

# STUDIES IN THE ASSESSMENT OF FOREIGN LANGUAGE VOCABULARY

ISTVÁN THÉKES



STUDIES IN THE ASSESSMENT  
OF FOREIGN LANGUAGE  
VOCABULARY

István Thékes

Gál Ferenc College, 2018

Financing:  
GÁL FERENC COLLEGE

Lektor:  
DR. ANIKÓ NÉMETH DR. THÉKESNÉ

Cover page design, typesetting:  
ILDIKÓ VERES

Cover page photo taken by:  
ISTVÁN THÉKES

Publishing: GERHARDUS PUBLISHING 2018  
Publisher in charge: DR. GÁBOR KOZMA, RECTOR  
Address: 6720 Szeged, Dóm tér 6.

ISBN 978-615-5256-34-9

© DR. ISTVÁN THÉKES

# CONTENTS

<b>INTRODUCTION</b> .....	7
<b>BRIEF OVERVIEW</b> .....	8
<b>1 INDIVIDUAL WORDS</b> .....	9
1.1 ASSESSING VOCABULARY .....	9
1.2 ASPECTS OF KNOWING A WORD IN A FOREIGN LANGUAGE ....	11
1.3 WORD FORM AND MEANING .....	13
1.4 FOREIGN LANGUAGE VOCABULARY TESTS .....	14
1.4.1 The Receptive Vocabulary Levels Test .....	16
1.4.2 Productive Vocabulary Levels Test .....	17
1.4.3 Vocabulary Knowledge Scale .....	18
1.4.4 Vocabulary Size Test .....	20
1.4.5 The X_Lex Test .....	21
1.4.6 Diagnostic online English and German receptive vocabulary size test for YLs .....	21
1.4.7 Principles of designing FL vocabulary tests for YLs .....	22
1.5 RESULTS OF DIAGNOSTIC TESTS ASSESSING YLS' VOCABULARY ...	24
<b>2 CORPUS LINGUISTICS</b> .....	27
2.1 THE USE OF CORPORA IN VOCABULARY RESEARCH .....	27
2.2 CORPORA AND THEIR DEVELOPMENT .....	27
2.3 WIDELY USED CORPORA .....	28
2.4 APPLICATIONS OF CORPORA IN FL VOCABULARY RESEARCH ..	29

<b>3 VOCABULARY LEARNING STRATEGIES .....</b>	<b>31</b>
3.1 DEFINING LANGUAGE LEARNING STRATEGIES .....	31
3.2 DEFINING VOCABULARY LEARNING STRATEGIES .....	34
3.3 RESEARCH ON LANGUAGE LEARNING STRATEGIES USED BY YLS .....	39
3.4 RESEARCH ON VOCABULARY LEARNING STRATEGIES USED BY YLS .....	42
3.5 DEVELOPMENT OF THE VOCABULARY LEARNING STRATEGIES QUESTIONNAIRE .....	45
3.6 METHODOLOGY .....	47
3.7 PARTICIPANTS AND PROCEDURE .....	47
3.8 INSTRUMENT .....	47
3.9 PROCEDURE .....	48
3.10 RESULTS AND DISCUSSION .....	48
3.11 CONCLUSIONS AND FURTHER INSTRUMENT DEVELOPMENT ....	53
<b>4 VALIDATING THE INSTRUMENT ASSESSING VOCABULARY .....</b>	<b>59</b>
4.1 METHODS .....	61
4.1.1 INSTRUMENT .....	61
4.1.2 PARTICIPANTS AND PROCEDURES .....	64
4.2 RESULTS .....	64
4.3 FREQUENCIES OF SCORE RANGES .....	68
4.4 CORRELATIONS ACROSS TASKS IN THE VOCABULARY TEST .....	70
4.5 DISCUSSION .....	72
4.6 CONCLUSION .....	74
<b>5 PHRASAL VERBS.....</b>	<b>75</b>
5.1 DEVELOPING AN INSTRUMENT ASSESSING YLS' KNOWLEDGE OF PHRASAL VERBS .....	75

5.2 PHRASAL VERB FREQUENCY LISTS .....	76
5.3 SELECTING THE ITEMS FOR THE COPHAVE TEST FOR YLS .....	78
5.4 CREATING THE COPHAVE TEST FOR YLS .....	82
5.5 DISCUSSION AND FUTURE RESEARCH .....	87
<b>END-NOTE .....</b>	<b>89</b>
<b>REFERENCES .....</b>	<b>101</b>

# INTRODUCTION

The domain of foreign language (FL) vocabulary assessment has experienced numerous research in the past years. However, hardly any attention has been focused on FL vocabulary assessment of young learners (YLS). In this book, English as a foreign language (EFL) vocabulary is investigated. The book is divided into five parts. In the first chapter, the literature on the assessment of individual words is revealed. In the second chapter the results of corpus linguistics are presented followed by chapter 3 focussing on word learning strategies. Chapter 4 is dedicated to a validation study of the assessment of word knowledge. The book is closed down by the presentation of the development of a tool assessing the knowledge of phrasal verbs.

## BRIEF OVERVIEW

The 1990s saw an increasing number of studies focusing on FL vocabulary learning and the literature has been growing ever since then by extending the knowledge on such areas as FL vocabulary assessment (Laufer, Elder, Congdon, & Hill, 2004; Nation, 2001; Schmitt, 1997), the FL mental lexicon (Singleton, 1999; Zareva, 2007), corpus studies (Horváth, 2001; Kilgarriff, 1997; Nation & Macalister, 2010) and vocabulary learning strategies (Nation, 1990; Schmitt, 2000). It has also been affirmed that vocabulary knowledge is a good predictor of reading comprehension (Henrikssen, Albrechtsen & Haastrup, 2004; Nassaji, 2003; Shiotsu & Weir, 2007) and general language proficiency (Boers, Demecheleer, Coxhead, & Webb, 2014; Henriksen, 1999; Zimmerman, 2004). It has been reported that receptive vocabulary knowledge predicts productive vocabulary knowledge (Laufer & Nation, 1999, p. 42). With the emergence of the lexical approach (Lewis, 1993) in language teaching, a new avenue was opened for vocabulary research. An expert on language teaching uncompromisingly concludes “*Without grammar very little can be conveyed, without vocabulary nothing can be conveyed*” (Thornbury, 2002, p. 18).

The learning and teaching of vocabulary is a popular research area in the FL learning literature. These two processes are in the center of attention of both scholars and teachers. Educators have been encouraged (Lewis, 1993; Thornbury, 2002) to promote intentional learning of words in the classroom. Since the early 1990s textbook and FL syllabus writers have laid special emphasis on integrating the results of vocabulary research into the curriculum (Fitzpatrick, Al-Qarni & Meara, 2008) since successful language learning is greatly determined by FL word knowledge (Schoonen & Verhallen, 2008).

The field of vocabulary and word knowledge has been investigated by several actors in scientific domains. Education researchers (Nagy, 2004), psycholinguists (Ellis & Beaton, 1997), neurolinguists (Paradis, 2004), and morphologists (Jackson & Zé Amvela, 2011) contribute to or exploit the empirical results of vocabulary learning research and assessment. In this chapter an insight is also given into the domain of vocabulary and it is stated what motivation and purpose have inspired me to conduct the research that not only assesses EFL of Hungarian young learners’ (YLS) vocabulary but it also seeks to explore the VLS used by the YLS.

Vocabulary is also considered as one of the strongest predictors of FL proficiency (Schmitt, 2008, p. 352). Significant correlations have been underlined between receptive vocabulary knowledge and FL reading comprehension (Henriksen et al., 2004; Koda, 1989; Laufer & Ravenhorst-Kalovski, 2010; Qian, 2002; Schmitt, Jiang & Grabe, 2011; Stæhr, 2009).

# 1 INDIVIDUAL WORDS

## 1.1 ASSESSING VOCABULARY

The focus of this section is the versatile nature of foreign language vocabulary assessment. Even though vocabulary assessment might be versatile, the same criteria apply to testing vocabulary as to other domains of language. It is posited by Nation (2001, p. 36) that reliability, validity, practicality and washback need to be considered when designing vocabulary tests. Washback is defined by Schiffko (2001, p. 832) as a positive or negative effect that derives from the classroom material judgment of students whether this material will appear on a subsequent test or not.

It is asserted (Read, 2000, p. 32) that there are two contrasting perspectives of vocabulary assessment. One viewpoint is that vocabulary items can be tested as a semantic field independent of context. The other view is that lexis must always be measured in context. Having pointed out these two mutually debunking points of view concerning vocabulary assessment, Read outlines the three dimensions of vocabulary testing. The dimensions elaborated on by Read are presented (Table 1).

TABLE 1. *Dichotomies of vocabulary assessment* (Read, 2000, p. 9)

---

<b>Discrete</b> A measure of vocabulary knowledge as an independent construct	<b>Embedded</b> A measure of vocabulary which forms part of the assessment of some other, larger construct
<b>Selective</b> A measure in which specific vocabulary items are the focus of the assessment	<b>Comprehensive</b> A measure which takes account of the whole vocabulary content of the input material
<b>Context-independent</b> A vocabulary measure in which the test-taker can produce the expected response without referring to any context	<b>Context-dependent</b> A vocabulary measure which assesses the test takers ability to take account of contextual information in order to produce the expected response

---

The first dimension, the discrete – embedded one, is elaborated on by Read (2000, p. 10) in the following way: discrete test items mean that they are separated from other components of language knowledge. Embedded vocabulary measure is one that has a contribution to assessment of a larger construct. An instance of this is reading comprehension questions following a text. Vocabulary items are not evaluated separately but simply form part of the measure of the learners' reading ability skills.

The second dimension, the selective – comprehensive one refers the range of vocabulary involved in the assessment. An instance of selective vocabulary measure is a test in which the target words are selected individually and then incorporated into the text. As opposed to the selective vocabulary measure, the comprehensive measure takes account of all the vocabulary content in the text.

The third dimension, the context-independent – context-dependent one is an old issue in vocabulary testing. The major distinction between context-dependency and independency is determined whether the test-taker has to rely on the text to induce the meaning of the vocabulary item or if they can just take the vocabulary test as if the different items were separate and in isolation.

It is very clear from the above dimensions outlined by Read that major distinctions must be made between kinds of approach towards assessing vocabulary. These issues should be the concern of test-givers. The issues emerging from language testing research need to be looked at from four different angles (Lehmann, 2009; Milton, 2011; Nation, 2013). Four major questions are proposed by Nation (2013) that need to be addressed: (1) why to test vocabulary? (2) what words to test? (3) what aspects of word knowledge to test? (4) how to test the various aspects of word knowledge?

The following questions are posed by Nation (2013, p. 171) when approaching the concept of vocabulary measurement: (1) what kind of vocabulary test is the best? (2) is it enough to ask learners if they know the word?, (3) should choices be given?, (4) should translations be used?, (5) should words be tested in context?, (6) how can we measure words that students do not know well?, (7) how can we measure the total vocabulary size? These questions are centered around the main issues of vocabulary testing by researchers: vocabulary in context, the way vocabulary is tested, the aspects of vocabulary knowledge and the selection of words that need to be tested. Different concepts of vocabulary assessment (aspects of knowing a word, word form, word meaning) will be in our focal point. After the explicit definition of these concepts, the major validated tests assessing vocabulary will be presented.

## 1.2 ASPECTS OF KNOWING A WORD IN A FOREIGN LANGUAGE

Knowing a word is complex and multidimensional in nature. Various aspects of knowing a word must be considered. As it was laid down in the previous section, breadth of vocabulary knowledge means how many words a person knows while depth refers to the knowledge of dimensions, e.g. synonyms, antonyms, contextual use, etc. The complexity of the concept of knowing a word is emphasized by Mukarto (2005, p. 153) who declares that *“learning even one FL word or a lexical item is a complex task. Naturally, learners’ knowledge of a word is not binary in nature, nor is it an all or nothing phenomenon.”*

Several dimensions have been identified that inform researchers and teachers how complex it is to determine what it means to know a word. When considering YLs, three facts are worth keeping in mind: (1) word knowledge is incremental, which implies multiple oral and written inputs (Nagy, Anderson & Herman, 1987, p. 238); (2) word knowledge is also multidimensional since a lot of words have different meanings, and (3) word knowledge is interrelated in that the knowledge of one lexical item is connected to another (Scott & de la Fuente 2008, p. 108).

According to Richards (1976, p. 77), knowing a word means (1) knowing its frequency and collocation, (2) knowing limitations of the use of that word, (3) knowing the syntax of this word; (4) knowing its basic forms and derivations, (5) knowing its associations with other words, (6) knowing its semantic value, (7) knowing many of the different meanings associated with the word.

Nagy and Scott (2000, p. 278) identified several new dimensions that describe the complexity of what it means to know a word. First, word knowledge is incremental, which involves many encounters with both spoken and written words in varying contexts (Nagy et al., 1987). Second, word knowledge is multidimensional because many words have multiple meanings and serve different functions in different contexts. Third, word knowledge is interrelated in that knowledge of one word connects to knowledge of other words.

These assumptions of word knowledge by Richards (1976) were later adopted (Nation, 2001). Form, position, function and meaning are the categories used by the scholar. Moreover, eight types of word knowledge are proposed to be considered: spoken form, written form, grammatical position, collocation, frequency, appropriateness, concept, and association. The components are presented in Table 2.

**TABLE 2.** *Components of word knowledge* (Nation, 2001, p. 16)

---

<b>Component</b>	<b>Receptive knowledge</b>	<b>Productive knowledge</b>
Spoken form	What does the word sound like?	How is the word pronounced?
Written form	What does the word look like?	How is the word written or spelt?
Grammatical position	In what patterns does the word occur?	In what patterns must we use the word?
Collocation	What words or types of words must we use with this word?	What words and types of words can we express before and after the word?
Frequency	How common is the word?	How often should the word be used?
Appropriateness	Where would we expect to find this word?	Where can this word be used?
Concept	What does the word mean?	What word should be used to express this meaning?
Association	What other words does this word make us think of?	What other words could we use instead of this one?

---

Bogaards (2000, p. 146) further claims that FL learners may learn the subsequent dimensions: form (spoken and written), meaning, morphology, syntax, collocates and discourse. As it is a clarified system of categories I will use Bogaards' categorization to elaborate on word form and meaning in the next section.

### 1.3 WORD FORM AND MEANING

Even though meaning has been considered the most important aspect of knowing a word, the notion of word form is gaining significance since evidence has proved that in the process of foreign language word processing. The knowing of the word form used to be considered a lower-level type of knowledge (Laufer et al., 2004); however, it has become an important feature in vocabulary learning.

Reading research has made advantageous contributions to indicating the importance of the orthographical word form. One study (Huckin & Bloch, 1993) entailed that orthographical similarity can mislead students in their guess of the meaning of the words. Readers mistook unknown (spooky) words for known (spoon) words that resembled one another. Notwithstanding the fact that the shape of the word and the visual features of the word can have a supporting effect, applying them in the recognition process has not proved to be the most effective way of learning foreign language vocabulary. In English as a FL research the bathtub effect is oft-cited. The bathtub is a visual metaphor indicating that the most remembered word parts are the beginnings and the endings of a word. This concept is a strong effect in English language; nevertheless it might not hold for other languages.

It is suggested by Goldstein (2004, p. 98) that being phonologically aware is important from the point of view of general vocabulary learning. It means that lower-level FL speakers need to rely more heavily on acoustic rules than native speakers. For instance, a native speaker would never mistake the word 'cub' for 'hub' since they can infer the word from context unlike non-natives who is likely to have a difficulty in interpreting the context.

Laypersons tend to believe that meaning of words equals definitions in dictionaries; nevertheless, the issue of meaning is more complex than that. Schmitt (2008, p. 82) argues that at the most basic level, meaning is the relationship between a word and its referent. Drawing on Bloomfield (1933), Drum and Konopak (1987, p. 77) emphasize that the relationship between the word and the referent is arbitrary.

The big white bear that lives in China and eats bamboo could be named 'napkin', a 'winner', or a 'melmel'. It is only common consensus that the label for this animal is 'panda' and it is this label that yields meaningful sense to this word. Words are usually labels for concepts which themselves involve our limited personal experience of the actual world reality. From the point of view of knowing the word, the notion

of meaning has a significant role. Most teachers and researchers consider a word learned if the form and the meaning are known. The first step towards foreign language vocabulary learning is the form-meaning link. This linkage has been studied in previous research from a number of aspects. Grainger and Dijkstra (1992) studied lexical neighbors. This notion 'lexical neighbor' is defined as words looking very similar in form but having a totally different meaning (e.g., pot, dot, cot, lot). They uncovered that learners are confused to distinguish among their meanings.

Having surveyed the aspects of knowing a word, I need to discuss one of the most researched areas of foreign language vocabulary: the form and meaning of vocabulary. In the next section, it will be looked at how it is possible to measure how deep and how broad learners' vocabularies are.

## 1.4 FOREIGN LANGUAGE VOCABULARY TESTS

Ever since vocabulary came into the focus of foreign language learning studies, assessment of word knowledge has been perceived as a fundamental issue in the research of this domain. This chapter provides an insight into how vocabulary is assessed and what types of validated and reliable instruments exist in the literature. Apart from presenting these instruments, special attention will be drawn to (1) the computerized versions of these instruments as in the 21<sup>st</sup> century diagnostic assessment is predominantly executed in an online environment (Laufer et al., 2004) and (2) whether the data collection instruments to be discussed have versions designed for YLs. It must also be highlighted that there is a consensus among scholars in foreign language vocabulary assessment that various modalities (see more in Laufer et al., 2004) of item assessment exist. Laufer et al. (2004, p. 218) claim that words may be measured from two perspectives: (1) the form-focused perspective that implies that the test-taker is able to retrieve the form of the word evidencing productive knowledge, (2) the meaning-focused perspective that entails the test-taker can retrieve the meaning of the word evidencing receptive knowledge. Laufer et al. (2004) refer to the productive-receptive dichotomy as active-passive knowledge. Four degrees of knowledge of meaning are discerned, on the basis of two dichotomous distinctions: providing the form for a given concept vs. providing the meaning for a given form; and recall vs. recognition (of form or meaning). These distinctions entail the following four modalities constituting a hierarchy

of difficulty: (1) passive recognition that encapsulates recognizing an item in e.g. a multiple choice test, (2) active recognition that encompasses a given definition and four items; in this modality the definition must be matched with the pertaining item, (3) passive recall that incorporates a sentence and the synonym of one item in the sentence must be given by the test-takers, and (4) active recall that comprises a description of items and the initial letters of the items are provided; test-takers are expected to produce the word. In a review article Schmitt (2014, p. 921) uses different terms for the same concepts. Passive recognition is termed meaning recognition; active recognition is named form recognition whereas passive recall is termed meaning recall and active recall is called form recall. In an attempt to provide instances, sample tasks will be given subsequently. In this book Schmitt's (2014) terminology will be utilized since the passive-active dichotomy is rather obsolete in the current literature (Webb, 2008, p. 82). Table 3 presents sample tasks of each of the four modalities.

TABLE 3. *Sample tasks of the four modalities*

<b>Recognition</b>		<b>Recall</b>
<b>Meaning</b>	<p><i>Instruction: Circle the equivalent of 'large'.</i></p> <p style="text-align: center;">large</p> <p>a) small                      b) tiny c) huge                         d) weak</p>	<p><i>Instruction: Finish the sentence with adequate words.</i></p> <p>When something is large, it is .....</p>
<b>Form</b>	<p><i>Instruction: Circle the word that suits the definition.</i></p> <p>something that is very big in size</p> <p>a) large                        b) angry c) hungry                      d) strong</p>	<p><i>Instruction: Finish the sentence with a word that starts with the given letter.</i></p> <p>Something very big in size l.....</p>

### 1.4.1 The Receptive Vocabulary Levels Test

The Receptive Vocabulary Levels Test is simply referred in the literature to as Vocabulary Levels Test (VLT). It operates with a discrete point measure. It requires meaning recognition. The test was developed by Nation (1990) and it was validated by Schmitt, Schmitt and Clapham (2001). Words are selected from such corpora as British National Corpus (Kilgarriff, 1997) and the CANCODE (Cambridge and Nottingham Corpus of Discourse in English) up to five levels: the first 2,000, 3,000, 5,000 and 10,000 most frequent words. These levels bear importance from a research-based perspective. The 2,000–3,000 levels contain high-frequency words whose knowledge is necessary for everyday communication. The 5,000 level is the minimal size which learners can conceive authentic texts. The 10,000 level, contains the most common low-frequency words (Webb, 2010). The fifth level is not grounded on any corpus but includes items from the University Word List (Xue & Nation, 1984).

The test-taker sees six words on the left-hand side and three definitions or synonyms on the right-hand side. They are expected to match the right-hand side items with three of the six words on the left-hand side. This means that the task contains three distractors. In the entire test each level comprises six clusters of six words. Table 4 presents one sample task of the VLT.

**TABLE 4.** *Sample task of the VLT* (Schmitt, Schmitt & Clapham, 2001)

*Instruction: Match three of the words from 1) to 6) with three definitions A) – C).*

---

1) bitter	
2) independent	
3) lovely	A) very small
4) merry	B) beautiful
5) popular	C) liked by many people
6) slight	

---

Since the test gives estimates of vocabulary size at 5 levels, it can be applied for placement purposes and for diagnosis of vocabulary gaps. Four parallel test versions were developed. The criterion of the development of the test was that

the definitions are succinct; the test could be completed in the fastest possible time and with the appropriate arrangement of the possibility of blind guesses could be diminished. In the online version of the VLT the test-taker is expected to write the listed six words next to the three definitions. The evaluation of the test is automatically completed. With the modified version of the online test, Vocabulary Online Recognition Speed Test (VORST) the speed of word recognition can also be examined (Laufer & Nation, 2001, p.21).

A version of the VLT designed for YLs has also been developed. Jimenez Catalan and Terrazas Gallego (2008) used the YL version of the instrument with young Spanish YLs of English. They modified the word selection process by involving such low-frequency words as names of animals (e.g., 'lion', 'ostrich', 'tiger') that YLs might know better than high-frequency words used by adults (e.g., 'beer', 'office', 'wine'). The researchers reported that the YLs' version of the VLT proved to be a valid measure of vocabulary assessment.

