.3.1. *The vocabulary size tests*

To replicate Molnar (2010) and Szabo (forthcoming) in a more controlled, online experimental paradigm, it has been decided that instead of the VLT, the current study will opt for the VST. There are several reasons for this. First, compared to the VLT, the format of the VST allows an online measure where speed of response can be recorded. Although this would be possible with the VLT if only one item out of three was presented at one time (see Laufer and Nation, 2001), the fact that participants would have already seen the 6 distractors, during the second and third items, the response time measures would be convoluted to interpret (see Harrington, 2006).

Extract from Level 1 of the VST (Nation and Beglar, 2007)

1. SEE: They saw it.

- a. cut b. waited for
c. looked at d. started

Second, as Nation (2001) also admits, the VLT is a diagnostic tool that allows teachers to investigate whether students need to focus on high or low frequency vocabulary, and perhaps how many academic words they know. Whereas, the VST was developed as a comprehensive written vocabulary size test that measures vocabulary knowledge on all fourteen levels of the BNC list (Nation, 2006), testing 10 items at each level. The VST has also been used widely in different studies in which its validity and reliability have been explored (see e.g. Beglar, 2010; Gyllstad, 2012; Gyllstad, Vilkaite and Schmitt, 2015). Compared to yes/no or checklist tests which use a self-report binary-choice format, the VST, as a form-meaning mapping multiple-choice test, allows for: '[...] a wide range of content to be sampled efficiently; (2) [...] to be used with learners from a variety of language backgrounds (i.e., many learners are familiar with the multiple-choice format); (3) control the level of difficulty of the items by demanding approximately the same degree of knowledge for each item (achieved through the consistent use of one set of item writing procedures); (4) [...] marking as efficiently and reliable as possible; and (5) [...] learners demonstrate knowledge of each item' (Beglar, 2010, p. 26). Although Wesche and Paribakht (1996) and Gyllstad et al. (2015) raise some concerns regarding multiple-choice tests like the VST, such as test takers could rely on their knowledge of the distractors, thus eliminating incorrect responses and increasing the chances of guessing, which according to Stewart (2014) and confirmed by Gyllstad et al. (2015) could be as high as 25%, it has to be noted that overall it is hypothesized that students guessing rate or the overestimation of their vocabulary scores will be proportional to their proficiency on all three vocabulary tests. In other words, this means that although the VST is not unproblematic, the ROVST and HUVST will have the same characteristics, therefore allowing us to compare vocabulary size and more importantly, investigate the relationship between these languages. Moreover, the lack of standardized vocabulary tests pose a challenge to researchers in the sense that we have to weigh the advantages and disadvantages of each test and find the best fit for the current purposes.

Third, we know from Beglar (2010) that using only 50% of the items (70), the test still remains valid and reliable, which allowed the author to reduce the number of items for the current study as testing 140 items in 3 different languages (420 items in total) would considerably reduce participants' interest and motivation in taking part in the study. This, based on Schmitt et al. (2001), would be impossible with the VLT as the authors call for increasing the number of items on the test. Additionally, at the design stage, the author was fully aware that according to the Classical Test Theory and also confirmed by Gyllstad et al. (2015) and Schmitt et al. (2001), in order to increase the validity of the test 'the more item the better' approach should be adopted. However, as the aforementioned emphasizes, the purpose of the current research is to

explore the interconnections between multiple languages and the intricacies of the multilingual mental lexicon.

This way, the VST has been split into two parts following Elgort's (2013) criteria, and avoiding items that have been deemed unreliable or misfitting according to Beglar's item analysis (2010). These problematic items and false cognates (3 on the entire test) have been replaced with the same part of speech wherever possible. The thus remaining 70 items (henceforth ENVST) distributed equally on the 14 frequency levels have been checked for cognateness in Hungarian and Romanian, word length as number of syllables, and part of speech.

Strictly following Nation and Beglar's (2007) criteria for item, response and distractor selection, for the ROVST Szabo's Romanian Word List was used (2015). In the case of the HUVST Varadi's frequency list was used, which is derived from the Hungarian National Corpus (2002). During the compilation stage every effort has been made to select distractors and responses that come from the first 2000 most frequent words on the same list, or are of higher frequency than the items themselves. Out of the 70 items on the ENVST, 31 have been identified as non-cognate in either Romanian or Hungarian. This defined the number of cognates to be selected for the ROVST and HUVST. Taken all three tests together, 93 out of 210 are non-cognates, which accounts to 56% of items being cognates equally distributed on all three tests and levels. Furthermore, on the ENVST and ROVST 50% (19 and 20 cognates out of 39) are only cognates in Romanian and English. The other 50% are multilateral cognates (including Hungarian as well). It has to be noted that this proportion does not represent the proportion of cognates in real language, but the cognate distribution on the original VST. The distribution of cognates on the Hungarian test slightly differs with 24 multilateral cognates versus 15 (Hungarian – Romanian) due to the lack of Hungarian words that are cognates in Romanian, but not in English (that also match the frequency level and selection criteria on the Hungarian frequency list).

Moreover, the nouns, verbs and adjectives are equally distributed on the three tests and, importantly, within each level, with 49 nouns, 12 verbs and 9 adjectives. Although the test may seem heavily noun-bound, as we know, some English words can easily change part of speech (e.g. She looked upset (*adj.*); the accusation upset him (*verb*); this was the greatest upset for them (*noun*)) and as the VST is testing knowledge of word family, it is assumed by Nation and Beglar (2007) that learners will be aware of the most common conjugations and derivations of the items. Throughout the test design it was also ensured that the number of syllables for each item is distributed proportionally. Thus, the number of syllables varies from 1 (min.) to 6 (max.), with $M = 2.2$ on the ENVST, $M = 2.8$ on the ROVST and $M = 2.4$ on the HUVST. This shows that the mean number of syllables for each test was roughly equal.

Finally, similarly to the VST, all items on the Romanian and Hungarian tests have been put in a non-defining sentence (see example above). Both tests have been checked by an additional educated native speaker from that language for cognateness,

item-distractor relationship (i.e. that there are no distractors that could be misleading or equally acceptable as responses) and errors.