### 1.4.2 Productive Vocabulary Levels Test

With regard to productive knowledge of vocabulary, Laufer and Nation (1995) developed an instrument that measures productive word knowledge. The test took its name after the VLT and the adjective 'productive' was added so that the type of test would be clearly discerned. The test requires form recall on part of the participants. Similarly to the Vocabulary Levels Test, the tasks are divided into frequency clusters: 2,000, 3,000, 5,000, 10,000. In this test sentences are seen by students. In each sentence only the first two or three initial letters of one word are provided. Students must write the missing part of the word. This test is originally named the Test of Controlled Productive Ability (TCPA), nowadays it is referred to as Productive Vocabulary Levels Test (PVLТ). A part of the instrument is presented in Table 5.

**TABLE 5.** *Productive Vocabulary Levels Test* (Laufer & Nation, 1999)

*Instruction: Complete the words by filling in the gaps with the proper letters*

---

He likes walking in the fo ..... because the trees are beautiful there.

He takes cr ..... and sugar in his coffee.

The actor took the st ..... to perform in the long-awaited play.

---

It is obvious from Table 4, that the sentences following one another are unrelated. The test format resembles a C-test to some extent. In the pilot study of the instrument it was reported by the researchers that the selection of the target words was determined with the aim of avoiding any ambiguity of the meaning of the words. Similar to Schmitt et al. (2001) four test versions were developed. It is worth noting that the test has been criticized from a construct validity point of view. It was pointed out by Read (2000, p. 66) that the instrument is unlikely to assess productive word knowledge. He argues that some of the items demand only recognition and some of them need more contextual clues than others, thus he is dubious whether the test assesses what it is meant to assess.

Abduallah, Puteh, Azizan, Hamdan & Saude (2013) used the PVLТ to assess the productive vocabulary of 480 ESL learners in Malaysia. The participants were 15 years old. Albeit they do not count as YLs, this study is the only one reporting on using the PVLТ as data gathering instrument with not adult learners. The online version of the PVLТ is found on Tom Cobb’s website: [www.lexтutor.ca](http://www.lexтutor.ca).

### 1.4.3 Vocabulary Knowledge Scale

A vocabulary measure which can serve the purpose of assessing depth of vocabulary is the Vocabulary Knowledge Scale (VKS) (Paribakht & Wechse, 1999). On the one hand, Schmitt (2008, p. 45) asserts this type of vocabulary measurement sheds light on what students know, rather than on what they do not know, by allowing them to indicate their partial knowledge of a lexical item. It may be more motivating than other types of tests. On the other hand, Schmitt (2010, p. 32) criticizes the instrument by claiming that defining depth can be executed with extreme difficulty. The format of this test is presented in Table 6. The scoring of the original test used by Paribakht and Wesche (1999) is presented in Table 7.

**TABLE 6.** *Vocabulary Knowledge Scale* (Paribakht and Wechse, 1999)

*Instruction: Indicate the level you know the word procrastinate.*

---

1) I don’t remember having seen this word before.
2) I have seen this word before, but I don’t know what it means.
3) I have seen this word before and I think it means .....
4) I know this word. It means .....
5) I can use this word in a sentence: .....

---

TABLE 7. *The original scoring system of the VKS (Paribakht and Wechse, 1999, p. 81)*

Self-report categories	Possible scores	Meaning of scores
I	→ 1	The word is not familiar at all.
II	→ 2	The word is familiar but its meaning is not known.
III	→ 3	A correct synonym or translation is given.
IV	→ 4	The word is used with semantic appropriateness in a sentence.
V	→ 5	The word is used with semantic appropriateness and grammatical accuracy in a sentence.

Lehmann (2009) modified the scoring system as it is reported in her doctoral dissertation based on the test developed by Goulden, Nation and Read (1990) which was originally developed for self-assessment. Instead of the highest possible score being 5, Lehmann (2009) assessed a word either as known or not known. Table 8 presents the scoring system modified by Lehmann (2009, p. 88).

TABLE 8. *The scoring system of the VKS modified by Lehmann (2009, p. 88)*

Self-report categories	Possible scores	Meaning of scores
I	→ 0	The word is not familiar at all.
II	→ 0	The word is familiar but its meaning is not known.
III	→ 1	A correct synonym or translation is given.
IV	→ 1	The word is used with semantic appropriateness in a sentence.
V	→ 1	The word is used with semantic appropriateness and grammatical accuracy in a sentence.

If a student reports the word is familiar but the meaning is not known, then it is worth no points. This instrument was applied by Lehmann (2009) for the assessment of university students as Paribakht and Wechse (2006) calibrated the VKS for this age group. The VKS has also been designed for YLs recently. The was used by Atay and Kurt (2006), Paribakht and Wechse (2006), and Jóhannsdóttir (2010) to assess YLs.

The online version of the test is also available on Tom Cobb’s website. I have no knowledge of any study that has ever used the online VKS, however it is an empirical question whether applying the online measure would change either the validity or the reliability of the test.

### 1.4.4 Vocabulary Size Test

The Vocabulary Size Test (VST) was developed and validated by Nation and Beglar (2007). It assesses the knowledge of the 14,000 most frequent English words. It implies the modality of meaning recognition similar to the VLT. One sentence is given in each task and one word is underlined in the sentence. Under the sentence four possible options are provided in a multiple choice format and the test-takers must settle upon which word is interchangeable with the underlined word. The test is available in online version and it renders the assessment of receptive vocabulary rapid and effective. It is a very similar test format to the one applied on the TOEFL test which is also taken in a computerized environment. Table 9 presents a sample task of the VST.

**TABLE 9.** *Sample task of the VST* (Nation & Beglar, 2007)

*Instruction: Choose one proper word from items a) – d) that best fits the word in bold*

---

DRIVE: He **drives** fast.

- a. swims
  - b. learns
  - c. throws a ball
  - d. uses a car
-

### 1.4.5 The X\_Lex Test

The X\_Lex Test is both a traditional and an online receptive vocabulary assessment tool whose origins stem from the 1980s. Meara and Buxton (1987) experimented with a Yes–No test that was named X\_Lex a few years later (Meara, 1992). The X\_Lex test includes a list of 50 words. Out of the 50 words 35 of the words are existing English words and 15 are non-existing items. The test-takers have to indicate whether they know the word or not by clicking on the right button. In case a non-existing word is indicated as unknown, the learner is penalized with minus points at the final evaluation. The test is available at [www.testyourvocab.com](http://www.testyourvocab.com) and as the final step anybody taking the test online can provide background data (number of years spent learning English, age, gender, etc.) as well; therefore the instrument looks into correlations between the final score and the given background variables.

### 1.4.6 Diagnostic online English and German receptive vocabulary size test for YLs

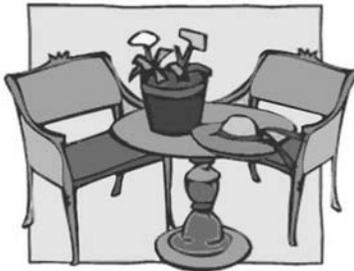
Most recently a FL vocabulary test has been developed and validated by the researchers of the University of Szeged (Vidákovich et al., 2013). The instrument is designed and calibrated to measure diagnostically the vocabulary size of 5<sup>th</sup> and 6<sup>th</sup> graders learning English and German as a FL. The selection of the target items was done on the basis of frequency lists and corpora and the test is unique in the sense that the words incorporated in the test are similar in the two languages. The instrument has a multiple choice test format in that the students see one picture and four words on the screen and they have to decide which word is described by the picture. Unlike the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) where only one word matches one picture in one task, in this test there is a likelihood that all four words match the picture or only one word can be matched with the picture; thus test-takers do not automatically exclude any correct item after solving one. The test-takers must click on the buttons next to each word and settle upon whether there is a match or not. The pictures are either simple or complex pictures and students must use identification or implication to figure out the correct answer. The test demands meaning recognition. The instrument has three versions in both languages. The instrument has been applied in an online environment on the eDia platform developed by the ICT specialists of the Institute of Educational Science at the University of Szeged (Molnár, 2013). The test-taking period is short as it takes

around ten minutes and apart from the test scores, background data can be processed immediately after the completion of the data collection instrument. Table 10 presents one task of the test.

**TABLE 10.** *Example of an item containing a simple picture (Vidákovich, Vígh, S. Hrebik, & Thékes, 2013)*

*Instruction: Choose from words a) – d) that best fit the picture on the left.*

---



- a) chair
  - b) plant
  - c) table
  - d) theatre
- 

### 1.4.7 Principles of designing FL vocabulary tests for YLs

Although most of the above-mentioned diagnostic data collection instruments have been originally designed to assess university students or adults there have been studies reporting on the testing of YLs' word knowledge as well. Diagnostic assessment of YLs' FL proficiency and word knowledge empowers teachers with a lot of classroom implications (McKay, 2006, p. 38).

It is typical of YLs that they use memorized chunks. Their knowledge is implicit in this sense; explicit learning ability that enables them to comprehend rules emerge around adolescence (Nikolov & Szabó, 2011, p. 32). Most of the YLs learn words rapidly (Orosz, 2009); nevertheless, after they are capable of recognizing words, the ability to use connotations, shades of meaning, synonyms and antonyms is only learned as a result of a long process of learning (Cameron, 2004, p. 32). Three fundamental facts have also been emphasized in the literature: (1) until the age of twelve students know only a limited (not more than 600–700) amount of words in an FL (Laufer, 1997, p. 143), (2) students hardly ever know the connotations (Schmitt, 2008, p. 352), and (3) YLs have limited awareness of the derivative forms of a word (Schmitt & Zimmerman, 2002, p. 160).

Before presenting the findings of studies assessing the word knowledge of YLs, I will elaborate on the characteristic traits and principles of diagnostic testing of FL in the context of YLs. Nikolov and Szabó (2011) outlined the principles of diagnostic testing of YLs (2011). These principles are based on the study by Alderson (2005). I will make an attempt to synthesize these principles which, I believe, are the most relevant from the perspective of vocabulary assessment of YLs.

- 1) the purpose of diagnostic tests is to identify the strengths and weaknesses of learners,
- 2) diagnostic tests must result in the treatment of difficulties arising during the learning process,
- 3) diagnostic tests must make it possible to analyze the score of each item in detail and to report the results; thus they provide feedback in detail and further steps can be taken,
- 4) diagnostic tests are low-stakes tests or bear no consequences so optimal achievement is not hindered by anxiety or any other affective factor,
- 5) diagnostic tests must take into consideration research on FL learning and in a wider sense the results of applied linguistics research,
- 6) diagnostic tests are more likely to be discreet-point tests than integrative, i.e., they focus on certain linguistic elements rather than on global abilities,
- 7) diagnostic tests are more likely to be less authentic than any other level-testing instruments;
- 8) diagnostic tests are more likely to focus on 'lower-level' linguistic abilities than on 'higher-level' abilities,
- 9) diagnostic tests assessing linguistic skills (listening, speaking, reading, writing) are more simple to develop than ones assessing grammar skills,
- 10) diagnostic testing is probably made more efficient by using a computerized platform.

Jang (2014), whose claims can also be subjected to FL vocabulary assessment, goes further by making the subsequent claims concerning the diagnostic assessment foreign language: assessment should (1) be cognitively rich enough to elicit knowledge and skills, (2) measure essential core skills, (3) promote positive learning and assessment experiences, (4) provide consistent and reliable information on proficiency, (5) promote students' ability to self-assess, (6) provide the support needed.

The two sets of principles laid out by Nikolov and Szabó (2011) and Jiang (2014) overlap in several points. It can be concluded that diagnostic vocabulary assessment of YLs ought to (1) give constant feedback, (2) provide information constantly to promote the learning process, and (3) remain low-stakes so that it can provide positive learning experiences.

When designing the diagnostic test Harding, Alderson and Brunfaut's (2015, p. 322) five recommendations were also taken into consideration: (1) it is not the test which diagnoses, it is the user of the test, (2) instruments, themselves should be designed to be user-friendly, targeted, and discrete in order to assist the teacher in making a diagnosis, (3) the diagnostic assessment process should take diverse stakeholder views into consideration, including learners' self-assessments, (4) diagnostic assessment ideally involves the diagnostic stages of listening/observation, initial assessment, use of tools, tests, expert help and decision-making, (5) diagnostic assessment should relate to some future treatment.

## 1.5 RESULTS OF DIAGNOSTIC TESTS ASSESSING YLS' VOCABULARY

Few studies have sought to explore the vocabulary size of YLs so far. In a study, the receptive vocabulary of Spanish 4<sup>th</sup> graders (n = 270) was diagnostically explored by Jiménez Catalan and Terrazas Gallego (2008). Students had learned English for three years in 3 lessons a week at the time of data collection. The VLT was utilized as the test up to the 2,000 most frequent words. The study discovered that the less frequent a word is the less chance students have of knowing it.

Atay and Kurt (2006) applied the VKS in a control-group treatment for the assessment of Turkish 6<sup>th</sup> graders (n = 62) in order to map their English word knowledge. The YLs' vocabulary development, elicited by post-reading activities, was measured. The researchers gave account of a well-functioning, reliable VKS test designed for YLs in this study. Schmitt (1998, p. 291) also confirmed the value of the instrument from the point of view of assessing YLs by stating that the VKS taps into the early stages of vocabulary learning.

Jóhannsdóttir (2010) also used the VKS to assess the vocabulary of 42 4th-graders in Iceland to map the vocabulary of the learners. Jóhannsdóttir (2010) had the learners take a Yes–No test on EFL words as well and was seeking to learn how reliable the vocabulary tests were and to find correlations among the two measures and motivation. The test proved to be of decent reliability (Cronbach's Alpha = .80)

and indicated significant correlations with the results scored on the Yes–No test and motivation. One of the major findings of the study was that students scored significantly higher points on the receptive test (Yes–No) than on the productive test measuring depth of word knowledge (VKS).

In Hungary receptive word knowledge was investigated by Orosz (2009, p. 184) using the paper-and-pencil X\_Lex test. Hungarian 3<sup>rd</sup>–6<sup>th</sup> graders (n = 253) took part in her study. The instrument contained 120 items with 100 real words selected from the 5,000 most frequent English words based on the BNC and with the addition of 20 non-existent words (e.g. ‘bable’, ‘lall’, ‘pentil’, ‘remlile’, etc.). By transforming the scores, the estimation was made that students knew 348 words in 3<sup>rd</sup> grade, 696 words in 4<sup>th</sup> grade, 1,177 in 5<sup>th</sup> grade and 1,457 in 6<sup>th</sup> grade.

In the pilot study 352 participants took the English test version (Vidákovich et al., 2013). The instrument proved to be robustly reliable (Cronbach’s Alpha = .91) and the test versions drew attention to strong relationships and significant correlations with one another. The instrument proved to distinguish well among the test takers. Relevant data were gained concerning the type of words high and low-achieving students know. High achievers know adjectives and verbs significantly better than low achievers whereas low achievers know significantly more nouns than any other word type. Table 11 presents all the relevant studies that have investigated YLs’ English as a FL word knowledge.

**TABLE 11.** *Studies investigating YLs’ EFL word knowledge*

<b>Study</b>	<b>Participants</b>	<b>Instrument</b>
Atay & Kurt (2006)	62 Turkish 6 <sup>th</sup> graders	Vocabulary Knowledge Scale
Jiménez Catalán & Terrazas Gallego (2008)	270 Spanish 4 <sup>th</sup> graders	Vocabulary Levels Test
Orosz (2009)	253 Hungarian 3 <sup>rd</sup> –6 <sup>th</sup> graders	X_Lex Test
Jóhannsdóttir (2010)	42 Icelandic 4 <sup>th</sup> graders	Vocabulary Knowledge Scale
Vidákovich et al. (2013)	127 Hungarian 6 <sup>th</sup> graders	Diagnostic Online English and German receptive vocabulary size test

Having given an overview of the domain of FL vocabulary assessment I discussed what it means to know a word form and what aspects of word knowledge must be tested. One of the fundamental tasks was not only to identify the main vocabulary tests but to choose the most applicable ones as numerous instruments exist. However, I have decided to select six of these tests as I have found them relevant from the point of view of our diagnostic vocabulary test development. The criteria of choosing these tests were the following: (1) they are validated and were proved to be reliable, (2) they preferably have a version adapted for YLs, (3) they are preferably computerized. Even though these tests are valid measure of word knowledge they have something in common: they lack the characteristic of measuring the interconnected aspects of word knowledge. They test only one construct: either receptive or productive word knowledge. I reckon a test that integrates all elements of word knowledge is necessary to map the vocabulary of learners. The notion of a multiple tests approach is supported by several scholars (Ishii & Schmitt, 2009; Laufer & Nation, 1999). It is argued that a more comprehensive picture of vocabulary knowledge ought to be provided.

## 2 CORPUS LINGUISTICS

### 2.1 THE USE OF CORPORA IN VOCABULARY RESEARCH

Corpus linguistics is a rapidly developing field of applied linguistics. A large amount of corpora are being developed all over the world for a lot of languages and for a lot of jargons also. For instance, one can see corpora of car mechanics jargon, spoken Scottish English jargon, etc. (Poplack, 1989). The application of corpora is a major empowerment not only for vocabulary learning and teaching researchers but for language teaching practitioners as well. Schmitt (2008, p. 42) underlines that it is unimaginable that any domain of research into vocabulary teaching, assessment or vocabulary syllabus design would do without the valuable information provided by corpus linguistics. While the compilation of different corpora had been a gigantic and imprecise effort before computers, nowadays exact data can be gathered with relative efficiency. This efficiency is really relative since the common endeavor of Cambridge University and the University of Nottingham, the CANCODE spoken corpus of British English took eight years to finalize by transcribing and coding five million words. Besides numerous English corpora, an attempt have been made to assemble corpora in most languages. With regard to Hungarian, Lengyelne (2006) sheds sufficient light on the status of Hungarian and other national corpora.

### 2.2 CORPORA AND THEIR DEVELOPMENT

The earliest corpora began appearing in the 1920s. It is hard to imagine how tedious it was to manually count the lexical items. Corpora comprising one million words were an extremely large number. From the 1960s on computers were utilized to assimilate corpora. The Brown University Corpus (Kucera & Francis, 1967) and Lancaster-Oslo/Bergen (LOB) Corpus (Hofland & Johansson, 1982) were two adequate instances of an attempt for collected corpora. From the 1990s on, the third-generation, as Schmitt (2008) calls them, of corpora has brought a large amount of development in quality and quantity.

Quantity in corpora is, nevertheless, not the only indicator of a good corpus. What goes into the corpus is also an important issue. It is claimed by Nation and Waring (1997, p. 12) that not all the words are equally worth knowing. To measure

the usefulness of a word, its ratio of occurrence also needs to be taken into consideration. Frequency is the simplest piece of information that can be retrieved from different corpora. How frequently a word occurs can determine the way textbook writers put together the syllabus if we approach the field from a teaching point of view.

Frequency is the most underlying concept that is examined in corpus linguistics. The most elementary thing that can be deduced from studying the language in a corpus is how many times a particular word occurs. The earliest corpora in research gave the frequency of a word as the first piece of information to researchers.

The General Service List (West, 1953) and University Word List (Xue & Nation, 1984) were outlined with the aim of measuring lexical richness in a new manner. This profile was then called the Lexical Frequency Profile (LFP) and was developed by Laufer and Nation (1995). The authors claimed they had the intention to exclude subjective judgments in the assessment of writing quality and it was efficient in the measurement of how vocabulary size was reflected in use. Laufer (2001, p. 248) also claims that the profile has no topic dependency and is a reliable measure as long as the topic is general and the writers are not required to apply any jargon. Jargon, in its nature, implies that a large amount of low-frequency words are used. The profile validated by these two researchers is not suited for assessment of any specific jargon. Nation also developed a software for Windows-based computers and it is named RANGE.

### 2.3 WIDELY USED CORPORA

In this section several corpora are presented from various perspectives such as an insight into their original purpose, their contents, and their area of usage. Five different corpora will be listed: (1) British National Corpus (BNC), (2) Contemporary Corpus of American English (COCA), (3) American National Corpus (ANC), (4) Cambridge and Nottingham Corpus of Discourse in English (CANCODE), and (5) the Child Language Data Exchange System (CHILDES).

Since the BNC and the COCA were considered for use in the selection of words during test development, I elaborate on these corpora. The BNC is available at <http://www.natcorp.ox.ac.uk/>. The development process of the BNC was published by Kilgarriff (1997). This corpus was launched in the early 1990s and was applied as a basis of vocabulary assessment to a large extent. Work began in 1991 and the first version was available for public use in 1994. It is the most cited corpus in word knowledge assessment and it is generally the basis of word selection in the development

of diagnostic vocabulary texts. The BNC is considered as a main source for anybody involved in language teaching. It contains more than 100 million words and has a large part of spoken corpus.

The COCA is available at <http://corpus.byu.edu/coca/>. It is claimed to be the largest freely-available corpus of English on its website. This corpus was developed and is under constant construction by Mark Davies of Brigham Young University. It is described as the first large, genre-balanced corpus of any language (Davies, 2010). The latest update was made in 2012. It contains 400 million words and is used broadly by researchers, linguists, teachers and translators. On the COCA website the user has to simply type in the word and the site generates all the necessary information (frequency, frequency rank) in less than a second.

## 2.4 APPLICATIONS OF CORPORA IN FL VOCABULARY RESEARCH

As was noted in this chapter, no vocabulary test development can do without the use of some kind of corpus. The items for all the major vocabulary tests listed in this chapter are selected from corpora. The selection of vocabulary for the Vocabulary Levels Test took place with the application of the BNC. The different frequency levels are determined by the BNC. The Productive Vocabulary Levels Test was also implemented on the basis of the BNC, whereas the selection of the words for the Vocabulary Knowledge Scale is based on the COCA.

In the YLs' vocabulary test development by a Hungarian research group (Vidákovich et al., 2013) both the British National Corpus and the Contemporary Corpus of American English were used. The researchers reported that the application of the COCA was a reasonable choice due to its larger size. Upon determining the list of words encompassed in the test the two corpora, the BNC and the COCA were compared. Two word lists were conceived and it was disambiguous that only minor difference existed between them.

Poole (2011) used the Vocabulary Knowledge Scale to assess the depth of word knowledge of university students and he gave an account of using the COCA for the selection of words. As cited in this chapter previously, Jimenez Catalan and Terrazas Gallego (2008) used the Vocabulary Levels Test to assess the vocabulary of Spanish YLs. They selected the lexical items from the BNC. Nation (2012) used both the BNC and the COCA for his Vocabulary Size Test. When examining the methods of determining what corpus to opt for the item selection for testing

instruments, it appears that the BNC and the COCA are the two corpora that researchers preferably apply for their work.

In this chapter I have emphasized the importance of the findings of corpus linguistics. I reasoned that no diagnostic FL vocabulary test can be developed without using corpora. It was pointed out that the item selection for the major validated vocabulary tests is corpus-based. The most important corpora have been listed and described in details. Important information can be gained from corpus linguistics with regard to frequencies. The service corpus linguistics can offer to foreign language vocabulary teaching and learning is enormous.

### 3 VOCABULARY LEARNING STRATEGIES

It has been argued in this book that research in the domain of vocabulary gained impetus in the 1980s. The popularity of word knowledge assessment has increased in the past 30 years. Even though the construct of learning strategies is another significant and well-researched domain in educational science, the intersecting construct, VLS are rather under-researched (Schmitt, 2000, p. 44). Hereby an attempt is made to fill in this gap. Vocabulary learning strategy research is important for two reasons: (1) the processes of language learning can be identified (Cohen 2003, p. 279); and (2) awareness of the strategies learners apply has enormous classroom implications since with strategy training the learning process can be made more efficient (Schouten-Van Parreren, 1992, p. 98). In this chapter I will give an outline of the various findings in the literature on VLS and present the empirical data that are relevant in terms of YLs' strategy use and training. First I will review the literature of the domain of language learning strategies (LLS), then I will seek to find a definition to VLS, finally I will synthesize previous research of YLs' VLS.

#### 3.1 DEFINING LANGUAGE LEARNING STRATEGIES

The past more than three decades have seen an enormous number of research into FL learning strategies. It must also be emphasized that at the outset of strategy research, strategies were thought of as conscious processes whereas nowadays they are considered semi-conscious operations (Cohen, 1990, p. 30). Semi-conscious operation means that the learners is not fully aware of their strategy use. Data on LLS can be gathered through self-reporting methods that might include interviews, written diaries and think-aloud protocols. Cohen also focuses on the conscious procedures by asserting that LLS are "*the conscious or semi-conscious thoughts and behaviors used by learners with the explicit goal of improving their knowledge and comprehension of a target language*" (Cohen 2003, p. 280). According to him, conscious thought is the intentional utilization of techniques whereas semi-conscious thinking encapsulates automatized, routine actions on the part of the learners.

As for the most amenable strategy to vocabulary learning, it is unanimously claimed in the literature that the most successful language learners do not use a great deal of strategies but they use only few of them, which might be only two

or three in number (Chamot, 2005; Cohen, 2003; Doró & Habók, 2013; Oxford, 1991). Cohen (2003, p. 282) argues from a perspective focusing on tasks that no single strategy will be amenable for all students or for all tasks, and students will apply the very same strategies in different ways.

During the past three decades in research drawing attention to FL learning and discussions on SLA theory, the emphasis has shifted from universal processes to the role of individual differences including cognitive and affective features. That is why light has been shed on the research of LLS (Chamot, 2005). In spite of the enhanced interest in LLS, defining strategies is still dubious (Doró & Habók, 2013; McDonough, 1999;) since several, sometimes contradictory, perspectives must be taken into consideration.

As for the theoretical background to strategy research, two major models have been used: (1) the information processing model of cognitive psychology (Bialystok, 1990); and (2) the communicative knowledge model of language knowledge whose executing components make it possible for learners to achieve their goals (Bachmann, 1990). Strategic knowledge is composed of metacognitive strategies whose executing functions enable the language learner to set goals, to evaluate and to plan (Bachmann, & Palmer, 1996).

Weinstein and Mayer (1986, p. 320) define strategies from a behavioral perspective by stipulating that *“learning strategies can be defined as the behaviors and thoughts that a learner engages in during learning and that are intended to foster the learners’ encoding process.”* O’Malley and Chamot (1990, p. 1) define LLS as *“special thoughts or behaviors that individuals use to assist them comprehend, learn, or retain new information”*. They discern three main types of strategies: metacognitive, cognitive and socio-affective. Their focus of attention is mostly on metacognitive strategies. LLS are defined by Ellis (1994, p. 226) who states that a strategy as a behavioral or mental activity in conjunction with some specific stages in language learning and to the process of language use. According to Ridley (1997, p. 231) strategies denote procedures which operate consciously or unconsciously in order to reach some kind of goal. Taking all the definitions of LLS into consideration, I regard Ellis’ concise definition the most applicable one; he says that strategy is a behavioral or mental activity related to some specific stages in language learning and to the process of language use. There are two reasons for this: (1) strategy must be looked at as a behavioral activity on the one hand; (2) on the other hand, it is also a cognitive activity during which learners want to learn new information.

A composite construct was developed by Oxford (1991) who classified strategies in a most comprehensive way. The dimensions of her Strategy Inventory of Language

Learning (SILL) are as follows: memory, cognitive, compensatory, metacognitive, affective and social. She defined LLS as “*operations employed by the learner to aid the learning, storage, retrieval, and use of information...; specific actions taken by the learners to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations*” (Oxford, 1991, p. 8). Three direct strategies and three indirect strategies were identified: the direct ones were memory, cognitive and compensation; the indirect ones were metacognitive, affective and social. It must be remarked that this type of classification is not supported by research.

According to Oxford (1991, p. 43), cognitive strategies involve the function of “*manipulation of the target language by the learner*”. Metacognitive strategies are related to a conscious overview of the learning process: planning, monitoring or evaluating. Memory strategies comprise the linking of new material to already existing knowledge. Social strategies involve interaction with peers and the teacher to track down the meaning of unknown vocabulary items. Compensation strategies comprise the use of synonyms, circumlocution, NL equivalents and guessing meaning. Finally affective strategies mean the reduction of anxiety and applying self-encouragement.

Nisbet, Tindall and Arroyo (2005, p. 105) imply that the SILL measures self-report behavior but it fails to measure autonomy. Their assertion was meant to include both adult and YLs. It is worth noting at this point that any questionnaire focusing on learning strategies suffers from this difficulty, namely what the students say they use as a strategy might not be in line with what strategy they actually use.

The large number of definitions in the literature are summed up in five points by Cohen and Macaro (2007, p. 27) who seek to determine a common intersection of the previously made claims:

- 1) the strategies that learners use can be documented;
- 2) a strategy is a construct that can be defined, and what it is and what it does can be described in practical terms;
- 3) strategies are important because they are associated with successful learning;
- 4) some learner types are more likely to use strategies or use them more successfully than other learner types;
- 5) strategies can be taught and learners, as a result, can develop more effective strategic behavior. As a consequence consciousness plays a major role in effective strategy use.

### 3.2 DEFINING VOCABULARY LEARNING STRATEGIES

The taxonomies of LLS differ in several domains but their most important features align to a great extent. The past 20 years have seen a considerable increase in studies on vocabulary learning and strategic thinking in FL learning. Yet, their intersection, strategies in vocabulary learning, has not attracted sufficient attention. Schmitt (2008) synthesized the literature on the topic. Several researchers had established categories and dimensions with regard to learning strategies. He also pointed out that there were several overlaps among the different strategies. In order to gain a clear insight into what different scholars consider the components of VLS, I have gathered all the relevant taxonomies. Table 12 presents six vocabulary learning taxonomies.

**TABLE 12.** *Taxonomies of VLS*

---

<b>Author</b>	<b>Taxonomies of VLS</b>
Nation (1990)	1) planning /choosing words, choosing the aspects of word knowledge, choosing strategies, planning repetitions/ 2) sources /analyzing the word, using context, consulting a reference source in NL and FL, using parallels in NL and FL/ 3) processes /noticing, retrieving, generating/
Gu & Johnson (1996)	1) metacognitive regulation 2) guessing strategies 3) dictionary strategies 4) note-taking strategies 5) memory strategies (rehearsal) 6) memory strategies (encoding) 7) activation strategies
Schmitt (1997)	1) discovery-determination 2) discovery-social 3) consolidation-social 4) consolidation-memory 5) consolidation-cognitive 6) consolidation-metacognitive

---

Author	Taxonomies of VLS
Lin (2001)	1) cognitive 2) metacognitive 3) affective-social strategies
Tseng, Dörnyei & Schmitt (2006)	1) self-regulatory commitment control 2) self-regulatory metacognitive control 3) self-regulatory satiation control 4) self-regulatory emotion control 5) self-regulatory environment control
Pavičič (2008)	1) strategies of formal vocabulary learning and practicing 2) self-initiated independent vocabulary learning 3) spontaneous (incidental) vocabulary learning
Schmitt (2008)	1) determination 2) social 3) memory 4) cognitive 5) metacognitive

Several scholars (Cohen, 1990; Nation, 1990; Oxford, 1991) gave a definition of VLS and determined their components. However, Nation (1990, p. 217) postulated that defining vocabulary learning strategy poses difficulty but a strategy is one that needs to encapsulate choice, be complex, require knowledge and benefit from training and increase the efficiency of vocabulary learning.

Nation's (1990) attempt was the first one in the field of vocabulary learning that had tapped into learning strategies. The subsequent general classification of strategies was established by Nation (1990): (1) planning, (2) sources: finding information about word, and (3) processes: establishing knowledge. Within these general classes of strategies he determined sub-types of strategies. According to Nation (1990), planning involves choosing words, choosing the aspects of word knowledge, choosing strategies and planning repetition. His second general class of strategies comprises analyzing the word, using context, consulting a reference source in NL and FL and using parallels in NL and FL, whereas his third general class consists of noticing, retrieving and generating words. It is worth noting,

nevertheless, that Nation's classification has never been validated with empirical data to the best of our knowledge.

A questionnaire was validated by Stoffer (1995) that bore the name: the vocabulary learning strategy inventory (VOLSI). Its item pool took vocabulary strategies into consideration based on nine dimensions. Stoffer (1995, p. 23) used his instrument with university students learning FLs and the VOLSI proved to be a reliable questionnaire (Cronbach's Alpha = .86). The most frequent strategy in Stoffer's study (1995) was linking to NL words similar in spelling and all in all, the fourth group of strategies (strategies for creating mental linkages) was indicated by the students as the most often used one. It was also revealed that learners who had previously received some kind of vocabulary learning strategy instruction used these strategies more frequently than those with no instructions whatsoever. The age of the language learners appeared to be significant on seven of the nine factors in that YLs tended to use fewer strategies than their older counterparts did. Gender differences, however, were not significant only by a small margin.

In the taxonomy outlined by Schmitt (1997) a new type of strategy was defined: determination strategies. Nine determination strategies were discerned in this taxonomy, a new variable compared to other instruments. According to this taxonomy, determination strategies facilitate the discovery of the meaning of new words through guessing from an NL cognate and from context, applying any kind of reference material and asking somebody. Cognates are words in two languages that take their origins in the same word family (Merriam-Webster, 2015). For instance, the word 'Vater' in German is a cognate of the English word 'father' or the word 'hamburger' has the same meaning in English as in Hungarian. Thus, it is quite simple for a German or a Hungarian learner of English to learn these items. Guessing from context has been a highly promoted method of learning words in the communicative era of language learning and instruction (Thornbury, 2004, p. 46). Guessing from context is likely to occur in different learning environments; nonetheless, it most commonly means inferring the meaning of an unknown word from its surrounding context. Bossers (1992, p. 251) claims that a substantial amount of the words that students learn occur through inferencing meaning from context; hence it can be asserted that contextual guessing is a major component of determination strategies. Out of the eight social strategies encompassed in Schmitt's taxonomy, five belong to the discovery-social and three to the consolidation-social dimension. The five discovery-social questionnaire items all inquire about students' asking their teachers or mates for help. Of the items, the 'Ask teacher for a synonym or paraphrase' is a common strategy amongst students in a classroom

environment (Schmitt, 1997, p. 202). Discovery-social strategies encapsulate requests for help whereas the consolidation-social strategies dimension involve interactions after the lexical item has been learned. As for consolidation-memory strategies, Schmitt (1997, p. 96) claims that these types of strategies traditionally known as mnemonics comprise the relating of already learned knowledge to the newly-learned words. Such strategies include imagery, an activity during which learners match words to pictures, relating words, normally synonyms to the new items (e.g., amazed-fascinated), grouping words in semantically related clusters and using physical action to memorize the lexical items.

The consolidation-cognitive strategies dimension consists of nine questionnaire items in Schmitt's instrument and they focus their attention on mechanical repetition and involve such traditional strategies as keeping a written vocabulary, writing word lists and using flash cards for the retention of words. The final dimension in the taxonomy is metacognitive strategies. When using metacognitive strategies, students evaluate their own learning process. Metacognitive and self-regulatory learning, two different kinds of strategies, are complex and interactive processes in which both motivation and self-regulatory activities play a role (Boakerts & Simon, 1995). Students learn some facts and some processes during the years that assist them remember something when it is necessary. Schmitt (1997, p. 224) was led to postulate that the most common metacognitive strategy is continuous studying of the same word over time.

In Schmitt's (1998) qualitative research with Japanese students of 14–40 years of age, he examined these categories thoroughly and came to the conclusion that the most frequently used ones were discovery strategies: using a bilingual dictionary, verbal repetition and guessing from textual context. Besides Schmitt's data gathering instrument, the vocabulary learning questionnaire (VLQ) compiled by Gu and Johnson (1996) is an instrument that examines different learner strategies in this sub-field of SLA. They distinguished three factors: (1) beliefs, (2) metacognitive strategies, and (3) cognitive strategies. Beliefs were not further fragmented into any other categories; however metacognitive strategies were split into self-initiation and selective attention. Cognitive strategies comprised initial handling, reinforcement and activation. The focus of their investigation was advanced learners of English. They ran a correlation study based on the data received from the questionnaire and students' scores on tests of vocabulary size. They intended to reveal what strategies went hand in hand with previous learning and they also aimed at finding out which clusters learners used. They found that self-initiation strategies and activation strategies correlated significantly with vocabulary size. Self-initiation

strategies were defined as ones involving the learner's autonomous decision to use any technique to learn a new word whereas activation was regarded as the intention to activate a strategy to learn a new word. Then, they distinguished five types of learners: (1) readers who dealt with words in context, (2) active strategy users who were hard working and motivated, (3) non-encoders who used no intentional memorization strategies, (4) encoders who used intentional memorization strategies, and (5) passive strategy users who hardly ever used any strategy.

Tseng, Dörnyei and Schmitt (2006) drew on work completed in educational psychology and focused on proposing a new psychometrically-based approach toward FL VLS. This construct of this new approach is grounded on the learners' self-regulatory capacity. It is a conceptual framework that highlights the learners' innate capacity which energizes their effort to personalize strategies efficient for them. Tseng et al. (2006, p. 98) claim that the underlying problem in strategy research is the diverse conceptualization of the notion. Determining the specific dimensions suited for specific age groups poses difficulty.

In his synthesis Schmitt (2008, p. 88) compiled a taxonomy of VLS by distinguishing five dimensions: (1) determination, (2) social, (3) memory, (4) cognitive, and (5) metacognitive. Schmitt (2008, p. 340) conceived two major factors when setting up his new taxonomy of VLS by creating two major factors: discovery and consolidation. Two sub-factors were added to the discovery factor: determination and social. The consolidation factor was widened with four sub-factors: social, memory, cognitive and metacognitive. It is worth noting that in spite of the fact that the labels of the dimensions in Nation's (1990) taxonomy differ from Schmitt's (2008), there is a considerable amount of overlap among the two researchers' dimensions. Nation's planning dimension aligns with Schmitt's determination and cognitive dimensions. Nation's dimension called 'sources' has an overlap with Schmitt's social factor to a great extent and the third dimension in the Nation taxonomy, processes, strongly aligns with Schmitt's memory and metacognitive dimensions.

Lin (2001, p.145) ran a case study with the participation of seven 15-year-old Taiwanese learners to investigate their VLS. Data gathering methods involved classroom observation, interviews and think-aloud protocols. More than 70 strategies were identified. These items were then identified as one of the three main types of strategies: cognitive, metacognitive and social-affective strategies. These types of strategies are identical in most of the instruments assessing VLS; nevertheless memory strategy and discovery strategy are not included in Lin's (2001) data collection instrument, which might question the validity of the instrument.

Jimenez Catalan (2003, p. 44) came to the conclusion that males and females differed in the use of strategies. She stipulated from her empirical data on a sample of 581 YLs (age = 11 years) of Basque (NL) and English as a FL that males and females both used similar strategies: using a bilingual dictionary, guessing from textual context, asking the teacher and saying the word out loud when studying. This finding corroborates Schmitt's (2008) results: discovery strategies are used more often than any other strategies.

In this section it has been discussed how vocabulary learning can be defined and the major components of the different instruments assessing strategy use have also been presented. The postulation has been made that there is an agreement among researchers in the literature that cognitive, metacognitive, social, affective and memory strategies are the main dimensions. Having revealed the construct of VLS in general, a YLs' perspective will be taken in the subsequent section.

### 3.3 RESEARCH ON LANGUAGE LEARNING STRATEGIES USED BY YLS

In this section the results of several studies will be presented with regard to YLs' LLS. Both international and Hungarian results stemming from data gathered with different instruments developed for YLs will be presented.

A key question of strategy research is the extent to which strategies contribute to the success of language learning. Firstly, it must be determined whether successful language learners use more strategies than less successful ones. Secondly, it must also be investigated whether a successful language learner uses more or fewer strategies. Another focus of language learning strategy research is related to the emergence of strategies. Chesterfield and Chesterfield (1985, p. 56) asserted that strategies emerge in a natural way. This finding is confirmed by Nikolov in her study involving young Hungarian EFL learners (Nikolov, 1999b, p. 228).

Gunning (1997) developed the SILL adapted for children (Children's SILL). The instrument comprises 30 items. The main classifications: memory, cognitive, compensation, metacognitive, affective and social strategies remained unchanged but the wording of the questionnaire items was altered, so YLs could easily comprehend it. In a study conducted with the application of the Children's SILL instrument it was concluded that YLs had a tendency of relying on compensation strategies to a great extent.

In a study involving adult and young language learners, the similarities and differences between the strategy uses of the two age groups were investigated by Pinter (2006). Students had to interact in pairs doing a 'Spot the difference' task. Pinter (2006, p. 624) was led to assert that adults controlled the task more effectively than YLs by using more cognitive and metacognitive strategies. Thus, it can be concluded that adult learners are likely to use strategies that are more empowering in the actual context. In another study (Nikolov, 2006) investigating YLs' strategy use whilst solving English as a FL tasks Hungarian 12-year-old children (n = 53) were examined with the method of think-aloud protocols in non-mediated verbalization (Gass, 1997). In this research Nikolov (2006, p. 38) uncovered that students used cognitive strategies in the majority of the cases, more specifically, translation. Besides translation, skimming and scanning the texts were also frequently used. The researcher also identified a lot of individual differences in the data received from think-aloud protocols. The assertion was also made that good performers did not necessarily use more strategies. In fact it was concluded that some of the high achievers in the high achievers did not report any strategy use. It was also noted, nevertheless, that some good performers used various types of strategies. It was claimed as a conclusion that high achievers might benefit more from strategy use and this finding confirmed previous research (see Purpura, 1991).

Nikolov (2003, p. 6) states that studies which investigate YLs have a broad variety of types of approach, an assumption that is confirmed by Szpotowicz & Szulc-Kurpaska (2012) and Mihaljevic Djigunovic (2010); nevertheless, it is also declared by Nikolov (2003, p. 6) that no significant correlation exists between ratio of occurrence in the use of strategies and language learning achievement.

The major factors of successful language learning by YLs have been examined in several studies. Apart from early exposure, attitude and motivation, one of the most important variable, was strategy use (Szpotowicz & Lidgren, 2011, p. 140). It was also pointed out that young language learners use similar communicative and cognitive strategies to adults. As part of a large-scale project called ELLiE (Enever, 2011) in which substantial data were collected on language use, classroom context, teaching style, the pace of learning and strategy use, Szpotowicz and Lindgren (2011) found that virtually all the students used transfer of words from their NL, which can be considered a cognitive strategy. In a study published by Szulc-Kurpaska (2000), eleven-year-old YLs' strategy use was examined in a communicative language game. When children faced dilemmas as far as language was regarded, they used formulaic chunks, formulated new, non-existing words, drew on their NL and turned to the teacher. The first three are cognitive strategies and the latter one is a social strategy.

In a study done with the participation of 61 5<sup>th</sup>- and 6<sup>th</sup> graders in Mexico as part of a large-scale English as a FL instruction program called National English Program for Basic Education (PNIEB), focus-group interviews were executed in order that conclusions could be drawn partly of YLs' strategy use. The researchers were seeking to find data concerning language learning outside the classroom (Sayer & Ban, 2014, p. 324). They revealed that children used numerous functions, sources and strategies to learn English. It was asserted that in the uses of English outside the classroom, sixteen distinct strategies were identified. Listening to popular songs, watching movies in English, playing video games, using the Internet and using Google Translate were among the most frequently used functions. Numerous students specifically reported that playing computer games, listening to present day celebrities are great empowerment for them in the process of language learning. As this study proved, the classroom is not the only learning environment for learning languages and using strategies as it has been previously hypothesized. The members of the Z generation are increasingly involved in out-of-school learning (Fúz, 2014).