3.3.2. *The questionnaire*

Most research to date that focuses solely on the target language, collects only basic information from the participants, such as their age, the age they started learning the target language and gender. As the focus in the present study is on the learners, the aim of the questionnaire was capture a more detailed picture of their language background and allow a comparison of their vocabulary scores with their perceived language knowledge. Moreover, as one of the primary aims of the study is to investigate the relationship between the typologically closer languages in terms of vocabulary size and cognate knowledge, the final questions on the questionnaire attempted to gather information on whether participants rely more on their L1 or L2 when using the L3.

Overall the questionnaire comprised 10 questions, 6 at the beginning of the study, 3 at the end with the fourth question being relevant only to the study group. The first question required students to tell at what age they started learning the 3 languages addressed in the study with the possibility to add more languages they knew. The following questions were Likert-scale questions, ranging from 1 to 5 (1 being 'not very', 3 being 'fairly' and 5 being 'very') and asked students to self-rate their proficiency and confidence levels in the 3 languages in different situations (e.g. everyday situations, formal settings and work). Questions 7 and 8 were multiple choice questions and focused on English and its relationship with the other languages in reading and writing contexts. Questions 9 and 10 was interested especially in students perception of cognates and how much they think Romanian helps them with English, where question 10 was especially addressed to the study group in form of an open question so they can reflect on the explicit instruction on cognates.

3.3.3. *Reliability of the lexical tests and the questionnaire*

The item reliability index for the ENVST was high, using Cronbach's alpha, at $\alpha = .83$, with no items significantly improving or worsening the score if dropped. The internal consistency of the test based on the 14 levels also shows high reliability at $\alpha = .8$ (lower $\alpha = .7$, upper $\alpha = .91$) and dropping any of the levels would only decrease the alpha coefficient. Concurrently, Cronbach's alpha for item reliability on the ROVST was calculated at $\alpha = .84$, with all items contributing in the $+0.01$ or -0.01 range if the item is dropped. The reliability, based on the scores, on the 14 levels also shows an acceptable level at $\alpha = .83$ (lower $\alpha = .74$, upper $\alpha = .92$). Due to a ceiling effect on the Hungarian test, which was expected as participants were Hungarian native speakers, the reliability scores for the HUVST items were only acceptable at $\alpha = .65$ (lower $\alpha = .5$, upper $\alpha = .79$) and for the 14 frequency levels at $\alpha = .56$. This indicates that the variance of items and correlations between items are reliable on the ENVST and ROVST, however this is not the case on the HUVST. Considering only the Likert-sale questions on the questionnaire ($N = 54$, no. of questions = 13), Cronbach's $\alpha = .87$, which indicates high internal consistency.

3.4. Procedure

Following ethical approval, the aforementioned tests and the questionnaire have been transferred to a free survey website (Qualtrics.com) that records the number of clicks to each question, the time taken between clicks and submission time. The three language tests, the ENVST, the ROVST and the HUVST, in this order respectively, were preceded by half of the questions from the questionnaire, recording name, age and general language background of the participants. Since the website does not allow for a controlled randomization, all tests have been taken in the above order without counterbalancing languages or items in the tests. Furthermore, since the study already involves an experimental paradigm, any significant differences in either the vocabulary scores or RT measures would have been difficult to decide whether they are a result of counterbalancing between languages or items, or the treatment condition. The data was collected during normal teaching hours at the Partium Christian University in a phonetics lab where each computer has the same screen size and speed, and the participants are familiar with their use. Due to the way the survey website works, and the lack of touch screen monitors or smart phones, the survey was completed using a standard two-buttoned mouse. This can be considered a limitation of the study, however it has to be noted that currently no RT measure allows for an exact or accurate measure of the duration of lexical accessibility in our brains. Additionally, a response time measure becomes relatively high if participants have to choose the correct option out of four definitions and a non-contextual sentence is also provided. Therefore, the RT measures in this study have to be taken as relative measures under the given conditions, which in the same time overcome the possibility of confusion over the 4 answer and a further 'pass' button on a regular keyboard reported by Tanabe (2016).

First, second and third year students were randomly allocated into two separate groups, a study (experimental) group and a control group. The study group ($n = 32$) received a cognate recognition task: a short text in Romanian, Spanish and English, in which they had to highlight cognates that they recognized. Following this, they also received a handout that outlined orthographic and structural similarities between similar and identical Romanian and English cognates, and participants also had the opportunity to ask questions. The aim of this exercise was to raise awareness of lexical similarities between the two languages. Both groups received a consent form and no participants were told that the vocabulary test is timed. Moreover, none of the participants reported any visual impairment or physical difficulties with using a standard desktop PC. The whole test was taken in one single session with no breaks and it took participants around 50 minutes on average to complete.

After ensuring that participants have been anonymized, the descriptive and inferential analysis have been conducted in R (R Core Team, 2015) using the ggplot2 package for all graphical representations.

4. Results

Descriptive statistics for the results on the three tests are provided in Table 1. The results, conform to the expectations, indicate that participants scored better and more consistently on the Romanian test than on the English test. The difference between the scores on the two tests was checked using a paired samples t-test, which indicates that the difference is statistically significant ($t(1, 53) = -4.63, p < .001$) with a large effect size $R^2 = .537$. The difference between the L1 and the L3/L2 is obvious from the very high difference in means ($M = 64$ compared to 41 and 46 out of 70) and the very consistent scores ($SD = 2.6$ compared to 8.5 and 7.9) on the HUVST.

Table 1 Descriptive statistics for overall scores

Test	N	M	SD	Min	Max
ENVST	54	41.07	8.46	19	59
ROVST	54	45.74	7.88	25	59
HUVST	54	64.41	2.58	57	68

Moreover, the results confirm that the mean scores are 59% and 66% on the ENVST and ROVST respectively, indicating that participants on average do not reach native speaker levels in their L2 and L3. To translate the scores into vocabulary size estimates, we need to bear in mind that the VST employs a 1:100 sampling rate at all 14 levels. In the present study however, this rate had to be reduced by half, meaning that the sampling rate decreased to 1:200. In other words, the scores need to be multiplied by 200. Thus, the vocabulary size estimates range between 3800 (min) and 11800 (max) with a mean of 8200 on the ENVST and 5000 and 11800 with a mean of 9200 on the ROVST. As expected, this range is considerably smaller in the L1, with a minimum of 11400 and a mean of 12800 on the HUVST. Due the fact that the aim of the current study was to explore the relationship between an L2 Romance language and L3 English, and the ceiling effect and lack of variance on the HUVST, only the results in the L2 and L3 will be compared, except in the case of RQ3 where speed of response is compared rather than vocabulary scores.