Playing online English games is another strategy that has recently been examined empirically. Butler, Someya and Fukuhara (2014, p. 265) investigated the effect online games exert on language learning. The use of an online English game called 'Jido-Eiken' developed by Japanese programmers in 1994 was scrutinized. This game is designed to teach learners words and common expressions. The complete game comprises nine elements. One is a car-racing game with multiple competitors. It is essential that language learning games be interesting for the students. The researchers identified features that are attractive for the learners. They must be motivated by challenge, curiosity and by control. In this game learners are awarded extra fuel once they answer an English language question. A total of 3,945 children took part in the study, aged 4 to 12. The main finding of the study was that the online game contributed to receptive word knowledge to a great extent and to productive word knowledge to a lesser extent. The YLs were divided into four age groups so that researchers could observe variations dependent on age: 4–5-year-olds, 6–7-year-olds, 8–9-year-olds and 10–12-year-olds. They uncovered that vocabulary learning through online games drastically increased in the 10–12-year-old group. Vocabulary was tested with the use of the VKS. A conclusion can be drawn from this study: playing online games and playing games is an efficient technique for both general language learning and vocabulary learning.

Another investigation on YLs' LLS was conducted by Doró and Habók (2013). The study used the SILL (Oxford, 1991) with 5<sup>th</sup> and 6<sup>th</sup> graders in Hungarian schools (n = 275). It must also be mentioned that the SILL was not originally developed to assess YLs but Doró and Habók (2013) adapted this instrument to the assessment

of YLs by rewording a few statements in the questionnaire. Six strategies were examined: metacognitive, compensation, memory, affective, social and cognitive. Although the main focus of the study was general language learning strategy use, the SILL questionnaire encapsulates a great number of items looking into the use of vocabulary learning. It was revealed that metacognitive strategies were the most frequently used ones by YLs, while compensation strategies were the least often used ones. By refining their findings with respect to gender, they found that girls used new English words in sentences more often and they told rhymes and repeated words to recall the meaning attached to them. They also frequently acted out situations or used mental images in order to memorize words more efficiently according to the self-report questionnaire. From these studies a definite conclusion can be drawn: YLs tend to use cognitive and metacognitive strategies more often than any other strategy in order to memorize new FL words. It has also been argued in this section that LLS are independent of the learning environment.

### 3.4 RESEARCH ON VOCABULARY LEARNING STRATEGIES USED BY YLS

This section is dedicated to the literature on YLs' VLS. The strengths and weaknesses of the relevant studies with a focus on the construct will be presented and the applied data gathering instruments will also be described. Hungarian studies will also be discussed.

Plenty of empirical research has been conducted on VLS; however, few involve YLs (Cameron, 2004). Cameron (2004, p. 92) was led to assert that effective strategies of vocabulary learning at the disposal of YLs are the subsequent: guessing meanings by using all information available in a picture or text, noticing grammatical information about words, noticing linkages to similar words in the NL and remembering where a word has been encountered. Cameron (2004, p. 93) also emphasizes that strategy use changes with age and there is a large variance in terms of what strategies they use and how they use them. I suppose YLs' word knowledge develops and their vocabulary increases when they are exposed to plenty of encounters with the words in speaking, listening, reading, and writing. Research has also indicated that learning words by young children occurs as a result of exposure most of the time (Baumann, Kame'enui, & Ash, 2003) so in many cases word learning is not related to strategies. It has been pointed out

that such strategies as extensive reading, intentional engagement in oral language and listening to adults, all for the purpose of learning a language, are efficient ways of YLs (Ellis, 1994).

Schouten-van Parreren (1992) examined 12–15-year-old Dutch learners of French as a FL. The VLS of reading from context for new lexis were studied. It was found that weak learners were unsystematic in their strategy use compared to efficient learners who used numerous resources to pick up meaning from context: illustrations, linguistic context, the topic, etc.

Nikolov (2003, p. 22) points out that learners use a wide range of strategies; nevertheless, it is also posited that conscious use of strategies were not typical of YLs. Albeit YLs' VLS have been theorized by several researchers, very few studies can be discovered with respect to this domain.

An instrument was developed by Pavičić (2008) to assess the construct as it was pointed out in Table 12: Vocabulary Learning Strategy Questionnaire for Elementary Schools (VOLSQUES). Three main dimensions were identified by Pavičić (2008). (1) strategies of formal vocabulary learning and practicing, (2) self-initiated independent vocabulary learning, and (3) spontaneous (incidental) vocabulary learning. The questionnaire comprised 27 items; every dimension contained nine items. The instrument was validated with item-analysis and factor-analysis with the participation of 300 Croatian children. She unveiled that strategies of formal vocabulary learning and practicing are used most often by YLs, especially, within this classification, ones that involve repetition.

Another study focusing on YLs' vocabulary learning strategy use was conducted by Griva, Kamaroudis and Geladari (2009). Greek-speaking 6<sup>th</sup> graders (n = 238) participated in the study. The researchers used both qualitative and quantitative measures. Besides a self-report questionnaire, think-aloud protocols were applied so that a deeper insight could be gained as far as word learning strategy use was concerned. In the self-report process, the participants were requested to write down the strategies they used frequently to learn words. Translating into the mother tongue, repeating orally and looking up words in a dictionary were reported as the most frequently used strategies. During the think-aloud protocols, the researchers also revealed that a metacognitive strategies were also a frequent instance of the attempt to learn new words. This finding is in line with the results of Doró and Habók (2013).

Coyle and Gomez Gracia (2014) sought to find whether the strategy 'listening to songs to learn new words' used by YLs would prove to be an efficient one. Spanish children of 5–6 years of age (n = 25) participated in the study. Vocabulary was

taught by means of songs in three sessions. Children were told to listen to the songs then to watch the teacher perform gestures related to the unknown words in the songs, and then to identify and link words to pictures. It was revealed that receptive vocabulary was enhanced and in the delayed pre-test a major finding was that some of the children performed better than on the post-test five weeks earlier. It was concluded that listening to songs is an amenable strategy that young language learners can use in the vocabulary learning process.

Hardi (2014) investigated Hungarian YLs' vocabulary learning strategy use within the framework of self-regulation. Her research in three phases. In the first phase she applied semi-structured interviews and classroom observations with a small number of students. In the second phase, she did a focus group interview and structured interviews. Following the interviews the researcher did a pilot study of her data-collection questionnaire that she developed. In phase 3, her validated questionnaire was used with the participation of 3<sup>rd</sup>–8<sup>th</sup> graders (n = 331). Looking up words in the dictionary, oral repetition, translating word to NL were strategies the participants reported applying in order to learn words. One of the main findings of her research was that there is a certain shift in YLs' VLS as they grow older. With the passing of time, FL learners tend to experiment with and apply new strategies. She documented that strategic vocabulary learning changes over time. Metacognitive strategies, even in the case of YLs (8–10-year-olds), were discovered to be frequent in self-reports. Self-regulated strategy use was found as early as 3<sup>rd</sup> grade and this indicates that self-regulation develops at an early age. This refutes the findings of Schmitt (1997) and Jimenez Catalan (2003) who had found that discovery strategies were more frequent strategies used by YLs to learn words.

In this chapter I have elaborated on VLS that were reported to be used by YLs. I have summed up the relevant findings of previous research concerning the strategy use of YLs. I started out by defining constructs of LLS, an overarching topic; then, I have narrowed my focus to VLS with a special focus on YLs' word learning strategies. It has been argued that YLs' word learning strategies are generally assessed with numerous instruments such as questionnaires, self-reports, think-aloud protocols and interviews. It has been asserted that strategy use changes with age especially among YLs and since word knowledge is a multidimensional construct learning strategies are also multi-faceted. From several studies a conclusion can be drawn that cognitive and metacognitive strategies tend to be the most often used ones by YLs. Cognitive strategies that involve translation, formal word learning are used for the discovery of meaning and metacognitive strategies such as repetition for the purpose of learning the new word are made use of in order to consolidate word meaning.

### 3.5 DEVELOPMENT OF THE VOCABULARY LEARNING STRATEGIES QUESTIONNAIRE

Having investigated the instruments assessing vocabulary learning strategies with special regard to those of YLs, a decision was made to consider Stoffer's (1995), Schmitt's (1998) and Pavičić's (2008) questionnaire items adapted from Oxford's SILL (1991) for a large item pool. The reason for this was that these instruments had been either used or adjusted for YLs vocabulary learning strategies were concerned. The pool also involved items that were considered worthy of being a component of a questionnaire assessing Hungarian YLs' vocabulary learning strategies. The items from all of the selected questionnaires were considered for inclusion in my new instrument. These four questionnaires appeared to have the most alignment with the construct and to be best suited for the development of vocabulary learning strategies questionnaire for YLs because these instruments had also been previously used to investigate YLs. I also added some items to my new instrument because new strategies had also come into the picture especially amongst YLs since social network sites became so popular. Some of the items were extended with different variations. For example, the item in Pavičić (2008) 'I watch English language TV shows spoken in English or go to movies spoken in English' was modified in the following way and was broken up into three different items: 'I watch English films with subtitles', 'I watch English films without subtitles and I watch English films with Hungarian subtitles.' Once the pool of the multitude of strategies were gathered, each and every item was examined as to whether they would fit into the instrument and into the Hungarian context. In the wake of this, the items were examined from the perspective of dimensions of vocabulary learning strategies.

Five factors were selected to be the composing parts of the questionnaire: cognitive, memory, metacognitive, determination and social. I decided not to have two major categories and six sub-categories as in Schmitt's taxonomy (1998) due to the fact the above-mentioned dimensions covered all the questionnaire items for a pilot study. Summing up the literature on this issue the five different strategies are defined as follows: (1) cognitive strategies involve the transformation of the target language by the learner, (2) memory strategies include the usage of old material and its linking to new knowledge, (3) metacognitive strategies are ones that exhibit evaluation and review of the cognitive processes by the learner, (4) determination strategies are used by the learners when faced with the challenge of discovering the meaning of a new word, (5) social strategies

Once the five dimensions were decided on, the questionnaire statements were carefully phrased with a view to the specific Hungarian learning environment and circumstances. Two experts were consulted during the process of questionnaire development. All the questionnaire items were thoroughly thought over with special respect to their wording so that they would reflect the construct. Items such as Schmitt's (1998) 'I use a word list to learn words' and Oxford's (1991) 'I use new words in a sentence so I can remember them' were adapted unchanged but these were rare cases in the questionnaire development process. The definition of the cognitive factor in my instrument is the same as Schmitt's (1998): cognitive strategies involve the mechanical repetition of word for the sake of retention. The cognitive factor comprised with eight items and each of them were meant to investigate how students try to retain the knowledge of newly-learned items by using them in a written sentence, in a spoken sentence, etc. The memory factor contained eleven different items. My working definition is based on Schmitt's definition: learners manipulate the words in order to memorize them. The metacognitive factor contained sixteen different statements as I considered it a significant factor to investigate. Metacognitive strategy is the conscious evaluation of the learners' strategies. All the statements focused on this conscious evaluation, manipulation and assessment of the vocabulary learning strategies used by the students. The involvement of the eight items assessing the use of the determination factor was also motivated by Schmitt's taxonomy (1998). Since guessing from context is a traditional, efficient and valid way of learning new words, this factor was indispensable in the instrument. The items were phrased with the aim of gaining the most possible information as to how and how often learners use context to learn new vocabulary. The social factor contained nine items inquiring into the learners' conscious use of the social media and their willingness to turn to their teachers or peers to learn the meanings of the new words. None of the cited data gathering instruments inquire into the use of info-communications technologies (ICT) to learn words, which is natural since at the time of their development ICT tools did not play as vital a role as now in education. This gap was meant to be filled in with statements added to the questionnaire. Three other statements not used by any of the cited researchers were also added as they were regarded as strategies typical in a Hungarian context. The questionnaire with the items in Hungarian and English is presented in Appendix 1.

### 3.6 METHODOLOGY

This study seeks to answer the following research questions

- 1) How do the questionnaire items function?
- 2) Does the factor analysis reflect the original dimensions?
- 3) What are the most frequent vocabulary learning strategies used by Hungarian young learners?

### 3.7 PARTICIPANTS AND PROCEDURE

The instrument was the 52-item vocabulary learning strategies questionnaire.

The pilot study was carried with the participation of 86 Hungarian 6<sup>th</sup> graders in primary schools in Budapest, Mezőtúr and Szeged in February 2014. All the students had studied English from their 4<sup>th</sup> grade (age 10) in three lessons a week.

Four classes were selected to be involved in the study. The headmasters and the English teachers had been requested to provide access to the learners two weeks before the data were taken.

### 3.8 INSTRUMENT

The self-reported vocabulary learning strategies questionnaire was used to gather data. Following the selection of the questionnaire items it was also decided that the data would be collected on a 4-value frequency scale: 'never', 'once a month', 'once a week', 'always.' The decision was made with the intention of forcing to students not to opt for a neutral answer. Since the foundations of my questionnaire were laid on Oxford's (1991) and Schmitt's (1998) data-gathering instruments, their way of data-gathering ought not to be left out of consideration. Oxford used a 5-value frequency scale from 'never' to 'always' and Schmitt's instrument also sought to reveal the frequency of the use of the statements the same way. Hence choosing the frequency scale made sense and it was also decided that instead of a 5-value scale, on which students can give a neutral answer, a 4-value scale would be used so that learners would by all means have to choose to give a solid answer.

### 3.9 PROCEDURE

I went to all the schools and presented the paper-and-pencil 52-item questionnaire to the learners. The questionnaire contained a brief description in Hungarian and a sample statement to which the answer was an obvious and predictable 'never': 'I skype with my Ugandan friend to learn new words.' This was done in order it would be evident for the learners what the correct answer was and what they were expected to do. The children took the questionnaire seriously and filled in it without any disturbance. When the completed questionnaires were collected, I asked each student to write down strategies that they use to learn words on their own. The lists of strategies of all the students were later considered at the pilot of the questionnaire and this resulted in new items being involved in the modified instrument used at the large-sample assessment.

### 3.10 RESULTS AND DISCUSSION

The reliability of the questionnaire was fairly high (Cronbach's Alpha = .91). It was also found that some of the items had 0 standard deviation. Every student indicated 'always' at the statement 'I use a vocabulary list to learn word'. Since this item provides the research with no information from an education scientific sense, it was decided that items having zero standard deviation would not be used in the final questionnaire. The correctness of the decision on adding the item 'My parents check if I have learned the new words by asking me' was confirmed since participants reported high frequency of this activity ( $M = 3.09$ ). Dictionary use also appeared to be a frequent activity used by students. Both the item 'I look up the new word in an English–Hungarian dictionary' and 'I look up the meaning of the word in an electronic dictionary' had high frequencies (3.09 and 3.03 respectively) as it had been previously assumed for Hungarian learners have a tendency of using dictionaries for the purpose of learning words. The activities 'I infer the meaning of the new word from context when reading' and 'I infer the meaning of the new words from spoken English', both related to inferring meanings from context also turned out to be often used by learners ( $M = 2.86$  and  $M = 2.84$ , respectively). This is a finding that is in line with and is confirmed by what Hardi (2014) found when investigating Hungarian learners' vocabulary learning strategy use: with Hungarian learners of a foreign language it is a popular tendency to infer the meanings

of vocabulary from context. In Hungarian schools writing down items with the purpose of memorizing them is also a frequent strategy, as a result participants indicated they often used this strategy ( $M = 2.87$ ). In Appendix 13 the descriptive statistics of some of the questionnaire items is presented.

**TABLE 13.** *Descriptive statistics of the vocabulary learning strategies questionnaire*

<b>Factor</b>	<b>Items</b>	<b>Mean</b>	<b>SD</b>
Cognitive 1	I use the new word in a sentence.	1.945	1.166
Cognitive 2	I write down new word many times.	2.876	1.254
Cognitive 3	I say the new word many times.	2.383	1.178
Cognitive 4	I use a vocabulary list to learn words.	4.000	0.000
Cognitive 5	I use the newly-learned word in speaking.	2.485	1.112
Memory 1	I draw the situation when I would use it.	1.554	1.022
Memory 2	I make a word list in order to remember it.	1.823	.478
Memory 3	I group the words in clusters based on their similarities.	1.374	.797
Memory 4	I link the new word to one with synonymous meaning.	1.222	.686
Memory 5	I link the new word to one with antonymous meaning.	1.513	1.000
Memory 6	I link the new word to one already known.	1.205	.544
Memory 7	I make picture word cards.	1.656	1.087
Metacognitive 1	I listen to English music so as to learn new words.	3.151	1.000
Metacognitive 2	I underline the important word.	2.054	1.224
Metacognitive 3	I circle the word that is important.	2.046	.547
Metacognitive 4	I watch English film with subtitles.	1.414	1.023
Metacognitive 5	I watch English films without subtitle.	1.652	.938
Metacognitive 6	I watch English films with Hungarian subtitle.	1.912	1.122
Metacognitive 7	I watch English cartoons.	2.154	1.124

<b>Factor</b>	<b>Items</b>	<b>Mean</b>	<b>SD</b>
Metacognitive 8	I read English newspapers so as to learn the words.	1.918	1.203
Metacognitive 9	I read English books.	1.412	1.021
Metacognitive 10	I play English computer games.	1.576	.947
Metacognitive 11	I read English cartoons.	1.264	1.166
Metacognitive 12	I read the English labels on every product.	1.538	1.012
Metacognitive 13	I use a new word in writing so as to remember it.	2.244	1.078
Determination 1	I look up the meaning of the new word in a printed dictionary.	1.982	1.286
Determination 2	I look up the meaning of the word in an electronic dictionary.	3.032	1.024
Determination 3	I remember where I saw the new word on the page of the textbook.	1.576	.712
Determination 4	I remember where I have heard the new word.	1.466	1.277
Determination 5	I look up the new word in an English–Hungarian dictionary.	3.098	1.192
Determination 6	I look up the new word in a monolingual dictionary.	2.276	.606
Determination 7	I try to remember the Hungarian equivalent of the new English words.	2.664	.944
Social 1	I ask the teacher what the new word means.	2.818	1.198
Social 2	I learn the new word with a classmate.	2.166	.604
Social 3	I ask my classmate in class what the new word means.	1.854	.943
Social 4	My parents check if I have learned the new words by asking me.	3.687	.677
Social 5	We learn the new words together in group work in class.	2.612	1.298

Looking at the frequency of strategy usage, effectual information can be inferred. The most frequently used vocabulary learning strategies, based on the results of the questionnaire, are the following as shown in Table 14.

**TABLE 14.** *The most frequently used strategies*

<b>Factor</b>	<b>Items</b>	<b>Mean</b>	<b>SD</b>
Cognitive 4	I use a vocabulary list to learn words.	4.000	0.000
Social 4	My parents check if I have learned the new words by asking me.	3.687	.677
Metacognitive 1	I listen to English music so as to learn new words.	3.151	1.000
Determination 2	I look up the meaning of the word in an electronic dictionary.	3.032	1.024
Determination 5	I look up the new word in an English–Hungarian dictionary.	3.098	1.192
Cognitive 2	I write down new word many times.	2.867	1.254
Metacognitive 16	I infer the meaning of the new word from context when reading.	2.866	1.168
Metacognitive 17	I infer the meaning of the new words from spoken English.	2.848	1.277
Social 5	We learn the new words together in group work in class.	2.612	1.298
Cognitive 6	I write or stick the meaning of words onto objects.	2.593	.978

Following the investigation of descriptive statistical data and the frequencies of the different items, item-analysis was carried out through looking at corrected item-total correlations. This value shows how each item correlates with the rest of task. It is a regularly used statistical method in pilot studies since a clear picture is outlined in terms of the functioning of the items. On a sample of 103 students, the reliability and the usefulness of the items with values under .194 are called into question (Falus & Ollé, 2008). Some of the items that fell near or under this .194 value. A decision was made item by item as to which items that fell under the value of .194 would be removed from the questionnaire and those that fell

near this value would further be examined. In Table 15 the items whose item-correlation values were under or near the value of .194 are presented. In response to RQ 1, the malfunctioning items are enlisted. Two items were examined and it was decided that they would be relevant items in the new questionnaire: (1) 'I make picture word cards' and (2) 'I ask my classmate what the new word means'.

**TABLE 15.** *Items with low item-total correlation values*

<b>Item</b>	<b>Item-total correlation value</b>
I write down the words many times.	.184
I say the new word many times.	.263
I write new word and its Hun. meaning into my vocabulary.	.004
I write or stick the meaning of words onto objects.	.122
I relate the new word to one with antonymous meaning.	.195
I make English–Hungarian word cards.	.164
I group the words in clusters based on their similarities.	.222
I read English comics.	.183
I draw pictures next to the word.	.097
I look up the meaning of the new word in a printed dictionary.	.046
I imagine a situation when I would use the word.	.145
I watch English film with English subtitles.	.168
I ask my teacher what the new word means.	.159
I ask my classmate what the new word means.	.227
I watch English cartoons.	.238
I learn the new word with a classmate.	.226
I remember where I have heard the new word.	.164
My parents check if I have learned new words by asking me.	.153
We learn the new words together in group work in class.	.242
I use twitter.	.185

Having investigated the descriptive statistics of the questionnaire results and having gained an insight into the items, factor analysis was conducted to check whether the five factors reflect the original conceptualization. In response to RQ 2, factor-analysis was conducted. After the factor-analysis had been run, it turned out that ten factors existed on the basis of the results. The KMO-index was .72 which was an indication that the strength of the correlation among the five dimensions make it moderately suited for factor analysis. The factor-loadings over the .50 factor-loading limit (Csíkos, 2003, p. 44) were taken into account. It must be noted that Pavičič (2008) took a .40 factor-loading limit in her vocabulary strategy learning research. The ten different factors were considered too many so a frequent procedure was taken in this case. Varimax factor rotation was performed in order that the number of factors would decrease.

### 3.11 CONCLUSIONS AND FURTHER INSTRUMENT DEVELOPMENT

Having looked at the results indicated by the descriptive statistics, the original five dimensions were exposed to factor analysis so that our primary concept would be justified. However, results of the factor analysis clearly showed that data do not reflect the original factors since the SPSS analysis indicated ten factors for the 52 items. Since the ten new factors were found to be overwhelmingly plenty, a decision was made to reduce the factors with Varimax factor rotation and the number of factors was reduced to four.