In order to further explore the data and address the research questions, the general assumptions for parametric tests will be reported. As the Shapiro-Wilk normality test indicates, in the case of the ENVST ($W = 0.98, p = .62$) and the ROVST ($W = .97, p = .30$) the assumption of normality is tenable in the case of both tests. To address RQ2, the assumption of homogeneity of variance between the study and control groups was tested using Levene's test. The results on the ENVST ($F(1, 52) = 2.04, Pr(>F) = .016$) and ROVST ($F(1, 52) = 2.14, Pr(>F) = .015$) indicate that the assumption of equal variance between the groups has been met.

4.1. Receptive vocabulary in the L2 and L3 (RQ1)

The first research question was addressed by comparing the overall and cognate scores on the two tests and analyzing the different features of the test, such as

construct validity. Vocabulary tests in general are based on the assumption that the most frequent words in a language will be more familiar to learners and as the frequency of words decreases, so does the proportion of words that the students actually know.

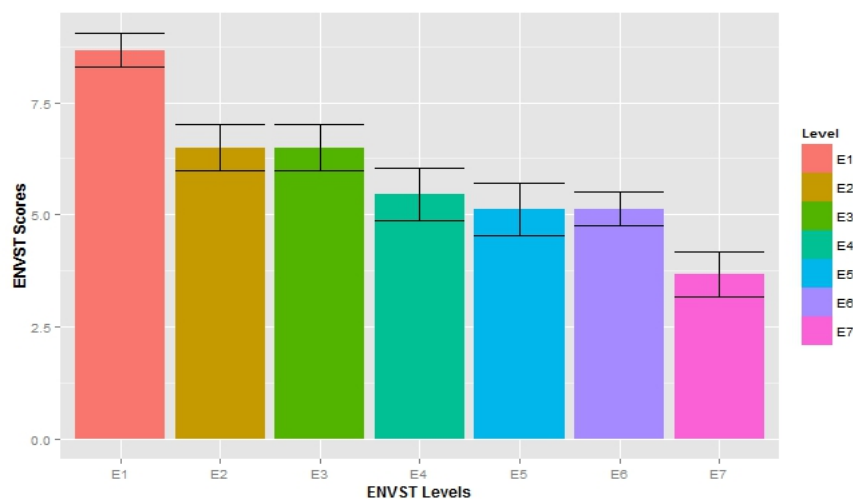


Figure 1 ENVST results by levels

Figure 1 above illustrates that students overall score better on higher frequency levels and as the items become less frequent, students' knowledge decreases as well. For ease of visual representation, level 1 and level 2 items are collapsed into one band (E1) and consistently so for all levels.

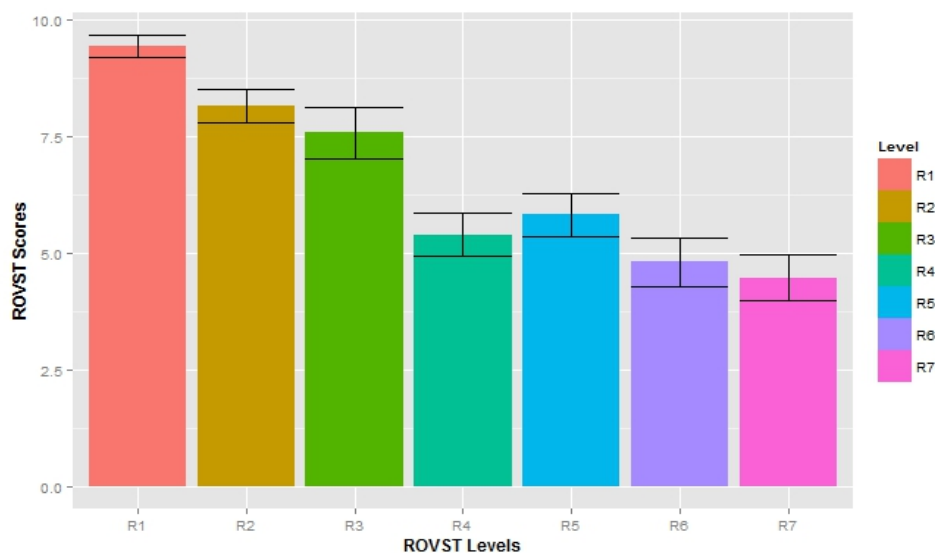


Figure 2 ROVST results by levels

It can be noted on Figure 2 that students' lexical profile follows roughly the same trend in their L2 as well. Taken together, this downward slope indicates that words

from lower frequency ranges will have a lower hit rate, indicating decreasing knowledge on these levels. This confirms that the tests' construct validity is acceptable and is in line with Meara's theoretical profile of a language learner (1992).

In order to compare participants' knowledge in the L2 and L3, scatterplots were consulted to establish whether there is a linear relationship between the variables. This is provided in Figure 3.