The interpretation of the four components reveal that statements focusing on metacognitive strategies such as underlining words and circling words load heavily on Factor 1. This factor is named functional since strategies inclusive of this component are related to functional use of the study material. It is also discovered that the factor loadings of repetitive strategy techniques such as rote-learning, repeating words to oneself, looking up words in a bilingual dictionary along with inferring meaning from context, watching English film with subtitles form one cluster in Factor 2. This factor is named the traditional factor since these strategies reflect traditional learning techniques dating from a long time. Reading newspapers and books using Facebook, reading labels on products, listening to music, watching films load heavily on Factor 3, named the authentic factor as all these strategies require encounter with authentic language used by native speakers. The last cluster is named the innovative factor. Such strategies as ‘analyzing and evaluating newly

learned words along with the use of word cards, inferring meaning from spoken context, playing video games loaded heavily on the fourth component. All these strategies have non-traditional way of learning innovative factor. In Table 16 the new factors are presented.

**TABLE 16.** *The new factors after Varimax rotation with the strongest factor-loadings reported*

Items	1	2	3	4
I make a word list to remember the words.		.605		
I make English–Hungarian word cards.				.678
I underline the important words.	.402			
I circle the word that is important.	.505			
I read English newspapers so as to learn words.			.606	
I use new words in my speaking so as to remember them.	.509			
I remember where I have seen the new word on the page of the textbook.		.606		
I use new word in a sentence.	.708			
I play with word games.				.702
I relate the new word to one with synonymous meaning.	.406			
I look up the meaning of the word in an electronic dictionary.				.403
I look up the meaning of the word in a monolingual dictionary.		.503		
I ask my classmate in class what the new word means.	.446			
I use facebook to learn English words.		.389		
I use skype to learn English words.		.489		
I link new word to an already known word.	.602			
I evaluate if I have really learned the word.				.396

Items	1	2	3	4
I watch English films with Hungarian subtitles.			.452	
I analyze parts of the word in order to find out its meaning.				.582
I infer the meaning of the new word in an English context when reading.		.602		
I try to remember the Hungarian equivalent of the new English word.		.556		
I listen to English music in order to learn new words.			.602	
I watch English films without subtitles.			.384	
I read English books to learn new words.			.478	
I play English video games to learn new words .				.581
I read English labels on all kinds of products to learn new words.				.652
I infer the meanings of the words from spoken English.			.346	
I look for English speaking friends in the social media.	.588			
I use the newly-learned word in speaking.	.624			
I use the newly-learned word in writing.	.445			
I ask my classmate in class what the word means in class.	.365			
I make picture word cards.				.426
I repeat the word to myself.		.398		

Items with low total-correlations have been taken out and their factor-loadings are not reported. In consequence 33 items remained in the final version of the questionnaire plus five new added items of which it was thought to fit well into the pertaining factors. The new items were the ones most frequently listed by the participants after filling in the questionnaire as they were requested to list the most frequently used strategies. The next list shows the new dimensions and their questionnaire items including the new added items from the pool of items given by the participants. The 4-value frequency scale remained unchanged and the developed questionnaire was used in the assessment on a large sample.

**The new dimensions, the pertaining items and their previous place in the original questionnaire factors of the vocabulary learning strategy questionnaire**

*Functional*

- I use the new word in a sentence.
- I circle the word that is important.
- I use a new word in writing so as to remember it.
- I link the new word to synonymous meaning.
- I use a new word in speaking so as to remember it.
- I underline the important word.
- I ask my classmate in class what the new word means.
- I look for English speaking friends in the social media.
- I link new word to one already known.
- I use the newly-learned word in writing.

*Traditional*

- I repeat the word to myself.
- I make a word list in order to remember it.
- I try to remember the Hungarian equivalent of the new English words.
- I remember where I have seen the new word on the page of the textbook.
- I infer the meaning of the new word from context when reading.
- I look up the new word in a monolingual dictionary.
- I rote-learn the words (new item).
- I look up the meaning of the new words in a bilingual dictionary (new item).
- I learn new words from my own vocabulary (new item).

*Authentic*

- I read English newspapers so as to remember the words.
- I infer the meaning of the new words from spoken English.
- I use facebook in English to learn new words.
- I skype in English to learn new words.
- I watch English films with Hungarian subtitles.
- I watch English films without subtitles.
- I listen to English music in order to learn new words.
- I read English books.
- I learn new word in order to say whatever I want (new item).

*Innovative*

- I make English–Hungarian word cards.
- I play with word games.
- I make picture word cards.
- I read English labels on all kinds of products to learn new words.
- I analyze parts of the word in order to find out its meaning.
- I play English video games.
- I look up the word in an electronic dictionary.
- I evaluate if I have learned the word.
- I take notes of the words when watching/listening to English programs (new item).

The results of the pilot study provides us with valuable information concerning not only vocabulary learning strategies but also FL learning and teaching in Hungarian classrooms and in different other learning environments. It became clear from the data that besides writing the words in a bilingual vocabulary, students are checked by the parents whether they have learned the new words or not. In Hungarian schools, especially until the end of primary school a considerable part of the parents puts special focus on their children's studies. The other eight most frequently used strategies reported by the students reflect the special features of Hungarian YLs learning FL words. Looking up words in either a printed or an electronic dictionary has always been a popular and favored strategy by students not just in Hungary but globally. Reading English comics appears to be a frequent activity, consequently a strategy, applied by the Hungarian YLs. This item was not adapted from any of the cited questionnaires in the literature but it was my own decision to involve it in the questionnaire. This result confirms and justifies the correctness of this decision. It is somewhat surprising that the participants indicated more frequent comics reading ( $M = 3.26$ ) than listening to English music with the purpose of learning words ( $M = 3.15$ ) since listening to English music is not only a daily activity but a popular trend among Hungarian 12-year-old children. It seems that comics is still a favored activity by the Hungarian primary school students. However, because of the low item-total correlation value of the item 'I read English comics', it was not included in the final questionnaire developed for online use. It has also been revealed, in accordance with my presupposition, that asking the teacher for the meaning of the new word, writing down the word many times, remembering the Hungarian equivalent of the new word and learning words in group work are within the most applied vocabulary learning strategies. However,

drawing pictures next to words was also reported to be a popular strategy ( $M = 3.01$ ). Visualization might have become such an everyday part of the children's lives that they use images in all situations as a way of learning new words. Drawing images appears to be a habitual activity amongst most of the students. It must also be kept in mind that in the questionnaire the learners had to indicate how often they use certain strategies. Value 3 on the 4-value scale meant 'every week'. Hence, the majority of students draws pictures next to words ( $M = 3.01$ ) every week but it does not necessarily mean that they draw images to all the words nor does it mean that they draw them next to even the half of the words.

Attention must also be paid to the strategies which had been supposed that to be more frequently used ones. They then turned out to be either hardly ever used or to have low standard deviation. Contrary to the relatively frequent activity, drawing pictures next to words ( $M = 2.09$ ), it turned out that students do not draw entire situations and do not use social network for vocabulary learning purposes consciously. Learning English words by using such social network sites as facebook and twitter and conducting English conversations on skype is still an undiscovered area amongst YLs. None of these strategies were reported to be used more on a weekly basis as the highest mean of these (use of facebook) is under the value 2. Even though a considerable amount of time is spent on the use of social network sites they do not yet see an opportunity to learn English words with their assistance. The metacognitive activity, evaluation of whether the word has been learned or not, is also an infrequent strategy among young language learners. With all probability, at the age of 12 students are not as mature as they could evaluate their own learning process successfully. In Appendix 2, the new questionnaire is presented with the items translated into English and it is also clarified which factor each item belongs to.

## 4 VALIDATING THE INSTRUMENT ASSESSING VOCABULARY

Nowadays it is a recognized fact among researchers that lexical knowledge is the main predictor of fluency in any language no matter what skill is concerned (Laufer, 2009; Nation & Meara, 2010; Webb & Sasao, 2013). The recognition of the central role of vocabulary has inspired numerous studies tapping into the necessary vocabulary size of the learners. In a study with a sample of high school students, Laufer (1997, p. 152) found that 95% of the words must be understood by the learners so that comprehension of the text will not be obstructed.

In a later study, Hu and Nation (2000, p. 426) uncovered that readers needed knowledge of 98% of the words in a text to achieve successful comprehension. In this study, the coverage of a fiction text was manipulated in a manner that words were replaced with nonwords. Four different coverage levels were conceived and FL readers' comprehension of these texts was assessed. At the 80% coverage level no learner achieved adequate comprehension and at 100% the majority did. With the knowledge of 90% and 95%, few participants gained adequate comprehension of the texts. Hu and Nation's (2000, p. 422) conclusion was that the required coverage was between 95% and 100% and proposed a figure of 98%.

These lexical coverage figures provide scholars and teachers with relevant information as they supply assistance for scholars in calculating the vocabulary size which speakers need in order to use language. For example, Nation (2006) calculated that a vocabulary size of 8,000 to 9,000 words is necessary to reach a 98% coverage level in written texts. Laufer and Ravenhorst-Kalovski (2010, p. 24) also estimated that it takes around 8,000 word families to allow a coverage of 98%. Even though the samples of these studies are learners older than 14 years of age, it might be assumed that the 98% rule is a valid figure as far as YLs are concerned. The results of the Laufer (1992) and Nation (2006) studies have exerted an influence on teaching methodologies. Developing relevant lexical knowledge is regarded now as a vital factor in the process of a YL learning a FL. Alderson and Banerjee (2002) even suppose that there is a critical lexical threshold below which the learner has extreme difficulties in comprehending and even taking part in comprehensible interaction.

Vocabulary is nowadays considered as one of the strongest predictors of FL proficiency (Schmitt, 2008, p. 352). Diagnostic vocabulary tests have been developed and validated in the past 30 years. The major characteristic feature of these instruments is that they test one dimension of knowledge. They either assess receptive or productive knowledge of words and hardly any attempt has been made towards an instrument

that assesses both receptive and productive FL word knowledge. Besides the vocabulary tests assessing either the receptive or the productive dimension of word knowledge, several instruments (e.g., Jiménez Catalán & Terrazas Gallego, 2008; Orosz, 2009) have been adapted to testing YLs' FL vocabulary. All the validated tests focus on the diagnostic assessment of adult learners. Two instances of assessment of YLs' English as a FL vocabulary were reported. (1) Orosz's assessment (2009) was carried out with a previously validated Yes/No test that might not provide sufficient data concerning Hungarian YLs' English word knowledge, (2) Vidákovich, Vígh, Sominé Hrebik and Thékes (2013) assessed Hungarian YLs' English and German as a FL receptive vocabulary with a diagnostic tool that the researchers had developed in an online environment.

Most of the validated diagnostic vocabulary tests are originally paper-and-pencil tests and there is a lack of vocabulary measures carried out in an online environment. Even though the paper-and-pencil vocabulary tests have been computerized and are available on Tom Cobb's website at [www.lex Tutor.ca](http://www.lex Tutor.ca), hardly any study has been published on the assessment of vocabulary executed with an online instrument except for Vidákovich et al. (2013).

As this summary shows, there is a need to develop a diagnostic online English as a FL vocabulary test assessing YLs' receptive and productive word knowledge.

Triangulating data is of utmost importance in educational research because the richness and complexity of the gathered information can be fully mapped out and explained by analyzing it from a different perspective (Cohen & Manion, 2000). Two instances of triangulation will be executed in the research. Concerning the vocabulary test, teachers will be requested to make an assumption of the score of an average 6<sup>th</sup> graders. Only those teachers will be called upon to participate that teach this age-group. Besides the online vocabulary test scores and the results gathered on teachers' estimations, think-aloud protocol will be implemented so that test solving techniques can be revealed. Considering the general purpose of the research the following research questions were phrased.

- 1) Which task of the vocabulary test proves to be the most simple and which proves to be the most difficult?
- 2) How do the different items function on the vocabulary test?
- 3) How do the young learners perform on the online vocabulary test?
- 4) How do the different tasks of the vocabulary test correlate with one another?
- 5) How do the high-achievers perform on the productive task of the vocabulary test?
- 6) How do teachers estimate the vocabulary size of 6<sup>th</sup> graders?

## 4.1 METHODS

### 4.1.1 Instrument

Three data-gathering instruments were used in the study:

- 1) the online vocabulary task comprising six tasks to map the English as a foreign language vocabulary of the students;
- 2) think-aloud protocol elicited during vocabulary test solving to map the students' thought processes;
- 3) a paper-and-pencil questionnaire for teachers related to the vocabulary test to gain insight into teachers' assumptions on students' vocabulary size.

A diagnostic complex vocabulary test was designed to assess learners' word knowledge. Most of the diagnostic vocabulary tests measure one dimension of vocabulary (Nation, 1990). They either tap into receptive or productive word knowledge. The diagnostic instrument consisted of seven different tasks as it is presented in Table 17.

**TABLE 17.** *Tasks in the diagnostic vocabulary test battery*

<b>Task</b>	<b>Receptive/ Productive</b>	<b>Language skill(s) and modality required Schmitt (2014)</b>
1) Listen to words and match them with pictures.	Receptive	Listening / Meaning recognition
2) Listen to definitions and match them with words.	Receptive	Listening / Form recognition
3) Match 6 written words with 3 pictures.	Receptive	Reading / Meaning recognition
4) Match written words with picture.	Receptive	Reading / Meaning recognition
5) Match written definitions with words.	Receptive	Reading / Form recognition
6) Write word next to picture.	Productive	Writing / Form recall

Up to this point vocabulary had been assessed with tests comprising tasks identical in format. Tests had either assessed receptive or productive word knowledge in one modality. The validity of none of the tests was called into question. However, questions may arise in case an instrument consists of several different tasks. There might be some skepticism whether an item assessed in listening mode would produce the same results as in reading mode. In my view, in an item pool containing 108 words, the overall result achieved in the complex test does not differ from that achieved, say, in a receptive vocabulary test comprising tasks of identical format. According to Melka Teichroew (1982, p. 244) the receptive-productive distinction is rather a continuum than two types of knowledge. It is also asserted that it is not clear where the threshold is found between receptive and productive knowledge (Goldstein, 2004). The deficiency in determining the place of this threshold evidences the fact that assessing a number of items in different modalities does not exert an influence on the results.

Besides taking corpus-based data into account, recommendations in the *Hungarian National Core Curriculum* (2007) and Nikolov (2011) were also considered in terms of grouping words based on topics and involving them in the list. The topics suggested were are (1) food and eating; (2) home and furniture; (3) shops and shopping; (4) travelling and transport; (5) jobs; (6) professions and sports. Nikolov (2011, p. 28) suggests 14 broader topics that should be considered by elementary school teachers for classroom practice and she also presumes that the lexis that is included in these topics might be the area of interest for the young language learners. Consequently, I added the most relevant vocabulary of these topics to the list of 2,000 words irrespective of word frequency rank. As a result, my list of words to be assessed was completed by the addition of another 2,000 word families summing it up to 4000 words. This decision is supported by the evidence found by Nation and Waring (1995) that the knowledge of the 4,000 most frequent words is the most critical aspect of communicating in a language.

For the measurement tool six of the major topics specified above were selected. There are two reasons for this decision. Not all of the 14 topics could be included in the test and after thorough supervision these six topics included the most frequent vocabulary of all the fourteen. I came to this conclusion after looking at the word lists of these topics and compared them with the frequency lists. Six tasks (Task 1–Task 6) of this complex vocabulary test were intended to assess breadth of vocabulary since most vocabulary tests (Meara, 2009; Nation, 1990; Read, 2000) assess this domain. One task (Task 7) was intended to assess depth of vocabulary. The required word knowledge for task solving was receptive in the first five tasks

and in Task 6 and 7 productive word knowledge was the requirement. The VKS was implemented in Task 7. Moreover, I reckoned that it would have been a heavy cognitive load for 6<sup>th</sup> graders if I had tested depth in more than one task.

The paper-and-pencil version of the vocabulary test was piloted in November 2013 with 103 participants. Item-analysis was conducted to see the functioning of the items and the tasks. With the tools of descriptive statistics results were analyzed and several decisions were made concerning the removal and replacement of items. First of all it was decided that Task 7 would be removed. The main reason for this was that this task showed negative correlations with some of the other tasks. Items with zero standard deviations were also removed and replaced and other instances of replacements occurred in case an item was under .194 (Fauls & Ollé, 2008). After item-analyzing and finalizing the pilot paper-and-pencil vocabulary test, I consulted the information-technology experts of the Institute of Educational Science of the University of Szeged. Assistance was provided by them in converting the finalized paper-and-pencil instruments into an online environment. The test was uploaded onto the online platform developed by Institute of Educational Science called eDia.

In the vocabulary test, all items were designated to three categories. Category 1 words were considered the easiest and Category 3 the most difficult. This classification was determined based on rank, frequency in textbooks used by 6<sup>th</sup> graders and professional recommendations. Out of the nine items the dispersion of the categories were the following: either four or five Category 1 words, either two or three Category 2 words and either one or two Category 3 words. Category 1 words are normally more frequent grounded on the BNC; however some words related to children's vocabulary with lower ranking were categorized higher than some higher ranked words in the BNC. Appendix 3 presents the items on the vocabulary test with their rank number in the BNC, frequency and their category.

Edia is a platform under constant development and is well-suited for efficient data gathering on a large sample. The voice files were also attached to the first two tasks of the vocabulary test. My voice, the researcher's, was recorded reading up the pertaining items. Every task contained a sample task that was presented to the students before they went about taking the test. Taking the vocabulary test took approximately 15 minutes. Students sat down in front of the screen with headsets over their ears so that they could hear the voice file of the first two tasks.

### 4.1.2 Participants and procedures

The sample was selected by the coordinators of the Institute of Educational Science. The Institute filed a request to schools in Hungary and twelve schools agreed to involve their students in the research. Participants were 282 Hungarian 6<sup>th</sup> graders. Sampling was non-representative; however this had not been an original goal.

The volunteering schools were given a passcode to be able to log into the eDia platform where the vocabulary test could be accessed. Data were gathered in November 2014 and data processing was performed with the use of the SPSS 17 software.

## 4.2 RESULTS

As described earlier, the vocabulary test contained 54 items. In all the six different tasks there were eleven items. One item was an exemplary item, one was a distracting item; as a result test-takers had to prove the knowledge of nine item. So, in every task the maximum achievable points were nine making the instrument a 54-point test. Reliability of the test proved to be acceptable (Cronbach's Alpha = .869). In Table 18 the descriptive statistics of the six tasks are presented.

**TABLE 18.** *Descriptive statistics of the six tasks in the vocabulary test*

	<b>Mean</b>	<b>SD</b>	<b>Reliability (Alpha)</b>
Task 1	6.393	2.039	.762
Task 2	3.804	2.534	.812
Task 3	6.135	2.347	.763
Task 4	2.756	2.292	.745
Task 5	2.763	2.293	.770
Task 6	3.380	1.934	.723

Laufer et al. (2004) argue form recognition is expected to be harder than meaning recognition. In the case of the two reading tasks, this argument proved incorrect. In spite of the fact that students performed below 30% in Task 5 ( $M = 2.763$ ), in Task 4 ( $M = 2.756$ ) they achieved even worse refuting the hypothesis that a form recognition task would be more difficult than a meaning recognition task.

Contrary to the paper-and-pencil pilot study, on the online test with a larger sample size, participants had the best achievement on Task 1. In the pilot study, Task 3 proved to be the task students which students achieved the best at. Nonetheless students proved to achieve the best on Task 1 and Task 3 during both test procedures. Both tasks are done in meaning recognition modality which is assumed to be the easiest in the hierarchy of modalities (Laufer et al., 2004; Schmitt, 2014). It must also be noted that students scored a lower number of overall test points in the online environment than in the traditional paper-and-pencil environment; however it is not the goal of this book to compare foreign language testing media. Another important finding is that the two reading tasks proved to be the most difficult of all six tasks. Task 4 that required task solving in the modality of meaning recognition and the use of reading skills appeared to be the most difficult for the test-takers whereas in Task 5 demanding task solving in the modality of form recognition and reading definitions and matching them with lexical items participants reached a bit higher number of points than in Task 4, a modality that is assumed to be simpler in the hierarchy. It needs also to be underline that in the task that necessitated the use of productive vocabulary, Task 6 in the modality of form recall, assumedly the most difficult modality, students scored significantly more points than in Task 4 and Task 5. This finding ought to be envisioned in a deeper way. In Task 5 students had to drag a line between the lexical item and the pertaining definition while in Task 6 a set of well recognizable pictures were at their disposal and they had to write one item next to picture. In an online environment it may be easier for students to recall words grounded on recognizing pictures than dragging a line between words and their definitions that might contain lexical items unfamiliar to them. It must not be left out of consideration that the productive task, Task 6, had the lowest reliability value whereas Task 2 in which learners were expected to match definitions they heard to words proved to be the most reliable task.

In order the results of the six tasks can be seen clearly, a histogram is presented. The numbers on the abscissa are the tasks. The maximum points to be achieved by the participants is altogether 2,538 since every participant had the chance of scoring nine points in each task and the number of participants were 282. The two reading tasks (Task 4 and Task 5) were the most difficult and the first listening task (Task 1) and a reading task in meaning recognition modality (Task 3) were the easiest. Having examined the six tasks, the descriptive statistics of all the items on the vocabulary test must inevitably be envisioned with particular regard to the item-total correlation values that give account of how each item behaves in a test. In Appendix 4 the descriptive statistics of the items on the test is presented.