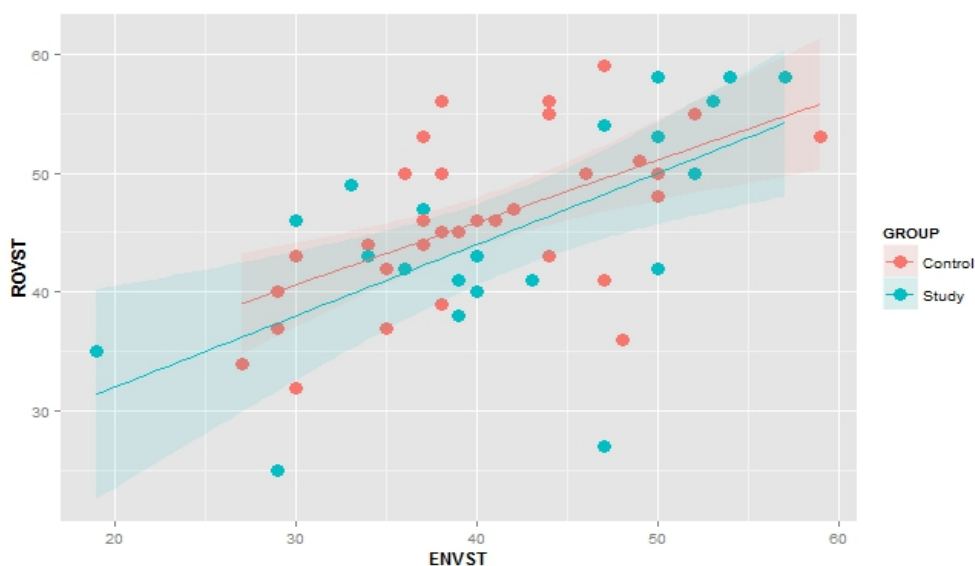


Figure 3 Scatterplot with overall results separated by group

Figure 3 confirms the expectations that there is a linear relationship between lexical knowledge in the L2 and L3. This is also supported by Pearson's correlation, which indicates that there is a statistically significant relationship between the overall scores on the two tests: $r = .592$, $p < .001$, $N = 54$. Similarly, if only cognate scores are examined, the results indicate a statistically significant correlation at $r = .629$, $p < .001$. Interestingly, the results also confirm that the higher the cognate score is, the higher the overall score will be ($r = .639$, $p < .001$ between L2 cognate score and L3 total score and $r = .622$, $p < .001$ between L3 cognate score and L2 total score).

Another feature of a valid vocabulary test is that it is capable of distinguishing between proficiency levels. Bearing in mind that the necessary assumptions, such as normality and homogeneity of variance are tenable, an ANOVA using contrasts was conducted to explore whether lexical proficiency in the two languages is affected by year of study. This confirms that scores on the ENVST significantly increase from year 1 and 2 to year 3 and that year is a significant predictor of the model ($F(2, 51) = 8.50$, $p < .001$; *eta squared* = .5, large effect). The contrasts between years suggest that the difference between 1st and 3rd year is somewhat significant ($t(2, 51) = 2.1$, $p = .04$) and the effect is very significant between 2nd and 3rd year ($t(2, 51) = 3.7$, $p < .001$). The fact that there is no difference between 1st and 2nd year students ($M = 38.87$, $n = 23$ and $M = 38.24$, $n = 17$) is due to the fact that the model only explains 25% of variance,

so the other differences can come from individual differences, hours of exposure, baseline vocabulary size at the start of university, motivation, etc., the effects of which might not be accentuated until students' 3rd year of study. Furthermore, there is no similar effect on the ROVST ($F(2, 51) = 2.89, p = .064; \eta^2 = .06$, small effect). This finding is congruent with the expectations as student's lexical knowledge in English is bound to increase by the 3rd year of their studies. However, since they do not receive any formal instruction in Romanian during their studies, their Romanian vocabulary size is not expected to increase by year of study.

4.2. The effect of explicit instruction on cognates (RQ2)

The second research question explores whether explicit instruction on cognates in the case of adult learners at an intermediate to high proficiency level has any effect on their vocabulary size in English.

Table 2 Descriptive statistics for the ENVST scores based on levels and group

Group		<i>N</i>	Total	E1	E2	E3	E4	E5	E6	E7
Study	<i>M</i>	22	42.18	8.64	6.36	6.91	5.73	5.50	5.27	3.77
	<i>SD</i>		7.6	1.76	2.17	1.41	2.03	2.20	1.39	2.00
Control	<i>M</i>	32	40.31	8.69	6.59	6.22	5.28	4.88	5.03	3.62
	<i>SD</i>		9.7	1.12	1.76	2.07	2.23	2.06	1.40	1.74
Total	<i>M</i>	54	82.49	8.67	6.5	6.5	5.46	5.13	5.13	3.69
	<i>SD</i>		8.5	1.40	1.92	1.85	2.14	2.12	1.39	1.83

Table 2 provides descriptive information on the difference between the performance of the study group and control group on the different frequency levels and in total. According to this, the study group reached higher means than the control group, however, based on a t-test this difference is not statistically significant $t(38.04) = -0.76, p > .05$ with a small effect size ($R^2 = .122$). The same outcome can be found on the Romanian test as well ($t(36.45) = 0.31, p > .05$ with a small effect size ($R^2 = .051$). This lack of significant effect of the instruction on cognates will be further discussed in the following section.

4.3. Lexical access (RT): size and frequency effects (RQ3)

Response times to hits (correct answers) indicate that on average it took 6 seconds to submit a response to a Hungarian item, while 13 and 10 seconds to the English and Romanian items respectively. This difference is illustrated in Figure 4 and if we compare this to the mean results in Table 1, it can be extrapolated that the more proficient students are in a language, the faster and more accurate their responses will be on the test. Therefore, speed of response can be a good indication of language proficiency in the case of multilinguals.

Following Laufer and Nation (2001) and based on Tanabe (2016), the participants have been equally divided into two groups (*High* and *Low groups*) using the median-split on their ENVST scores in order to establish whether the group with a larger

vocabulary size also showed faster accessibility in the tested foreign languages. Table 3 illustrates just this. It becomes obvious from the mean scores and paired t-tests that there is a statistically significant difference between the two groups' L3 ($t(26) = 9.10$, $p < .001$) and L2 vocabulary scores ($t(26) = 2.87$, $p < .01$). Furthermore, in support of size effect, the t-tests also indicate that students with a larger vocabulary also had a faster response time in English ($t(26) = -2.03$, $p = .04$), but not Romanian ($t(26) = -1.14$, $p = .26$). The reason for this probably lies in the fact that the more proficient the group is in a language, the less variability can be detected in their RT scores.