**TABLE 19.** *Itemwise descriptive statistics of the vocabulary rest*

<b>Item</b>	<b>Task</b>	<b>Mean</b>	<b>SD</b>	<b>Item-tot corr.</b>
monkey	1	.706	.456	.338
lion	1	.635	.482	.270
airplane	1	.507	.500	.317
tram	1	.709	.454	.405
swimming	1	.858	.349	.334
helicopter	1	.862	.345	.336
ship	1	.890	.443	.352
camel	1	.858	.232	.426
skating	1	.592	.492	.430
supermarket	2	.585	.493	.386
theatre	2	.862	.345	.404
bake	2	.359	.480	.382
cinema	2	.477	.500	.475
eat	2	.320	.467	.409
hospital	2	.206	.405	.449
learn	2	.253	.435	.406
play	2	.658	.475	.469
sell	2	.534	.499	.420
boat	3	.712	.453	.427
drink	3	.683	.466	.394
drive	3	.676	.468	.486
heavy	3	.737	.441	.431
leg	3	.475	.500	.302
hit	3	.932	.252	.264
pocket	3	.800	.400	.448
quick	3	.682	.466	.513

<b>Item</b>	<b>Task</b>	<b>Mean</b>	<b>SD</b>	<b>Item-tot corr.</b>
small	3	.432	.496	.290
busdriver	4	.371	.484	.276
waiter	4	.675	.469	.497
cook	4	.418	.494	.485
firefighter	4	.368	.483	.438
hairdresser	4	.246	.437	.333
mechanic	4	.150	.357	.269
pilot	4	.161	.369	.340
plumber	4	.136	.331	.335
tailor	4	.193	.392	.277
bedroom	5	.676	.471	.204
cook	5	.414	.493	.232
cup	5	.422	.495	.224
curtain	5	.383	.485	.207
dining room	5	.242	.431	.201
open	5	.151	.358	.265
shelf	5	.164	.365	.226
talk	5	.142	.344	.261
wash	5	.181	.387	.282
cake	6	.237	.420	.266
cheese	6	.514	.501	.261
chicken	6	.446	.497	.276
coffee	6	.824	.386	.265
fish	6	.378	.484	.255
hotdog	6	.164	.371	.295
icecream	6	.586	.494	.019
cucumber	6	.192	.314	.332
sausage	6	.162	.364	.288

Some manifest data needs to be deeply envisioned. It ought to be highlighted that the item-total correlation values of all items except for 'icecream' proved to be acceptable, i.e., above the .194 limit (Falus & Ollé, 2008). Even though some items were in the vicinity of this critical value (e.g., 'lion', 'sausage', 'hit' and 'busdriver') the instrument does not appear to suffer from low item-total correlation values, thus it can be claimed that the entire instrument yields valid results. It is a remarkable fact as well that items with low item-correlation values have an even distribution across tasks. No task has more than one item that strikingly behaves improperly. This provides evidence for the fact that the tasks requiring different task-solving modalities have equal strength and assess the same construct.

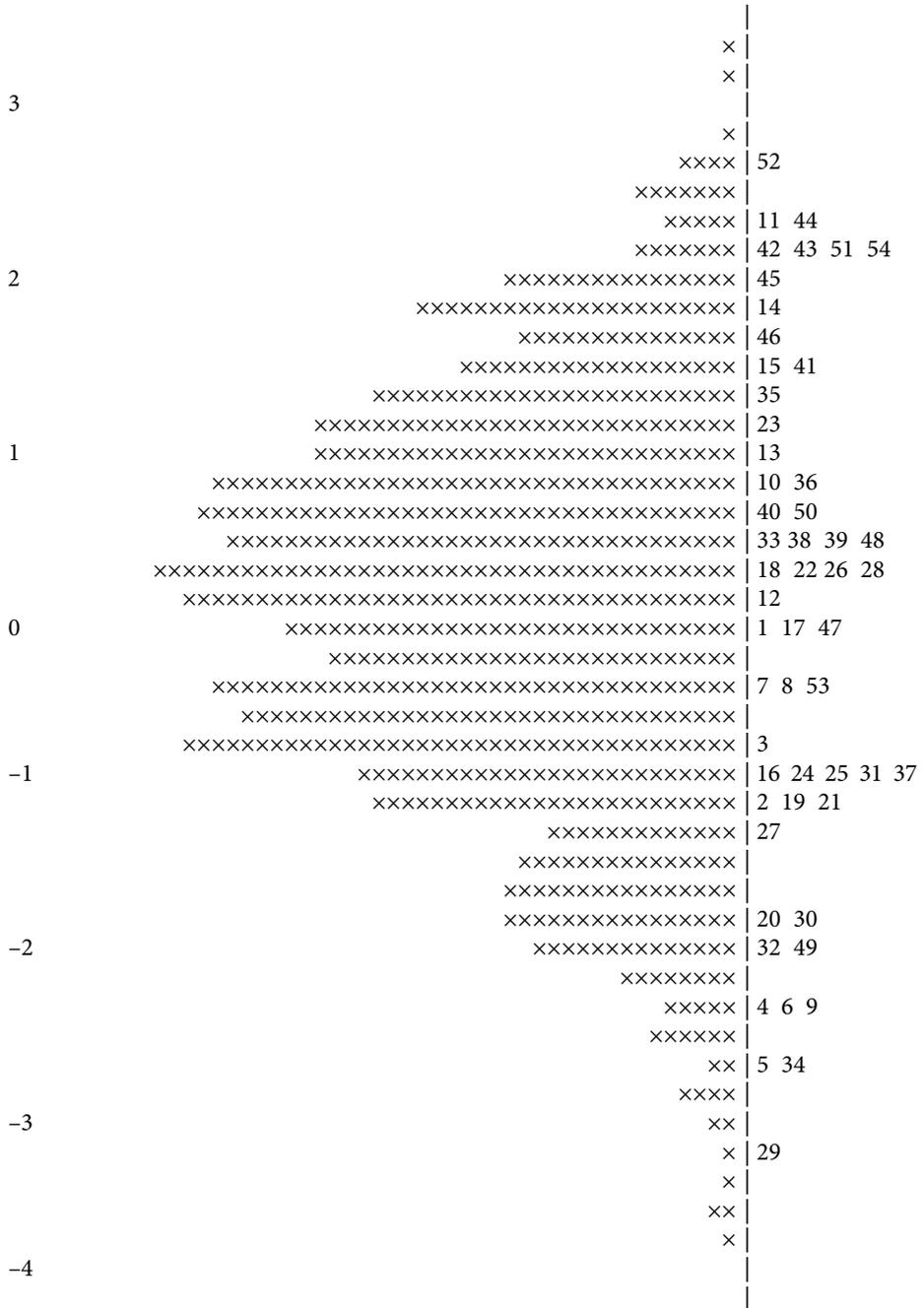
Having examined the results of the items with the means of classical test theory, the applicability of the tools of modern test theory was also considered. The Rasch-model was applied to gain a deeper insight into the reliability of the test and the difficulty of each item. The value used in modern test theory, EAP/PV, of .912 yielded evidence of high reliability. In Figure 1 the item difficulty values are presented. The logit values are shown from -4 to 3. The values below zero indicate easy items and those above zero indicate difficult items. The further an item is in the positive range, the more difficult it was in the test and vice versa, the further the item is in the negative range, the easier it proved to be in the vocabulary test. As 54 items were assessed in the test, the same number of items are scaled by logits. Based on the model, the assertion can be made that the test has a relatively normal distribution. Most of the items are in or near logit 0, which indicates a normal distribution. As regards easiness and difficulty of items, easy and difficult items are represented almost in an equal ratio, so the test differentiates properly. Each x represents 0.4 students.

Using item response theory, the distribution of the items can also be examined visually. Figure 9 precisely shows the distribution. The top right part of the graphs almost reflects identically the left bottom part, which yields the information that the assessment tool provides valid data.

### 4.3 FREQUENCIES OF SCORE RANGES

Having analyzed the items in all tasks, the distribution of the score ranges must be envisioned so that a clear picture can be received as far as students' achievement is concerned. Table 42 presents the score ranges and the number of students pertaining to them. Before going into any discussion, it is observable that the test differentiated properly among students with the number high-achievers being more than that of low-achievers.

FIGURE 1. *The item difficulty values of the vocabulary test*



The maximum point to be received was nine in each of the six tasks, making 54 the overall maximum possible total score. No student achieved 54 points; however twelve reached a remarkable score of 46–48 points. Ten knowledge ranges were determined with five point units except for the top range that was calibrated to the above-mentioned 46–48 since no higher score than 48 was observed. The number of the worst-achieving students, within the range of 1–5 was four and by doing a slight extension to the range of 1–10, the cumulative number of low-achievers is twelve, which is an acceptable number on a sample of 288. This means that not even the 10% of the students achieved below ten points.

By examining the other extremity, the high-achievers, it can be stated that the number of the high-achievers, number of students within the range of 41–48 is 10, which means that not even 5% of the students scored more points than 41. It is inevitable to note that 23 students, almost exactly 10% of the sample scored over 36 points.

As it is expected from a properly differentiating diagnostic test, most students achieved in the range of 40%–60%. The 50% of the total points is 27, which means that in the range of 26–30 points 63 students can be found and 53 students reached the range of 31–35 points. Out of 288 test-takers 116 of them achieved in the average range of 26–35 points, which means that nearly one-third of the sample had an average achievement.

Having analyzed the test score at the item and student levels, it is of paramount importance to examine the correlations among tasks so that deeper relationships can be revealed at task level.

#### 4.4 CORRELATIONS ACROSS TASKS IN THE VOCABULARY TEST

The diagnostic instrument assessing word knowledge, as it has been described so far, comprised six tasks. The first two tasks were two listening tasks in meaning and form recognition modality. The third task was a reading task in meaning recognition modality that expected test-takers to match items with pictures. The fourth and the fifth task were reading tasks in meaning and form recognition modality, respectively whereas the sixth task was a productive writing task in form recall modality. The correlations among these tasks were investigated to see whether the reading tasks had strong relationship with one another and whether the two listening tasks showed any correlations. It was also examined how significantly Task 6 correlated with the rest of the tasks. Table 20 presents the correlation matrix of the six tasks.

TABLE 20. *Correlations among tasks of the vocabulary test*

	Task 2	Task 3	Task 4	Task 5	Task 6
Task 1	.501**	.434**	.337**	.065	.149*
Task 2		.557**	.530**	.012	.115
Task 3			.517**	.068	.070
Task 4				.368*	.051
Task 5					.476**

\*\**. Correlation is significant at the .01 level (2-tailed).*

\**. Correlation is significant at the .05 level (2-tailed).*

Task 1 and Task 2, the two listening tasks show a significant correlation ( $r = .501$ ,  $p < .01$ ), meaning that no matter whether the modality is meaning recognition or form recognition, the two tasks measure the same construct. Task 4 and Task 5 also correlate significantly with a slightly weaker relationship ( $r = .368$ ,  $p < .05$ ). Two similar tasks which required the students to match pictures with the items, Task 1 and Task 4 correlate significantly ( $r = .337$ ,  $p < .01$ ); however the listening task, Task 2, requiring learners to match items with definitions does not show any correlation with the reading task, Task 5, requiring learners also to match definitions with items. It is intriguing to observe that two related tasks in terms of task solving function have very weak relationship and insignificant correlation with each other within the same test. This result reflects the assumption (Vidákovich et al., 2013) that listening to and reading definitions might be two totally different task solving functions. Furthermore, it is hard to rely on previous research data as young learners' vocabulary had only been assessed in only one modality in each testing instrument. Vocabulary knowledge in different modalities had not been assessed; thus no comparable data are accessible.

By investigating the correlations of Task 6, the productive writing task in form recall modality, crucial information can be procured. Task 6 has a weak relationship with Task 1 but the correlation is significant ( $r = .149$ ,  $p < .05$ ). This means that a task requiring the use of a receptive skill, listening, has a stronger relationship with a productive task than with another task also requiring reading skills. Task 6 is also significantly correlated to Task 5. This root of this relationship might be found in the fact that words in these two tasks were ones denoting household items and

activities (Task 5) and food (Task 6). These items form a set of words that are usually learned in a collected cluster. The chapters comprising these two sets of words in the course-books used in schools are in close vicinity to each other. Learners that know words meaning food are likely to know those meaning household activities and learners who are not aware of household vocabulary are probably unaware of food vocabulary in a recognition modality, let alone in a form recall modality.

## 4.5 DISCUSSION

Task 1, the listening task of meaning recognition modality, proved to be the easiest ( $M = 6.39$ ) and the most difficult task was Task 4, a reading task of meaning recognition modality ( $M = 2.75$ ). It was asserted during data analysis that a task of form recall (Task 6), a supposedly difficult task, proved to be easier ( $M = 3.38$ ). In response to RQ 2, item-total correlation values were evaluated. This value is calculated to see if any of the items do not have responses that vary in line with those items for other tests in the population. In other words, this calculation is performed to check if any item is inconsistent with the averaged behavior of the other items. The minimum of this item-total correlation value, according to the literature, is .194. None of the items, except for 'icecream' (.018) fell below this value. In case a test is under development, it is suggested that the items below .194 should be discarded. In this case there is no possibility to replace 'icecream' so it is not taken out of consideration; however in further research a new item will be implemented in Task 6. Some very low values are come across in, for example, the item the most learners knew, 'hit' had a value of only .264. 'Lion' was also fairly inconsistent with the rest of the test with a value of .270. In an instrument with 54 items, one item not being consistent with the rest of the items might be acceptable. However, it is a striking finding that in Task 5 all of the items' total-correlation values are below .300 but above the .194 limit. Task 5 proved to be the most difficult task as it was stated earlier. Task 5 correlated significantly with Task 4 and Task 6 and had a weak relationship and insignificant correlation with the rest of the tasks. Since none of the items in Task 5 are of unacceptably low item-total correlation values, it can be asserted that Task 5 fits in well with the entire test.

In answer to RQ 3, the sample was divided into score ranges of five point units. Four students fell within the score range of 1–5 points and eight students within

the 6–10 point units. This means that twelve students knew fewer than ten words. Even though they had been learning English for two years, at the time of test-taking they had a vocabulary of around ten words. It is incredibly low. As for the average achievers, within the score ranges of 21–25, 26–30 and 31–35, 168 students are found out of the 288 test-takers. By carefully envisioning the badly-achieving, the average-achieving and the well-achieving parts of the sample, a normal distribution can be noticed, which means that the criterion of the classical test theory of proper differentiation is realized. The six tasks showed significant correlations with one another with the exception of Task 5 and Task 6. Task 5 had a weak relationship with Task 2 ( $r = .012$ ) and a strong relationship but no significant correlation with Task 1 and Task 3 ( $r = .065$  and  $r = .068$ , respectively). Task 6 had a weak relationship with Task 2 ( $r = .115$ ) and no significant correlation with Task 3 and Task 4. It was earlier pointed out in this book that it is hard to find a reason for the near zero relationship between Task 5 and Task 2 because they were of the same modality (form recognition) and the task was the same: matching words with definitions. The only difference was the skills required to solve the tasks: listening and reading. It was supposed that the productive task in form recall modality would be the most difficult task and as such it would be a major differentiating factor among the participants of different word knowledge. As it was discussed earlier in response to RQ 1, Task 6 did not prove to be the most challenging task. However, I intended to know how high-achievers performed on this particular task to gain better insight into the organization of their vocabulary.

High-achievers had a mean of 4.784 on the productive task, which means that they reached nearly 50% on this task. It is a low value compared to the number of points they reached on Task 1, Task 2 and Task 3. None of them had the maximum nine points on this task and one of the high-achievers on the overall test has as few as two points. This result gives evidence to the fact that a form recall modality task is difficult and most of the Hungarian 6<sup>th</sup> graders are not prepared to use words or word clusters in production. The classroom implication can be concluded that even learners of good ability must be trained for productive use of the foreign language so that their communicative skills can be improved. Having compared the results of what teachers assumed and what students achieved, it can be asserted that teachers of English of 6<sup>th</sup> graders generally overestimate the word knowledge of students. Besides envisioning the descriptive statistics of teachers' assumptions and students' test scores, the t-test also confirmed the existing over-estimation of students' word knowledge.

## 4.6 CONCLUSION

The investigation of YLs' English as a foreign language vocabulary size was a major endeavor since an online data-gathering instruments had to be developed and created. Having conducted a pilot study with the two instruments, item-analysis and several statistical procedures were executed in order that a properly functioning test would be used on large sample for the sake of unveiling correlations and of gaining a deeper insight into the organization of vocabulary.

It was of utmost importance to triangulate the data. With regard to the vocabulary test, a think-aloud protocol was conducted with a student of average language proficiency and school achievement, and teachers were also requested to make their assumptions as to the probable achievement of 6<sup>th</sup> graders on the test.

With regard to the results, the listening task of meaning recognition modality, proved to be the easiest and the most difficult task was Task 4, a reading task of meaning recognition modality. It was asserted during data analysis that a task of form recall (Task 6), a supposedly difficult task, proved to be easier than Task 4 and Task 5. To gain a clear picture of the functioning of the items, total-correlation values were also envisioned. None of the items, except for 'icecream' (.018) fell below a critical value.

Having divided sample was divided into score ranges of five point units, a more sophisticated dataset was gained. Four students fell within the score range of 1–5 points and eight students within the 6–10 point units. This means that twelve students knew fewer than ten words. Even though they had been learning English for two years, at the time of test-taking they had a vocabulary of around ten words. As for the average achievers, within the score ranges of 21–25, 26–30 and 31–35, 168 students are found out of the 288 test-takers. By carefully envisioning the badly-achieving, the average-achieving and the well-achieving parts of the sample, a normal distribution can be noticed, which means that the criterion of the classical test theory of proper differentiation is realized.

## 5 PHRASAL VERBS

### 5.1 DEVELOPING AN INSTRUMENT ASSESSING YLS' KNOWLEDGE OF PHRASAL VERBS

Hungarian 6<sup>th</sup> graders are 12-year-old learners, most of whom, have three or four 45-minute lessons in English in primary schools. The majority of the public schools do not provide more than four English lessons for students per week (Fazekas, 2009, p. 4). However, most of the students in Hungary have more exposure to English than the three or four occasions determined in the school curriculum. They attend private language lessons or courses organized by local language schools in the afternoons and they are exposed to a large amount of English by using the Internet.

Besides being exposed to vocabulary learning in school and private lessons, students are also believed to learn vocabulary by listening to songs on Youtube and reading posts on social media sites. In the schools where I conducted the pilot study I interviewed the teachers to make sure I am aware of what course-books were used. Information was provided that they used course-books published by either of the three major publishers: Cambridge University Press, Oxford University Press and Pearson. It is characteristic of these course-books and workbooks that they have well-designed chapters and contain a great deal of visual material. As for course-books, the investigation of how teachers apply the course-books in practice and what methods they use to teach English is also an important aspect. In an empirical study conducted by Nikolov (2008), it was pointed out that the observed teachers in the research tended to use the grammar-translation method (Harmer, 2012), and their native Hungarian to explain grammar and vocabulary meaning. It was asserted that English words were generally taught with their Hungarian equivalents and very few motivating techniques are applied to help learners learn vocabulary.

All students in primary school start learning it in 4<sup>th</sup> grade and normally they attended three Russian lessons a week until 12<sup>th</sup> grade, the final grade in public education. Because of the difficult orthography and lack of willingness on the part of the Hungarian learners to identify themselves with the language, Russian was not a popular subject and most of the learners left public education without being able to communicate in this language (Nikolov, 2007). Since Hungary became a democratic country (the first democratic elections were held in 1990) Russian teachers have been retrained into English teachers. Learning English and German has

become popular, especially English as lingua franca, since the turn of the millennium (Csizér & Dörnyei, 2005). Nowadays, an increasing number of children start their FL studies prior to the mandatory age of ten. Due to parental pressure more and more YLs begin to study English before grade 4 (Nikolov & Szabó, 2011, p. 16). The most popular FL is English but German, French, Italian, Spanish and Chinese are also offered in some schools.

Hungarian YLs gain access to English words from three main sources: (1) public school classroom, (2) private lessons, and (3) incidental instances of hearing or reading words. As for the classroom, vocabulary input can be received by the student from teacher talk. One study puts English education in Hungary into focus (Nikolov, 1999a). A lot of useful observations can be made based on this study. One is that teachers in Hungary use NL in the majority of the classroom time and tend not to use pictures, videos or songs to teach language and vocabulary, a method that learners would prefer according to their report. On the basis of Nikolov's (1999a) data originating from a series of classroom observations in Hungary, I reckon that English words are predominantly taught with the use of NL. I observed lessons prior to the pilot study of the vocabulary test and saw the techniques of grammar-translation method used in the classrooms. Words were basically taught with one technique: an unknown word arose from context and the teacher gave the Hungarian meaning. There is a likelihood that the lack of variety in teaching techniques limits the chances of the learners to learn words rapidly and efficiently in the classroom.

## 5.2 PHRASAL VERB FREQUENCY LISTS

Phrasal verbs (PVs) are almost unanimously defined in the literature (Merriam-Webster, 2015; Cambridge Dictionary, 2016) as a phrase (as 'take off' or 'look down on') that combines a verb with a preposition or adverb and that functions as a verb whose meaning differs from the combined meanings of the individual words.

Similar to the BNC and the COCA, attempts have been made to create a reliable pool of phrasal verbs based on frequency. As Garnier and Schmitt (2015) note frequency of occurrence is an appropriate indicator of usefulness. There are variations in the estimation of PVs. McCarthy and Dell (2004) state that there are over 5,000 PVs. Gardner and Davies (2007) go as far as stipulating that there are a total of 12,508 PVs in English. This is an indication of the need to set up frequency lists of PVs.

Gardner and Davies (2007) used the BNC as the only data-source to establish the list of the most frequent PVs. They composed the inventory of the most frequent PVs. They asserted that the top 20 lexical verbs that one finds in PV constructions make up 53.7% of all PVs; moreover they also pointed out that these 20 lexical verbs account for half of the PVs in the BNC when one combines them with only eight particles. It must be noted that Gardner and Davies (2007) emphasized the highly polysemous nature of PVs by pointing out that lexical items on their list have a 5.6 meaning sense average figure.

Liu (2011) endeavored into establishing a brand new inventory of the most frequent PVs by reviewing previous lists including Gardner and Davies (2007). Liu (2011) examined 8,847 PVs on the basis of frequency and ended up incorporating 150 PVs into the list. According to him, there is no significant difference in the use of PVs between British and American English and in contrast to the continuously growing number of words, PV use has remained relatively unchanged over the past decades.

The most recent list of PVs called PHaVE List was compiled by Garnier and Schmitt (2015). First and foremost, they took Gardner and Davies' (2007) and Liu's (2011) list into consideration when analyzing the included PVs. They considered all the PVs used by Gardner and Davies (2007), plus 48 other items extracted by Liu from the COCA. Thus, their list added up to 150 words. Garnier and Schmitt (2015) sincerely admitted to knowing that 150 PVs is not sufficient considering the enormous amount of PVs in English. However, they made the decision that they intended to make the list as useful as possible for learners and teachers of English. So, their limitation of the list to no more than 150 words was supported by a pedagogical perspective motivated by practicality. They also wanted to avoid the list being too long, which is a prerequisite for a frequency list to be truly meaningful (Liu, 2011, p. 667). They further suggested that learning only these PVs would be beneficial for the students (Garnier and Schmitt, 2015, p. 651). As for what information to provide in their PHaVE List, they decided to give the following information: (1) meaning and the connotations of these words; (2) meaning sense frequency percentages; and (3) example sentences. The ordering of the items, similar to Liu's (2011) was made by frequency order. The sources they used to compile the PHaVE list besides Gardner and Davies (2007) and Liu (2011) were well-known and established dictionaries (e.g., Oxford Dictionaries, Merriam-Webster, Collins COBUILD Phrasal Verbs Dictionary, etc.) and a lexical database (WordNet Search 3.1). They noted that PV dictionaries include more sophisticated and refined distinctions than general dictionaries. As for the corpus they chose, it was the COCA

since it offered the researchers the four following advantages: (1) it is very large; (2) it is balanced across several genres; (3) it is frequently updated; and (4) it is freely accessible. Following the corpus analysis procedure and an inter-rater reliability procedure, the final list was compiled. An example is provided here with the lexical item ‘work out’ (Table 21).

**TABLE 21.** *Example of the word ‘work out’ in the Garnier and Schmitt (2015) PHaVe List*

Meaning sense	Sample sentence
Plan, devise or think about STH carefully or in detail (33%)	We still need to work out the details of the procedure.
Exercise in order to improve health or strength (23%)	He works out at the gym 5 times a week.
(+well/badly) Happen or develop in a particular way (15%)	Everything worked out well in the end.
Prove to be successful (12.5%)	Despite our efforts, it just didn’t work out.

### 5.3 SELECTING THE ITEMS FOR THE COPHAVE TEST FOR YLS

First and foremost, I contacted Garnier and Schmitt’s PHaVe List (2015) upon compiling the items for the CONDITEYOLE. This decision was supported by three main reasons: (1) there is no more profound and carefully compiled list of PVs than that of the two researchers; (2) the list was created from a pedagogical perspective; (3) the list comprises the most frequent PVs in English, several of which YLs can be expected to know.

Garnier and Schmitt (2015) go into detail about the polysemous nature of PVs. They argue that no-one knows anything whether students are aware of the different meaning senses of the polysemous PVs. Polysemy is problematic in the assessment of PV knowledge, especially in case of YLs since they are not supposed to know

synonyms and multiple connotations of PVs. In this respect, I made the decision to include exclusively the most frequent meaning sense of the PVs.

Selecting the items from the PHaVE List would have been hardly sufficient. Course-books that YLs use had to be consulted and lessons had to be observed to explore the amount of PV knowledge that can be expected from the YLs. Since the new diagnostic test assessing YLs' PV knowledge will first be used amongst Hungarian 6<sup>th</sup> graders, it was necessary to check which PVs they were expected to learn from 4<sup>th</sup> grade until 6<sup>th</sup> grade. Four teachers of English in four different Hungarian primary schools were called upon to list the course-books they use with their students that start learning in 4<sup>th</sup> grade. Plus, they were also requested to underline the PVs that they have taught or the students are likely to know during the three years of the YLs' English studies. Apart from their help, I also checked all the course-books listed by the four teachers and made a list to myself which I eventually compared with the lists of the four teachers. In case a PV was listed in at least three of the five lists (four teachers plus me), it was included in a final list. This inventory comprised 88 PVs.

In spring 2016, I also observed eight English lessons of 6<sup>th</sup> graders in four primary schools in Szeged, Hungary and paid attention to which PVs the teachers mentioned in their speech, may it be part of an instruction and I also listened to the students' speaking and took note of the PVs that they used verbally in class. Based on classroom observations, another list was thus created that included all the PVs mentioned in class by either teachers or students. This inventory comprised 42 PVs.

At this point I had three lists at my disposal: (1) Garnier and Schmitt's PHaVE List (2015), (2) the list compiled by four teachers and me, (3) the list grounded on classroom observations. The three inventories were compared and any overlap amongst the latter two lists and the PHaVE list resulted in including that particular item in the ultimate list that would serve as the pool of items for the test battery. This procedure led to 47 PVs. Following this I checked the PHaVE List again and completed the ultimate inventory with the seven most frequent PVs on the PHaVE List that were not included amongst the 47 PVs. Thus, the inventory eventually comprised 54 items.

Similar to Thékes' (2016) procedure, the items were divided into three categories. Three word categories were established on the basis of the BNC list and the amount of occurrence of a particular word in the course-books. The necessity of creating categories is underlined by the fact that major vocabulary tests (Nation, 2001;

Laufer & Nation, 1995) include items selected on the basis of layered word list. Three perspectives served as the basis of classifying words into categories: (1) word frequency based on the PHaVE List, (2) occurrence of the words in course-books used by Hungarian YLs, (3) personal judgement on the assumed difficulty of the word. The process of determining word categories is presented below.

Every PV in the test was given a difficulty index calculated from the sum of the three perspectives. Points were given on a scale of one to three based on the perspectives of classifying the PVs. In terms of each perspective a minimum of one point and a maximum of three points were given to the PVs. One point was the indication of easiness and three points were that of difficulty. Personal judgement was done prior to consulting the PHaVE List and course-book occurrence so that prejudice would be avoided. In case I assumed a word easy, I gave it one point; I gave two points to a word I assumed of average difficulty and three points were given to the words that were supposed to be the most difficult. I conducted my judgement on the basis of fifteen years of experience of teaching EFL. As regards the frequency perspective, the word was given one point if it was among the 2,000 most frequent PVs in the PHaVE List, it was given two points if it was between the 2,000 and the 4,000 most frequent words. In case it was outside the 4,000 most frequent PVs, it was given three points. As concerns course-book occurrence, I consulted the course-books used by the students and investigated my item pool with a focus on how frequently the words appear in the books. I wrote ticks next to the words on my list. Afterwards I counted the ticks and gave points to the PVs in the following way: one point to more than six ticks, two points for a number of ticks between three and five, and finally three points for ticks between one and three.

From the process described above it is clear that each word could be given the minimum of three points and the maximum of nine points. The summed points were considered the difficulty indices of the PVs. Based on these difficulty indices, the categories of the PVs were determined. Category 1 contains the simplest and Category 3 contains the most difficult items.

As mentioned above the final inventory of PVs comprise 54 words. The goal was to create the three categories with equal number of items so that in the analysis of the results the knowledge of the words of different categories could be compared. Table 22 presents the PVs that will be used as items in the tasks of the COPHAVE Test for YLs. The category is also indicated. The list is in alphabetical order.

TABLE 22. *The PVs and their category based on index points*

Phrasal Verb	Category	Phrasal Verb	Category
break out	2	hold up	2
bring back	2	keep on	2
build up	3	look around	2
clean up	3	look back	2
close down	3	look out	2
come back	1	look up	1
come down	2	move in	3
come in	1	move on	2
come on	1	move out	3
come out	1	pick up	1
fill out	2	put back	3
find out	1	put on	3
get back	1	set out	3
get down	3	show up	1
get in	3	shut down	3
get out	1	shut up	2
get up	1	sit down	1
give in	3	slow down	3
give up	1	sort out	3
go ahead	2	stand up	2
go back	1	start out	3
go off	3	step back	2
go on	1	take off	1
-go out	1	throw out	3
go up	2	wake up	2
grow up	1	walk out	2
hang up	3	write down	2

## 5.4 CREATING THE COPHAVE TEST FOR YLS

As presented in Table 10, eighteen items were classified into either of the three categories. As discussed above, a decision was made to create three different tasks in the test battery. Six equivalent tests were developed, each comprising three tasks of the same format. Each task in each test consists of nine items. This means that a test comprises 27 PVs. One PV occurs in three tests out of the six but always in a task that requires knowledge of a different modality. Thus, every PV is assessed in all of the three modalities (meaning recognition, form recognition and form recall). This makes it possible to compare the achievements in the three modalities both at item and at student level.

As pointed out, a 27-item test consists of three tasks and each task involves nine items. Students can reach nine points in each task; thus their maximum possible achievement is 27 points in each test. When composing the tasks, I made sure that every task would contain equal number of items of the three categories. In each task three items were selected from Category 1, three items were selected from Category 2, and three items were selected from Category 3. This way, not only the task-solving modalities can be compared but the categories as well. In sum, equivalency among the tests was assured by selecting equal number of items of each category and by assessing each PV in each modality. Herby one test out of the six is presented..

### TASK 1

*Instruction: Choose the phrasal verb that matches the meaning of the verb in bold. See example.*

*Example: He is sitting now but he will **be on his feet** in a second.*

- A) *bring over*
- B) *get off*
- C) *get down*
- D) *stand up*

*Correct answer: D)*

- 1) Uncle Jim **will go to his house and give us** our lawn-mower.  
A) bring back  
B) get up  
C) get down  
D) go ahead
- 2) Prices usually **increase** year by year  
A) come back  
B) get up  
C) go up  
D) move in
- 3) My mom has travelled to England but will **return** soon.  
A) bring back  
B) come back  
C) give in  
D) go ahead
- 4) The DJ has asked everybody to **move onto the dancefloor**.  
A) bring back  
B) get down  
C) get up  
D) go up
- 5) The football player is on the field; he has to **lift his body** and play on.  
A) get up  
B) go ahead  
C) go off  
D) move in
- 6) Her grandpa has been smoking. He will need to **stop** doing so.  
A) come back  
B) give up  
C) go up  
D) move in

7) – Can I have some of your hamburger?

– Sure, **help yourself**.

- A) get up
- B) give in
- C) go ahead
- D) go up

8) The criminals are threatening the city council that a bomb will **explode** outside their office center.

- A) get down
- B) give in
- C) go ahead
- D) go off

9) The family cannot wait to start to **live in their new house**.

- A) come back
- B) get up
- C) go off
- D) move in

**TASK 2**

*Instruction: Choose the correct preposition after the verb. See example.*

*Example: He is sitting now but he will stand ..... soon.*

- A) in
- B) on
- C) over
- D) up

*Correct answer: A)*

1) I need new information. I have to find ..... what they are up to.

- A) back
- B) on
- C) out
- D) over

- 2) I have been waiting here. I hope I will get ..... within an hour.
- A) back
  - B) down
  - C) in
  - D) over
- 3) I tried to convince the president last night. I hope he will give .....  
and we can have his signature.
- A) back
  - B) down
  - C) in
  - D) over
- 4) The child was acting badly. The teacher asked him to go ..... .
- A) in
  - B) into
  - C) out
  - D) over
- 5) If you fix a meeting on Facebook messenger, you have to show .....  
and talk to your friend.
- A) in
  - B) on
  - C) up
  - D) through
- 6) If your plastic mineral water bottle is empty, you have to throw it .....  
not to pollute the environment.
- A) across
  - B) down
  - C) in
  - D) out

7) Try to remember the cell phone number of your friend. Write it .....  
not to forget.

- A) down
- B) in
- C) into
- D) over

8) Try to walk ..... quietly so the baby will stay asleep.

- A) back
- B) in
- C) out
- D) over

9) When you ride your bike, never look ..... so that you will always  
focus on the road.

- A) back
- B) in
- C) over
- D) through

**TASK 3**

*Instruction: Write the proper preposition into the gaps. See example.*

*Example: When his Iphone rings, He will stand ..... and walk to pick it up.*

*Correct answer: up*

- 1) The airplane will take ..... in a minute and head to Beijing.
- 2) My Blackberry alarm app will wake me ..... at exactly 6 o'clock.
- 3) Sarah will put ..... he new Hello Kitty denim vest this morning.
- 4) When my children grow ....., they will not watch Wild Kratts cartoons.

- 5) The monkey has climbed the tree in the zoo not he has to come .....  
from above.
- 6) Corn flakes is always great to start ..... the day.
- 7) Come ..... boys, we can win this match.
- 8) Vincent talks too much in class, sometimes he needs to shut .....  
so everyone can focus on the lesson.
- 9) We are going to Venice on holiday. On the first day, we set .....  
in the morning.

As for further principles taken into account, three important aspects of the test battery need to be mentioned: (1) the sentences in the tasks were phrased in a context familiar to 6<sup>th</sup> graders and motivating for them, which is an important tenet of language tests for YLs (Nikolov, 2011), (2) in the two receptive tasks, the items to be selected are listed in alphabetical order so that students will not surmise any hidden trace behind the order of listed items; thus validity is strengthened, and (3) the sentences are short but interpretable so that the proper PV can be elicited. Furthermore, the format of Task 1 resembles that of the VST (for description, see above), that of Task 2 takes the model of the diagnostic online English and German receptive vocabulary size test for YLs except that no picture is applied and Task 3 is basically the PVST with the difference that instead of word parts, prepositions must be invoked.

## 5.5 DISCUSSION AND FUTURE RESEARCH

Having six equivalent tests assessing young learners' knowledge of PVs makes it possible to begin the piloting process of the tests. As in my doctoral dissertation (Thékes, 2016) the complex vocabulary was piloted, the same procedure will be taken in case of the PHAVE Test for YLs. The fact that each PV is assessed in three tests out of the six in the test battery, it will be made possible to compare the knowledge of PVs in different modalities. It will be empirically investigated whether form recall is the most difficult modality as Laufer et al. (2004) surmise.

First a paper-and-pencil pilot study will be conducted with the participation of around 120 6<sup>th</sup> graders, each sitting for three of the six tests. Test versions will be assigned randomly. Students' achievements will be examined and besides that, item level analysis of the test results will be conducted with classical statistical procedures and with the Rasch-model. Item-total correlation values will also be investigated to explore how the items behave in the tests. Checking these values, a clear picture will be revealed whether any change needs to be made on the test concerning any of the items.

Once the statistical analysis is finished, the tests will be uploaded onto an online platform called eDia (Molnár, 2014) and a large-scale online assessment will be done with 350 students. Besides the quantitative statistical procedures, think-aloud protocols will also be performed with twenty participating students so that more profound data will be revealed. The PHAVE Test for YLs will be a valid tool assessing Hungarian 6<sup>th</sup> graders; however reproduction studies will also be made possible to be run at an international level.

Having developed a diagnostic test assessing the knowledge of phrasal verbs, I will provide facilitation for teachers' work. Teachers will have a new instrument at their disposal with which they can identify individual differences among students with respect to the acquisition of phrasal verbs. Teachers will, thus be shown the right road to applying more efficient methods to teach PVs.

## END-NOTE

The purpose of this book was to reveal the relevant literature on the assessment of foreign language vocabulary and word learning strategies. Since the assessment of vocabulary has to be based on the results of corpus linguistics, an ample analysis of different corpora is presented. With consideration to the fact that there is a growing trend of assessing the knowledge of collocations and phrasal verbs, focus was cast on the latter. The process of the development of a diagnostic instrument measuring the knowledge of PVs was described. The book was intended to be the sum of my previous research and to shift the attention from individual words onto the potential endeavors to unearth the construct of the assessment of PVs.

### APPENDIX 1. *The vocabulary learning strategies questionnaire*

<b>Dimension</b>	<b>Item in Hungarian</b>	<b>English translation</b>	<b>Previous instrument inclusive of item</b>
Cognitive 1	Az új szót mondatban használok.	I use the new word in a sentence.	Stoffer (1995), Pavičič (2008), Schmitt (1997)
Cognitive 2	Az új szót sokszor leírom.	I write down new word many times.	Pavičič (2008), Schmitt (1997)
Cognitive 3	Az új szót sokszor kimondom.	I say the new word many times.	Stoffer (1995), Schmitt (1997)
Cognitive 4	Szótárfüzetet használok a szavak tanulására.	I use a vocabulary list to learn words	Stoffer (1995), Pavičič (2008), Schmitt (1997)
Cognitive 5	Az újonnan megtanult szót beszédben használok.	I use the newly-learned word in speaking.	Stoffer (1995), Schmitt (1997)
Cognitive 6	Az újonnan megtanult szót írásban használok.	I use the newly-learned word in writing.	Stoffer (1995), Schmitt (1997)

<b>Dimension</b>	<b>Item in Hungarian</b>	<b>English translation</b>	<b>Previous instrument inclusive of item</b>
Cognitive 7	Tárgyakra ráírom vagy ráragasztom az angol jelentésüket.	I write or stick the meaning of words onto objects.	Stoffer (1995), Pavičič (2008)
Cognitive 8	Szójátékokat játszok.	I play word-games.	added item
Memory 1	Elképzelek egy helyzetet, amikor használnám a szót.	I imagine a situation when I would use the word.	Schmitt (1997), Pavičič (2008)
Memory 2	Szólistát csinállok, hogy emlékezzek a szóra.	I make a word list in order to remember it.	Stoffer (1995)
Memory 3	Csoportosítom a szavakat hasonlóságuk alapján.	I group the words in clusters based on their similarities.	Stoffer (1995), Schmitt (1997)
Memory 4	Hasonló jelentésű szóhoz kötöm a megtanulandó szót.	I link the new word to one with synonymous meaning.	Schmitt (1997), Pavičič (2008)
Memory 5	Ellentétes jelentésű szóhoz kötöm a megtanulandó szót.	I link the new word to one with antonymous meaning.	Schmitt (1997)
Memory 6	Az új szót ismert szóhoz kapcsolom.	I link the new word to one already known.	Stoffer (1995), Schmitt (1997)
Memory 7	Képes szókérttyákat készítek.	I make picture word cards.	Stoffer (1995), Pavičič (2008), Schmitt (1997)
Memory 8	Angol magyar szókérttyákat készítek.	I make English–Hungarian word cards.	Stoffer (1995), Pavičič (2008), Schmitt (1997)
Memory 9	Magamban elismétlem a szót.	I repeat the word to myself.	Stoffer (1995), Pavičič (2008), Schmitt (1997)

<b>Dimension</b>	<b>Item in Hungarian</b>	<b>English translation</b>	<b>Previous instrument inclusive of item</b>
Memory 10	A szó mellé képeket rajzolok.	I draw pictures next to the word.	Stoffer (1995)
Memory 11	Felmérem, hogy megtanultam-e az új szót.	I evaluate if I have really learned the word.	Stoffer (1995), Pavičič (2008)
Metacognitive 1	Angol nyelvű zenét hallgatok, hogy új szót tanuljak.	I listen to English music so as to learn new words.	Pavičič (2008)
Metacognitive 2	Aláhúzom a fontos szót.	I underline the important word.	Stoffer (1995), Pavičič (2008), Schmitt (1997)
Metacognitive 3	Bekarikázom a szót, amely fontos.	I circle the word that is important.	Stoffer (1995), Pavičič (2008), Schmitt (1997)
Metacognitive 4	Angol nyelvű filmeket nézek angol felirattal.	I watch English film with English subtitles.	Stoffer (1995), Pavičič (2008)
Metacognitive 5	Angol nyelvű filmeket nézek felirat nélkül.	I watch English films without subtitle.	Stoffer (1995), Pavičič (2008)
Metacognitive 6	Angol nyelvű filmeket nézek magyar felirattal.	I watch English films with Hungarian subtitle.	Stoffer (1995), Pavičič (2008)
Metacognitive 7	Angol nyelvű rajzfilmeket nézek.	I watch English cartoons.	added item
Metacognitive 8	Angol nyelvű újságot olvasok a szó tanulására.	I read English newspapers so as to learn the words.	Stoffer (1995), Pavičič (2008)
Metacognitive 9	Angolul olvasok könyvet.	I read English books.	Pavičič (2008)

<b>Dimension</b>	<b>Item in Hungarian</b>	<b>English translation</b>	<b>Previous instrument inclusive of item</b>
Metacognitive 10	Angol nyelvű számítógépes játékokat játszok.	I play English computer games.	Pavičič (2008)
Metacognitive 11	Angol nyelvű képregényeket olvasok.	I read English comics.	added item
Metacognitive 12	Elovasom az angol nyelvű feliratokat mindenféle termékeken.	I read the English labels on every product.	Stoffer (1995)
Metacognitive 13	Azért használok írásban új szót, hogy emlékezzek rá.	I use a new word in writing so as to remember it.	Stoffer (1995), Schmitt (1997)
Metacognitive 14	Azért használok beszédemben új szót, hogy emlékezzek arra.	I use a new word in my speaking so as to remember it.	Stoffer (1995), Schmitt (1997)
Metacognitive 15	Elemzem egy új szó részeit, hogy rájőjjek a jelentésére.	I analyze the meaning of new words so as to realize its meaning.	Stoffer (1995), Pavičič (2008), Schmitt (1997)
Metacognitive 16	Olvasáskor a szövegkörnyezetből következtetem ki a szó jelentését.	I infer the meaning of the new word from context when reading.	Stoffer (1995), Pavičič (2008)
Metacognitive 17	Angol nyelvű beszédből következtetem ki a szó jelentését.	I infer the meaning of the new words from spoken English.	Stoffer (1995), Pavičič (2008), Schmitt (1997)
Determination 1	Nyomatott szótárból keresem ki az új szó jelentését.	I look up the meaning of the new word in a printed dictionary.	Stoffer (1995), Pavičič (2008), Schmitt (2008)