Table 3 Vocabulary size and mean RT

Group	ENVST				ROVST			
	Vocabulary Size		RT		Vocabulary Size		RT	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
High ($n = 27$)	47.96	4.85	11.69	3.38	48.92	7.80	9.55	2.28
Low ($n = 27$)	34.18	4.88	13.58	3.42	42.56	6.70	10.20	2.10

Figure 4 furthermore shows that there is a slight difference between the response times to cognates and non-cognates. In order to test whether the difference overall between response time to correct answers on cognates and non-cognates is significant, t-tests were conducted. In the case of the ENVST, $t = -.689$, $p = 0.491$, $R^2 = .016$ (cognate $M = 12.43$ and non-cognate $M = 12.64$), which indicates that there is no difference between the means. However in the case of the Romanian test ($t = -2.23$, $p = 0.03$, $R^2 = .048$, cognate $M = 9.61$ and non-cognate $M = 10.12$) and the Hungarian test ($t = -2.27$, $p = 0.02$, $R^2 = .039$, cognate $M = 6.27$ and non-cognate $M = 6.54$) there is a significant difference between the means, indicating that even in the case of multiple choice tests cognates are likely to be recognised faster than non-cognate items, albeit the advantage of cognates is not as accentuated as in yes/no studies.

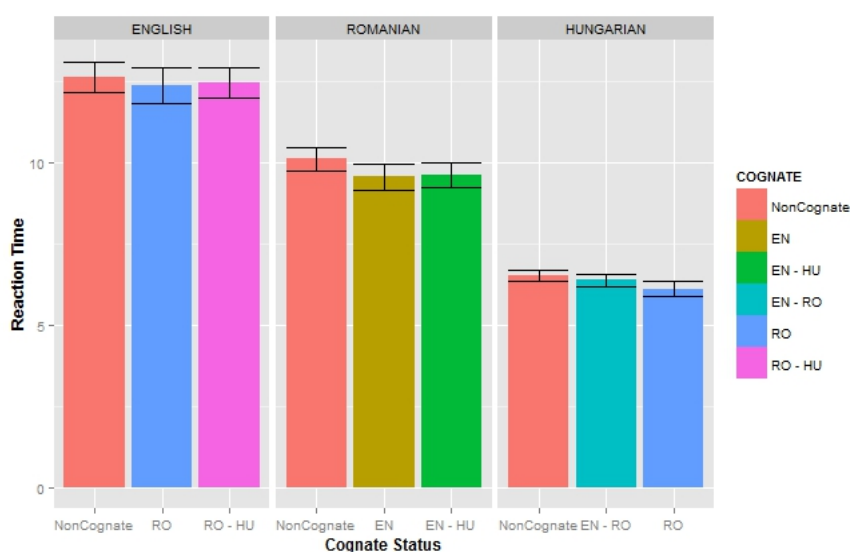


Figure 4 RT on correct answers divided by language and cognate status

Previous research also suggests that speed of access becomes slower for lower frequency items (Harrington, 2006; Laufer and Nation, 2001, Tanabe, 2016). This is investigated in Table 4.

Table 4 Response times on the tests based on frequency levels

Test	N	Total	T1	T2	T3	T4	T5	T6	T7	
ENVST	M	54	651.8	91.4	112.2	114.3	126.9	131.7	129.3	123.6
	SD			31.2	39.5	41.6	40.4	44.3	42.5	40.4
ROVST	M	54	829.3	75.0	76.8	89.6	129.7	90.4	99.67	90.6
	SD			19.8	21.3	21.4	42.6	30.5	31.4	34.1

* Please note: mean response times are expressed in seconds. T1 corresponds to response times on band one and two (first 2000 words)

This concurs with previous studies in suggesting that high frequency items will have lower response times. Generally, this is the tendency on both tests. For instance, on the ENVST compared to the first band where the average speed of response is 91 seconds, on and above the fourth band (7k and 8k word levels and beyond) this somewhat gradually increases to around 130 seconds. Whereas on the Romanian, the lowest average speed is 75 seconds on T1 and it gradually increases to above 90 seconds, except for T4 which stands out having the highest time of 130. If this is compared to Figure 3 where on the same frequency band (R4) a slight dip in scores can also be noticed, which suggests that some of the items on this level might be more unfamiliar for some reason to students than on the surrounding bands. A closer analysis suggests though that the multilateral cognate *monopol* with only 6 hits on the Romanian test had a misleading distractor and the majority of the students chose this (*governed by a single party*) instead of the correct answer (*exclusive right*), which explains the dip on R4 in Figure 3. On the other hand, the response times on the 10 items on this level indicate that 3 items, 2 non-cognates and a Romanian – English cognate have slightly higher response times, which suggest words on the 4th level behave differently in terms of number of responses and mean response times for different reasons. Nevertheless, considering the length of the test and the multiple choice format, by and large, the assumption that response time proportionally increases with word frequency seems justified.

4.4. Results on the questionnaire (RQ4)

To explore the relationship between students' self-rated proficiency and confidence levels in the two foreign languages and their vocabulary scores, Spearman's rank correlation was used. These confirm that the higher the students' vocabulary score on the English test: the more proficient they feel in that language ($\rho = .50, p < .001$); the more confident they feel using it in everyday situations ($\rho = .52, p < .001$) and in formal settings ($\rho = .52, p < .001$). Similarly, the analysis confirms that there is a moderate correlation between the ENVST scores and students' self-rated confidence levels in using English in a work situation ($\rho = .47, p < .001$), and if they moved to an English speaking country ($\rho = .41, p < .001$).

Considering the ROVST and the self-ratings, the results indicate that the test is a significant predictor of: self-rated proficiency ($\rho = .50, p < .001$), confidence levels in everyday ($\rho = .53, p < .001$) and formal ($\rho = .50, p < .001$) situations. Moreover, there is a weak negative correlation between ROVST scores and the age they started learning Romanian ($r = -.27, p = .05$), however this is not the case in English. The reason for this is that regardless of the age they started learning English, students might have had different amounts of exposure to English, for instance some might have been in “intensive English” classes (6-8 hours / week) in high school.

These results show the vocabulary size tests are good indicators of how proficient and confident students feel in using a foreign language, at least in this cultural context (e.g. Japanese students tend to be very modest on Likert-scale type questionnaires).

Table 5 Descriptive statistics for self-rating questions

Proficiency self-rating										Confidence in everyday situations										
English					Romanian					English					Romanian					
<i>N</i> = 54	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
%	4	7	41	35	13	6	24	30	35	6	6	15	28	28	24	13	26	20	28	13
Confidence in formal settings										Efficiency at work in an international company										
English					Romanian					English					Romanian					
<i>N</i> = 54	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
%	6	19	50	20	6	26	28	24	20	2	4	22	46	22	6	7	15	31	30	17

Table 5 summarizes the responses in percentages to proficiency and confidence questions from the questionnaire[†]. This reveals that only 11% of the students consider themselves not proficient in English, compared to 30% in Romanian. Taking into account the fact that all students have lived most of their lives in Romania, this is a somewhat surprising finding, however it is in line with findings in Szabo (forthcoming). Regarding confidence ratings in everyday situations (e.g. in a restaurant, shopping, chatting to friends) and formal settings (e.g. job interview, presentation, business meeting) it is suggested that 39% and 54% students respectively do not feel confident enough in using their L2, despite the fact that their vocabulary score is higher on the ROVST than on the English test. In comparison, only around 25% of the students do not feel confident in English and the majority rated themselves as fairly to very confident in both situations. Perhaps desire is mixed with reality, nevertheless, 74% of the students claim that they could use English efficiently at an international company, compared to 78% when it comes to Romanian.

[†] For the tests, questionnaire or more information, please contact the author: csaba.szabo@open.ac.uk

Questions 7, 8 and 9 had the aim of gathering information on how students perceive the links between their languages and what strategies they use. According to the answers, when looking for an English word while writing, 83% of the students think of suitable Hungarian words and translate those, instead of relying on their Romanian knowledge. This finding, besides confirming the findings (91%) in Szabo (forthcoming), reflects the fact that most of the students have Hungarian native speaker English teachers in schools and thus the link between these two languages becomes stronger and, arguably, with a faster or more direct lexical access connection within the network. Concurrently, when reading, 48% of the students translates unknown L3 words into Hungarian and 46% tries to guess from context. 6% chose the ‘other’ strategy option and no student chose options *b* (‘I try to think if I’ve seen the word in Romanian’) or *c* (‘I try to find a Romanian translation’). This clearly shows that the L1 to/from L3 is a very strong link, despite the similarities between the L2 and L3. The last question (9) reveals that, contrary to students’ strategies while reading or writing in English, over half of them (54%) think that Romanian is useful and only 11% consider it to be confusing or misleading.

Question 10 provided students with the opportunity to comment on the cognate-awareness tasks and express their opinion on how much they think Romanian would or would have helped them in learning English. The large majority of students considered the explicit instruction on cognates and the task useful. Their comments highlight the fact due to the similarities between Romanian and English, learning their L3 is definitely enhanced by Romanian, some of them also pointing out that this is probably the reason why Romanian native speakers are better in English than Hungarians (confirmed by Molnar, 2008). Due to the cognate facilitation effect, many of the participants point out that learning English through Romanian would have been easier, but it also depends on the level of the L2. Some of the following examples from students’ comments reflect these issues:

First of all I liked the text, since it is about Madagascar. (Yupiii) Making the comparison between the 3 texts I came to the conclusion that these languages share much more than I have imagined before. When I make translations I never think in Romanian however now it seems it is a good idea. [...] On the other hand I believe, Hungarians cannot speak Romanian on a level to learn another language on its basis [...].
SS50

It would be very useful because having similar base or spelling with Romanian words would lead to a logical acquisition of English vocabulary and it would make it easier to retain the words and use them correctly. SS25

In sum, the analysis of the questionnaire seems to indicate that there is no discrepancy between the test results and how students perceive their proficiency and confidence levels in English and Romanian. Furthermore, in contrast with the proximity between the two additional languages and the fact that most students accept that Romanian can be helpful when dealing with (learning, processing and using) English, their strategies indicate that simply making L2 and L3 connections is

unnatural, which in turn suggests that the study has teaching and learning implications.

5. Discussion

With vocabulary size tests being widely used, especially form-meaning mapping multiple choice tests, such as the VST, but studies mainly exploring the lexicon of learners of English as a single construct without taking into account prior lexical knowledge, the present study set out to explore the mental lexicon of multilinguals through three concomitant issues regarding vocabulary size. First and foremost, although it is known from studies such as Cenoz (2001) that transfer to the target language will typically come from the typologically closer language and that students' proficiency in one language can be a good indicator of their proficiency in the other languages (e.g. Cook, 1992; Cenoz and Genessee, 1998 and Cenoz, 2000), the fact whether vocabulary tests are sensitive enough to detect such differences has been relatively scarcely explored. Furthermore, albeit many raise concerns regarding the overestimation of vocabulary size in the case of Romance languages, a reasonable way to control for this overestimation and quantify the influence between such languages and English has not yet been fully investigated. The results presented hereby confirm Molnar (2010) and Szabo's (forthcoming) findings in several aspects. On the one hand, it shows that there is a sizable correlation between L2 Romanian and L3 English lexical knowledge, suggesting that the higher the vocabulary size in one language, the higher it is likely to be in the other language. By way of extension, since vocabulary is a significant predictor of overall language proficiency, the assumption that low proficiency in the L2 will result in similarly low proficiency in the L3 seems to be reasonable and justified. On the other hand, it appears that accounting for cognate knowledge by ensuring that the original test design is rigorously followed on all tests is a meaningful method for ensuring that cognate knowledge is controlled on all tests. As this study employs the VST instead of the VLT and therefore the test items differ, the fact that overall lexical dexterity and cognate knowledge indicate a strong relationship between L2 and L3 is further justified. Moreover, based on these and Szabo's (forthcoming) findings, it can be asserted that cognates are equally recognized indifferent of their L2 or L3 form. One finding that diverges from the latter study is whether students' vocabulary size is higher or lower in English than Romanian. While the present findings indicate that there is a significant difference and students have a larger vocabulary size in Romanian, Szabo (forthcoming) did not find a significant difference between the two scores and even the mean scores suggested a slightly higher English proficiency. The reason for this might lie in the fact that the two populations are different, so students' language background can differ, or the variation is entirely due to the two different tests, something that needs to be addressed in future studies.

Second, following Molnar (2010) and Szabo (forthcoming), the findings reported here support the Parasitic Hypothesis for the explicit instruction and the cognate awareness task did not lead to greater vocabulary gains in the case of the study

group, pointing in the direction that adult learners at intermediate or above levels recognize cognates intrinsically and no explicit instruction is required. Naturally, as Szabo (forthcoming) also concludes, albeit cognate recognition, in full accord with the least effort principle (Zipf, 1949) and other economic reasons that might be governing the underlying factors of mental representation, processing and use (Hall, 2000), can be regarded as automatic, it also has to be noted that the vocabulary size tests might not be sensitive enough to detect the effect of such short awareness-raising tasks from which there is no direct vocabulary gain.

Third, since studies examining the interrelations between lexical proficiency in multiple languages, or using the VST per se for that matter, and speed of response have heretofore been non-existent, the aforementioned findings shed light on a number of important factors. On the one hand, based on the current findings, it can be corroborated that using the multiple-choice format with a temporal method, language proficiency can be accounted for, i.e. the larger the vocabulary size is in one language, the faster the response times will be in that language, confirming the size effect. On the other hand, the findings also confirm the frequency effect in not just English, but Romanian as well, this being evidence for construct validity as well. Furthermore, the study also confirms a cognate facilitation effect as both in Hungarian and Romanian cognates were recognized faster with a higher item facility ratio. The fact that there was no significant difference between response times to cognates and non-cognates on the ENVST has to be further investigated however.

Similarly to Szabo (forthcoming), the questionnaire also reveals a number of important aspects of students' language background, self-rated proficiency and confidence levels, and thus complements well the results found with the vocabulary tests. The results indicate that vocabulary size is a significant predictor of students' self-rated proficiency and confidence levels in different situations and this is valid for not just the ENVST, but the ROVST as well, reconfirming the importance of vocabulary as a core element of additional language learning. Furthermore, despite the unquestionable lexical similarities between the two additional languages, the data points to a strong connection between the L1 and L3. The answers suggest that according to students' perception, in the case of reading and writing, they link words in their mental lexicon to concepts in their mother tongue and do not actively access the L2 for form-meaning connections. This can indicate that although there is a strong link between the L2 and L3 due to cognates, the underlying factors behind word recognition and processing seems to be automatic and not something that is done conscientiously.

6. Conclusion

The aim of this paper was to explore the relationship between vocabulary size and cognate knowledge in the L1, L2 and L3 from three different perspectives. First, the results indicate that, in the case of Romance language speakers, there is a strong relationship between vocabulary knowledge in the L2 and English L3 not only on a

cognate level, but overall as well. Therefore, it has been argued that as vocabulary is considered the sine qua non of language learning and proficiency, the more proficient a learner is in the L2, the more proficient s/he is likely to be in a typologically proximate language as well, in this case English. Furthermore, the results, aligning with Szabo (forthcoming), indicate that as both the VLT and VST comprise over 50% of cognates for Romance speakers, such words have to be accounted for when testing a mixed group of language learners. From a multilingual perspective, it has been shown that prior lexical knowledge will have an influence on English vocabulary scores due to these similarities between languages, which in turn are believed to carry profound influence for pedagogy and language learning. Although we can only speculate at present, drawing on prior lexical, and perhaps linguistic knowledge, at the very early stages of teaching English might create and increase L2 to L3 links, which in turn can facilitate language learning and increase speed of access, at least in the written modality. Future studies, could investigate consequently cognate recognition and the effect of vocabulary size and frequency in the aural modality as well.

Second, given the strong L2-L3 link, in order to understand the underlying processes behind cognate recognition, the study investigated the effect of explicit cognate instruction on vocabulary scores in the case of intermediate to proficient adult learners of English. The results converge with previous studies and suggest that cognate recognition is an automatic process, therefore if a word is recognized in any of the known languages, its meaning can be transferred to the other languages as well, thus supporting the Parasitic Model. However, it has also been suggested that previously employed tests might not be sensitive enough to measure the effect of such instruction and perhaps future studies could reveal more in a longitudinal experimental design conducted with participants at different age and proficiency levels.

Third, as Laufer and Nation (2001) consider that speed of lexical access is directly related to fluency, this paper also elaborated on the findings provided by a relative measure of form-meaning recognition time. The results indicated that the temporal measure rarely employed in lexical studies has the potential of defining language proficiency and reinforcing the importance of the frequency effect. In order to explore whether fluency can be directly linked to vocabulary size should be perhaps explored in future studies in which the number of cognates on the test are proportionally distributed on all frequency levels. Furthermore, we do not know if speed of lexical access is affected by background languages or age, and there are only a handful of studies suggesting tentatively that lexical proficiency and speed of access are inversely proportional. Thus, studies exploring or replicating the above issues would rightly aim to try and establish a threshold of vocabulary knowledge beyond which cognate recognition or a decrease of reaction time becomes uniform.

In conclusion, the present study has confirmed that foreign or second language knowledge, and by extension teaching, learning or researching additional languages incorporated into the mental lexicon, is not as straightforward as it has been previously believed. Prior lexical knowledge will have an influence on additional

language acquisition and therefore teasing apart the different aspects and interconnections between different or similar background languages and developing robust methodologies that allow such investigations is highly pressing. Finally, it has been shown that the methodology employed in this study and the issues addressed are furthering our understanding of the mental lexicon of the multilingual.

Acknowledgements

I would like to express my gratitude to my supervisors for encouraging and supporting my research and The Open University and CREET for the funding provided. I am also grateful to Dr Imma Miralpeix and the two anonymous reviewers for their thoughtful comments on earlier versions of this paper.

References

- Ard, J. and Homburg, T. (1983). Verification of language transfer. In S. M. Gass, & L. Selinker (Eds.), *Language transfer in language learning: Issues in second language research* (pp. 157-176). Rowley, MA: Newbury House.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Barac, R., & Bialystok, E. (2011). Cognitive development of bilingual children. *Language Teaching*, 44(1), 36–54. <http://doi.org/fwdmxt>
- Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, 21(6), 522–538. <http://doi.org/d8m5mh>
- Cenoz, J. (2000). Research on multilingual acquisition. In J. Cenoz, & U. Jessner (Eds.), *English in Europe: The acquisition of a third language* (pp. 39-53). Clevedon, UK: Multilingual Matters.
- Cenoz, J. (2001). The effect of linguistic difference, L2 status and age on crosslinguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (9-20). Clevedon, UK: Multilingual Matters.
- Cenoz, J. (2013). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language Teaching*, 46, 71-86.
- Cenoz, J., & Genessee, F. (1998). Psycholinguistic perspectives on multilingualism and multilingual education. In J. Cenoz, & F. Genessee (Eds.), *Beyond bilingualism. Multilingualism and multilingual education* (pp. 16-32). Clevedon, UK: Multilingual Matters.
- Chen, X., Ramírez, G., Luo, Y. C., Geva, E., & Ku, Y. M. (2012). Comparing vocabulary development in Spanish- and Chinese-speaking ELLs: The effects of metalinguistic and sociocultural factors. *Reading and Writing* 25, 1991-2020.
- Cobb, T. (2000). One size fits all? Francophone learners and English vocabulary tests. *Canadian Modern Language Review*, 57, 295-324.
- de Angelis, G. (2007). *Third or additional language acquisition*. Clevedon, UK: Multilingual Matters.
- De Angelis, G., & Selinker, L. (2001). Interlanguage transfer and competing linguistic systems in the multilingual mind. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic*

- influence in third language acquisition: Psycholinguistic perspectives* (pp. 42-58). Clevedon, UK: Multilingual Matters.
- García, G. E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly*, 26, 371-92. <http://doi.org/fpmhpb>
- Gass, S. M., & L. Selinker. (2001). *Second language acquisition* (2nd ed.). Hillsdale, N.J.: Lawrence Erlbaum.
- Gyllstad, H., Vilkaite, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL International Journal of Applied Linguistics*, 166, 276–303.
- Hall, C. J. (2002). The automatic cognate form assumption: Evidence for the parasitic model of vocabulary development. *International Review of Applied Linguistics in Language Teaching*, 40, 69–87. <http://doi.org/cntmwz>
- Harrington, M. (2006). The lexical decision task as a measure of L2 lexical proficiency. *EUROSLA Yearbook*, 6, 147-268.
- Helms-Park, R., & Perhan, Z. (2016). The role of explicit instruction in cross-script cognate recognition: The case of Ukrainian-speaking EAP learners. *Journal of English for Academic Purposes*, 21, 17-33. <http://doi.org/bq3n>
- Laufer, B. & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, 54, 469–523. <http://doi.org/fv2rvk>
- Laufer, B., & Nation, P. (2001). Passive vocabulary size and speed of meaning recognition. *EUROSLA Yearbook*, 1, 7–28.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.
- Meara, P. & Milton, J. (2003). *X_Lex, The Swansea Levels Test*. Newbury: Express.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*, Bristol, UK: Multilingual Matters.
- Milton, J. (2013). L2 vocabulary acquisition, knowledge and use. *EUROSLA Monographs Series*, 2, 57–78.
- Milton, J., & Daller, H. (2007). The interface between theory and learning in vocabulary acquisition. In *Annual conference of the European second language association*, Newcastle, UK
- Miralpeix, I., & Meara, P. (2014). Knowledge of the written word. In J. Milton & T. Fitzpatrick (Eds.), *Dimensions of vocabulary knowledge* (pp. 30–44). Basingstoke, UK: Palgrave Macmillan.
- Molnar, T. (2008). Second language versus third language vocabulary acquisition: A comparison of the English lexical competence of monolingual and bilingual students. *Toronto Working Papers in Linguistics (TWPL)*, 33, 1-16.
- Molnar, T. (2010). Cognate recognition and L3 vocabulary acquisition. *Acta Universitatis Sapientiae, Philologica*, 2, 337–349.
- Nagy, W. E., García, G. E., Durgunoklu, A., & Hancin-Bhatt, B. (1993). Spanish-English bilingual children's use and recognition of cognates in English reading. *Journal of Reading Behavior*, 25, 241–59.
- Nation, P. and Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.
- Nation, P. (1983). Testing and teaching vocabulary. *Guidelines*, 5, 12-25.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

- Nation, P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59–82. <http://doi.org/djfssx>
- Pellicer-Sánchez, A., & Schmitt, N. (2012). Scoring Yes–No vocabulary tests: Reaction time vs. nonword approaches. *Language Testing*, 29(4), 489–509. <http://doi.org/bq3m>
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Reynolds, B. L., Wei-Hua, W., Hui-Wen, L., Shu-Yuan, K., & Ching-Hua, Y. (2015). Towards a model of advanced learners' vocabulary acquisition: An investigation of L2 vocabulary acquisition and retention by Taiwanese English majors. *Applied Linguistics Review*, 6(1), 121-144. <http://doi.org/bq3k>.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing* 18, 55-88.
- Stewart, J. (2014). Do multiple-choice options inflate estimates of vocabulary size on the VST? *Language Assessment Quarterly*, 11(3), 271–282.
- Szabo, Cz. (2015). Introducing a Romanian Frequency List and the Romanian Vocabulary Levels Test. *Current Issues in Linguistic Variation: The 14th international conference of the Department of Linguistics*, Vol. 2, University of Bucharest: Bucharest University Press, 304 – 316.
- Szabo, Cz. (forthcoming). Multilingual vocabulary knowledge and cognate recognition in the L2 and L3.
- Szubko-Sitarek, W. (2015). *Multilingual lexical recognition in the mental lexicon of third language users*. Berlin: Springer.
- Tanabe, M. (2016). Measuring second language vocabulary knowledge using a temporal method. *Reading in a Foreign Language*, 28(1), 1.
- Treffers-Daller, J., & Milton, J. (2013). Vocabulary size revisited: The link between vocabulary size and academic achievement. *Applied Linguistics Review*, 4(1), 151–172.
- Treffers-Daller, J. (2011). Operationalizing and measuring language dominance. *International Journal of Bilingualism*, 15(2), 147–63.
- Treffers-Daller, J., Helmut Daller, David Malvern, Brian Richards, Paul Meara, & James Milton (Eds.) (2008). Introduction: Special issue on knowledge and use of the lexicon in French as a second language. *Journal of French Language Studies*, 18(3). 269–276.
- Varadi, T. (2002). The Hungarian National Corpus. *LREC 2002*. 385–389.
- Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *Canadian Modern Language Review*, 53(1), 13–40.
- Willis, M., & Ohashi, H. (2012). A model of L2 vocabulary learning and retention. *The Language Learning Journal* 40(1). 125–137.
- Zipf, G. (1949). *Human behavior and the principle of least effort*. New York: Addison-Wesley.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the Journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

This page intentionally left blank