<b>Dimension</b>	<b>Item in Hungarian</b>	<b>English translation</b>	<b>Previous instrument inclusive of item</b>
Determination 2	Elektronikus szótárból keresem ki a jelentést.	I look up the meaning of the word in an electronic dictionary.	added item
Determination 3	Megjegyzem hol láttam az új szót a tankönyv oldalán.	I remember where I have seen the new word on the page of the textbook.	added item
Determination 4	Megjegyzem hol hallottam az új szót.	I remember where I have heard the new word.	added item
Determination 5	Angol–magyar szótárból nézem ki az új szót.	I look up the new word in an English–Hungarian dictionary.	Schmitt (1997)
Determination 6	Az új szót egynyelvű angol szótárból nézem ki.	I look up the new word in a monolingual dictionary.	Pavičič (2008)
Determination 7	Pórbálok az új angol szó magyar megfelelőit is megjegyezni.	I try to remember the Hungarian equivalent of the new English words.	Pavičič (2008)
Social 1	A tanárt kérdezem meg, mit jelent az új szó.	I ask the teacher what the new word means.	Schmitt (1997), Pavičič (2008)
Social 2	Osztálytárssal tanulom az új szót.	I learn the new word with a classmate.	Schmitt (1997), Pavičič (2008)
Social 3	Órán, a társam kérdezem meg, mit jelent az új szó.	I ask my classmate in class what the new word means.	Schmitt (1997), Pavičič (2008)

<b>Dimension</b>	<b>Item in Hungarian</b>	<b>English translation</b>	<b>Previous instrument inclusive of item</b>
Social 4	A szüleim kikérdezik tőlem a szavak jelentését.	My parents check if I have learned the new words by asking me.	added item
Social 5	Órán csoportmunkában együtt tanuljuk a szavakat.	We learn the new words together in group work in class.	Stoffer (1995), Pavičić (2008)
Social 6	Angolul tudó barátot keresek a közösségi oldalakon.	I look for English speaking friends on the social network sites.	added item
Social 7	Angolul használok a facebookot.	I use facebook in English.	added item
Social 8	Angolul twitterezek.	I use twitter in English.	added item
Social 9	Angolul skypeolok.	I skype in English.	added item

**APPENDIX 2.** *The newly-developed self-reported vocabulary learning strategies questionnaire*

\* *F – Functional, T – Traditional, A – Authentic, I – Innovative;*

\*\* *1 – never, 2 – once a month, 3 – once a week, 4 – always*

<b>Factor*</b>	<b>Item in Hungarian</b>	<b>Item in English</b>	<b>How often do you do these activities to learn words?***</b>
A	Szólistát csinállok, hogy emlékezzek a szóra.	I make a word list to remember the words.	1 2 3 4
I	Angol–magyar szó-kártyákat készítek.	I make English–Hungarian word cards.	1 2 3 4

<b>Factor*</b>	<b>Item in Hungarian</b>	<b>Item in English</b>	<b>How often do you do these activities to learn words?***</b>			
F	Aláhúzom a fontos szót a szövegben.	I underline the important words.	1	2	3	4
F	Bekarikázom azt a szót a szövegben, amit fontosnak tartok.	I circle the word that is important.	1	2	3	4
A	Angol nyelvű újságot olvasok a szavak tanulása céljából.	I read English newspapers to learn words.	1	2	3	4
T	Megjegyzem hol láttam az új szót a tankönyv oldalán.	I remember the page where I have seen the new word.	1	2	3	4
F	Azért használok beszédemben új szót, hogy emlékezzek arra.	I use the newly-learned word in speaking to remember it.	1	2	3	4
F	Az új szót mondatban használom.	I use new word in a sentence.	1	2	3	4
A	Angolul használom a Facebookot, hogy angol szavakat tanuljak.	I use Facebook to learn English words.	1	2	3	4
I	Szójátékokat játszok.	I play with word games.	1	2	3	4
F	Hasonló jelentésű szóhoz kötöm a megtanulandó szót.	I link new word to one with synonymous meaning.	1	2	3	4
I	Elektronikus szótárból keresem ki a szó jelentését.	I look up the word in an electronic dictionary.	1	2	3	4
T	Az új szót egynyelvű angol szótárból nézem ki.	I look up the new word in a monolingual dictionary.	1	2	3	4

Factor*	Item in Hungarian	Item in English	How often do you do these activities to learn words?***			
F	Órán a társam kérdezem meg, mit jelent az új szó.	I ask my classmate in class what the new word means.	1	2	3	4
A	Angolul skypeolok, hogy angol szavakat tanuljak.	I use skype to learn English words.	1	2	3	4
F	Az új szót ismert szóhoz kapcsolom.	I link new word to one already known.	1	2	3	4
I	Felmérem, hogy megtanultam-e az új szót.	I evaluate if I have really learned the word.	1	2	3	4
I	Elemzem egy új szó részeit, hogy rájöjjek a jelentésére.	I analyze parts of the word in order to find out its meaning.	1	2	3	4
A	Angol nyelvű beszédből következtetem ki a szó jelentését.	I infer the meaning of the new words from spoken English.	1	2	3	4
T	Próbálom az új angol szó magyar megfelelőjét is megjegyezni.	I try to remember the Hungarian equivalent of the new English words.	1	2	3	4
T	Szavakat azért tanulok meg, hogy könnyebben kommunikáljak.	I learn new words to communicate better.	1	2	3	4
A	Angol nyelvű filmeket nézek magyar felirattal, hogy szavakat tanuljak meg.	I watch English films with Hungarian subtitles to learn new words.	1	2	3	4
A	Angol nyelvű zenét hallgatok, hogy új szót tanuljak.	I listen to English music in order to learn new words.	1	2	3	4
A	Angol nyelvű filmeket nézek felirat nélkül.	I watch English films without subtitles.	1	2	3	4
A	Angolul olvasok könyvet.	I read English books.	1	2	3	4

Factor*	Item in Hungarian	Item in English	How often do you do these activities to learn words?***			
I	Angol nyelvű számítógépes játékokat játszok.	I play English video games.	1	2	3	4
T	Elovasom az angol nyelvű feliratokat mindenféle termékeken.	I read English labels on all kinds of products to learn new words.	1	2	3	4
T	Olvásáskor a szövegkörnyezetből következtetem ki a szó jelentését.	I infer the meaning of the new word from context when reading.	1	2	3	4
F	Angolul tudó barátot keresek a közösségi oldalakon.	I look for English speaking friends in the social media.	1	2	3	4
F	Az újonnan megtanult szót írásban használom.	I use the newly-learned word in writing.	1	2	3	4
A	Mikor angol nyelvű műsort nézek/hallgatok jegyzetlem a szavakat.	I take notes of the words when watching/listening to English programs.	1	2	3	4
F	Az újonnan megtanult szót beszédben használom.	I use a new word in speaking so as to remember it.	1	2	3	4
I	Képes szókétyákat készítek.	I make picture word cards.	1	2	3	4
T	Magamban elismétem a szót.	I repeat the word to myself.	1	2	3	4
T	Kétnyelvű szótárból nézem ki a szó jelentését.	I look up the meaning of the new words in a bilingual dictionary.	1	2	3	4
T	Saját szótárfüzetből tanulom a szavakat.	I learn new words from my own vocabulary.	1	2	3	4
T	Bemagolom a szavakat.	I rote-learn the words.	1	2	3	4

**Appendix 3.** *Ranks, frequencies and categories of words*

<b>Item</b>	<b>Task</b>	<b>BNC Rank</b>	<b>Frequency</b>	<b>Category in the test battery</b>
monkey	1	5,317	1,067	2
lion	1	3,722	1,828	1
airplane	1	2,002	4,505	1
tram	1	5,878	722	2
swimming	1	5,861	906	2
helicopter	1	4,240	1,517	1
ship	1	1,384	6,974	1
camel	1	2,912	658	3
skating	1	6,200	421	3
supermarket	2	4,052	1,621	2
theatre	2	1,882	4,917	2
bake	2	5,773	930	3
cinema	2	3,461	2,026	2
eat	2	662	15,446	1
hospital	2	604	16,898	2
learn	2	432	23,394	1
play	2	245	38,053	1
sell	2	494	20,902	1
boat	3	1,317	7,373	2
drink	3	1,129	8,926	2
drive	3	618	16,477	1
heavy	3	970	10,439	1
leg	3	858	11,858	1
hit	3	998	10,098	1
pocket	3	1,809	5,172	2
quick	3	5,817	918	3

<b>Item</b>	<b>Task</b>	<b>BNC Rank</b>	<b>Frequency</b>	<b>Category in the test battery</b>
small	3	183	51,626	1
busdriver	4	1,264	7,806	1
waiter	4	5,534	998	2
cook	4	4,199	1,541	2
fire(fighter)	4	719	14,379	1
hair(dresser)	4	682	15,020	1
mechanic	4	3,543	1,948	2
pilot	4	2,159	4,117	1
plumber	4	6,422	322	3
tailor	4	6,826	280	3
bedroom	5	1,626	5,865	1
cook	5	4,199	1,541	2
cup	5	831	12,294	1
curtain	5	2,621	3,119	2
dining room	5	6,068	853	3
open	5	392	25,614	1
shelf	5	3,078	2,419	2
talk	5	310	30,930	1
wash	5	1,854	5,027	1
cake	6	2,299	3,773	1
cheese	6	2,783	2,864	1
chicken	6	3,072	2,426	1
coffee	6	1,461	6,614	1
fish	6	1,017	9,901	1
hotdog	6	no data	no data	3
(ice)cream	6	2,930	2,638	2
cucumber	6	6,800	780	3
sausage	6	5,560	990	2

## REFERENCES

- Abduallah, K. I., Puteh, F., Azizan, A. R., Hamdan, N. N., & Saude, S. (2013). Validation of a controlled productive vocabulary levels test below the 2000-word level. *System*, 21(2), 352–364. doi:10.1016/j.system.2013.03.005
- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Alderson, J. C. & Huhta, A. (2005). The development of a suite computer-based diagnostic test based on the Common European Framework. *Language Testing*, 22(3), 301–320. doi: 10.1191/0265532205lt310oa
- Atay, D., & Kurt, G. (2006). Elementary school EFL learners' vocabulary learning: The effects of post-reading activities. *The Canadian Modern Language Review*, 63(2), 255–273. doi: 10.1177/1362168808089921
- Bachmann, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachmann, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Baumann, J. F., Kame'enui, E. J., & Ash, G. (2003). Research on vocabulary instruction: Voltaire redux. In J. Flood, D. Lapp, J. R. Squire, & J. Jensen (Eds.), *Handbook of research on teaching the English language arts* (pp. 752–785). Mahway, NJ: Lawrence Erlbaum.
- Bialystok, E. (1990). *Communication strategies: A psychological analysis of second-language use*. Oxford: Basil Blackwell.
- Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.
- Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effectiveness of exercises on verb-noun collocations. *Language Teaching Research*, 18(1), 50–70. doi:10.1177/1362168814541752
- Bogaards, P. (2000). Testing L2 vocabulary knowledge at a high level: the case of the Euralex French Tests. *Studies in Second Language Acquisition* 21(2), 128–160.
- Bossers, B. (1992). *Reading in two languages: a study of reading comprehension in Dutch as a second language and in Turkish as a first language*. Rotterdam: Drukkerij Van Driel.

- Butler, Y, Someya, Y., & Fukuhara, E. (2014). Online games for YLs' foreign language learning. *English Language Teaching Journal*, 68(3), 265–275. doi: 10.1093/elt/ccu008
- Cameron, L. (2004). *Teaching languages to young learners*. Cambridge: Cambridge University Press.
- Carroll, J. B., & Sapon, S. (2002). *Modern Language Aptitude Test: Manual 2002 edition*. Bethesda, MD: Second Language Testing, Inc.
- Chesterfield, R., & Chesterfield, K. (1985). Natural order in children's use of second LLS. *Applied Linguistics*, 6(1), 45–59.
- Chostelidou, D., Griva, E., Ioannidis, T., & Panitsidou, E. (2012). Multilingual learning for specific purposes: Identifying language strategies, awareness and preferences. *Procedia*, 46, 1419–1423.
- Cohen, A. (1990). *Language learning: Insights for learners, teachers, and researchers*. New York: Newbury House/Harper Collins.
- Cohen, A. (2003). The learner's side of foreign language learning: Where do styles, strategies, and tasks meet? *International Review of Applied Linguistics in Language Teaching*, 41(4), 279–291. doi: 10.1515/iral.2003.013.
- Cohen, A., & Macaro, J. (2007). *Language learner strategies: 30 years of research and practice*. Oxford: Oxford University Press.
- Cohen, L., & Manion, L. (2000). *Research methods in education*. New York: Routledge.
- Common European Framework of Reference for Languages: learning, teaching, assessment*. (2002). Cambridge: Cambridge University Press.
- Coyle, Y., & Gomez Gracia, R. (2014). Using songs to enhance L2 vocabulary acquisition in preschool children. *ELT Journal*, 68(3), 276–285. doi: 13.1138/0122-05637.82.1.463
- Csizér, K., & Dörnyei, Z. (2005). Language learners' motivational profiles and their motivated learning behaviour. *Language Learning*, 55(4), 613–659. doi: 10.1111/j.0023-8333.2005.00319
- Davies, M. (2010). *The orpus of contemporary American English*. Provo, Utah: Brigham Young University Press.
- Doró, K., & Habók, A. (2013). Language learning strategies in elementary school: The effect of age and gender in an EFL context. *Journal of Linguistics and Language Teaching*, 4(2), 25–37.

- Drum P., & Konopak B. (1987). Learning word meanings from written context. In B. McKeown, M., Curtis (Eds.), *The nature of vocabulary development* (pp. 73–87). Lawrence Hillsdale, NJ.: Erlbaum Pub.
- Enever, J. (2011). The ELLiE study: Policy. In J. Enever (Ed.), *ELLiE. Early Language Learning in Europe* (pp. 23–42). London: British Council.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, 43(4), 559–617. doi: 10.1111/j.1467-1770.1993.tb00627.
- Everitt, B. S. (2002). *The Cambridge dictionary of statistics*. Cambridge: Cambridge University Press.
- Fazekas, M. (2009). *Felmérés a 2008-ban Magyarországon érettségizettek idegennyelv-tudásáról*. Budapest: OKM (Manuscript).
- Field, A. (2005). *Discovering Statistics Using SPSS*. 2nd ed. London: Sage.
- Fitzpatrick, T., Al-Qarni, I., & Meara, P. (2008). Intensive vocabulary learning: a case study. *Language Learning Journal*, 36(2), 239–248. doi: 10.4304/jltr.4.2.209-219
- Fontecha, A.F. (2014). Receptive vocabulary knowledge and motivation in CLIL and EFL. *RLyLA*, 9, 23–32.
- Gardner, R. C., & MacIntyre, P. D. (1993). A student's contributions to second language learning. Part II: Affective variables. *Language Teaching*, 26, 1–11.
- Garnier, M., & Schmitt, N. (2016). Picking up polysemous phrasal verbs: How many do learners know and what facilitates this knowledge? *System*, 59(1), 29–44.
- Gass, S. (1997). *Input, interaction and the second language learner*. Mahwah, NJ.: Lawrence Erlbaum.
- Goldstein, B. (2004). *Bilingual language development and disorders in Spanish-English speakers*. Baltimore, Maryland: Paul H. Brooks Publishing Co.
- Goulden, R., Nation, I. S. P., & Read, J. (1990.) How large can a receptive vocabulary be? *Applied Linguistics*, 11(3), 341–363. doi: 10.1093/applin/11.4.341.
- Grainger, J., & Dijkstra, A. (1992). On the representation and use of language information in bilinguals. In R. J. Haris (Ed.), *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: Elsevier Science Publishers.

- Griva, E., Kamaroudis, S., & Geladari, A. (2009). YLs' vocabulary strategies employment during second language learning. *Synergies*, 2, 21–36.
- Gunning, P. (1997). *The learning strategies of beginning ESL learners at the primary level*. MA dissertation. Montréal: Concordia University.
- Gu, Y., & Johnson, R. K. (1996). *VLS and language learning outcomes*. University of Hong Kong.
- Hardi, J. (2014). *Assessing YLs' strategic L2 vocabulary learning in the framework of self-regulation*. PhD dissertation. Szeged: University of Szeged.
- Harding, L., Alderson, J. C. & Brunfaut, T. (2015). Diagnostic assessment of reading and listening in a second or foreign language: Elaborating on diagnostic principles. *Language Testing*, 32(2), 317–336. doi: 10.1177/0265532214564505
- Henrikssen, B., Albrechtsen, D., & Haastrup, K. (2004). The relationship between vocabulary size and reading comprehension in the L2. In D. Albrechtsen, K. Haastrup & B. Henriksen (Eds.), *Writing and Vocabulary in Foreign Language Acquisition* (pp. 129–140). Copenhagen: Museum Tusulanum.
- Hofland, K., & Johansson, S. (1982). *Word frequencies in British and American English*. London: Longman.
- Horváth, J. (2001). *Advanced writing in EFL. A corpus-based study of processes and products*. Pécs: Lingua Franca Csoport.
- Huckin, J., & Bloch, T. (1993). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 21(2), 181–193. doi: 10.1016/j.esp.2015.04.002.
- Hulstijn, J. H. (1992). *De ideale taalleerder. Handelingen elfde colloquium Neerlandicum* (pp. 51–66). Den Haag: Internationale Vereniging voor Neerlandistiek.
- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–86). Cambridge: Cambridge University Press.
- Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327–339. doi: 10.1111/j.1540-4781.1996.tb01614
- Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(4), 539–558. doi: 10.1111/0023-8333.00164

- Ishii, T. & Schmitt, N. (2009). Developing an integrated diagnostic test of vocabulary size and depth. *RELC Journal*, 40(1), 5–22. doi: 10.1177/0033688208101452
- Jackson, H., & Zé Amvela, E. (2011). *Words, meaning, vocabulary. An introduction to English lexicology*. London: Bloomsbury.
- Jiang, E. E. (2014). Assessing English language learners in K-12 Schools. *Education Matters Journal of Ontario Institute for Studies in Education*, 2(1), 72–80.
- Jimenez Catalan, R. M. (2003). Sex differences in L2 VLS. *International Journal of Applied Linguistics*, 13(1), 54–77. doi: 10.1093/applin/21.1.47.
- Jimenez Catalan, R. M., & Terrazas Gallego, M. (2008). The receptive vocabulary of English foreign language YLs. *Journal of English Studies*, 5(1), 173–191.
- Jóhannsdóttir, Á. (2010). English in the 4th grade in Iceland. Exploring exposure and measuring vocabulary size of 4<sup>th</sup> grade students. *Menntakvika 1*(1), 1–20.
- Koda, J. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Language Annals*, 22(6), 529–540. doi: 10.1111/j.1944-9720.1989.tb02780
- Laufer, B. (1997). What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: description, acquisition and pedagogy* (pp. 140–155). Cambridge: Cambridge University Press.
- Laufer, B. (2001). Quantitative evaluation of vocabulary: what it is good for and how it can be done. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara, & K. O'Loughlin (Eds.), *Experimenting with Uncertainty* (pp. 241–250). Cambridge: Cambridge University Press.
- Laufer, B. (2005). Instructed second language vocabulary learning: The fault in the 'default hypothesis. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language learning* (pp. 311–329). Berlin: Walter De Gruyter.
- Laufer, B., & Nation, I. S. P. (1999). A vocabulary size test of controlled productive ability. *Language Testing*, 16(1), 33–51. doi: 10.1177/026553229901600103
- Laufer, B., & Nation, I. S. P. (2001). Passive vocabulary size and speed of recognition. *EUROSLA Yearbook 1*, 7–28.
- Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge? *Language Testing*, 21(2), 202–226. doi: 10.1191/0265532204lt277oa.

- Laufer, B., & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15–30. doi: 10.1002/tesq.140..
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26. doi: 10.1093/applin/22.1.1.
- Lehmann, M. (2009). *Assessing English majors' vocabulary at the University of Pécs*. PhD thesis. Pécs: University of Pécs.
- Lengyel, M. T. (2006). Gyakorisági szótárak. Magyarországi helyzetkép. [Word frequency dictionaries. Hungarian perspective]. *Könyvtári Figyelő*, 52(1), 45–58.
- Lewis, M. (1993). *The lexical approach*. Cambridge: Cambridge University Press.
- Lin, L. (2001). *Taiwanese children's EFL vocabulary learning strategies*. MA thesis. Hsinchu, Taiwan: National Chinghua University.
- McDonough, S. (1999). Learner strategies. *Language Teaching Journal*, 32(1), 1–18. doi: 10.1017/s0261444800013574
- McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.
- Meara, P. (1992). Vocabulary in a second language. *Reading in a Foreign Language*, 9(1), (Complete issue).
- Meara, P. (2009). *Connected words*. Amsterdam: John Benjamins Publishing.
- Meara, P., & Buxton, B. (1987). An alternative to multiple choice vocabulary tests. *Language Testing*, 4(2), 142–154.
- Meara, P., & Milton, J. (2003). *X-Lex: The Swansea Vocabulary Levels Test*. Newbury: Express Publishing.
- Melka Teichroew, F.J. (1982). Receptive vs. productive vocabulary: a survey. *Interlanguage Studies Bulletin* 6(2), 5–33.
- Merriam-Webster.com*. 2015. <http://www.merriam-webster.com> (8<sup>th</sup> April, 2015).
- Mihaljević Djigunovic, J. (2010). Starting age and L1 and L2 interaction. *International Journal of Bilingualism* 14(3), 303–314. doi:10.1177/1367006910367847
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol: Multilingual Matters.
- Milton, J. & Fitzpatrick, T. (2014). *Dimensions of vocabulary knowledge*. Basingstoke: Palgrave Macmillan.

- Molnár, Gy. (2013). Számítógépes játékon alapuló képességfejlesztés: egy pilot vizsgálat eredményei. [Computer game-based skill development. Results of a pilot study]. *Iskolakultúra*, 21(4), 3–12.
- Molnár, Gy. (2015). A képességmérés dilemmái: a diagnosztikus mérések (eDia) szerepe és helye a magyar közoktatásban. [The dilemmas of ability assessment: the role and place of diagnostic (eDia) in Hungarian public education]. *Génius Műhely Kiadványok*, 2, 16–29.
- Mukarto, F. X. (2005). *The patterns of semantic mapping development of English verbs acquired by Indonesian*. PhD dissertation. Sains Malasia: University Sains Malasia.
- Nagy, J. (2004). A szóolvasó készség fejlődésének kritériumorientált diagnosztikus feltérképezése. [The criterion-referenced, diagnostic mapping of the development of word-reading skills]. *Magyar Pedagógia*, 104(2), 123–142.
- Nagy, W., Anderson, R., & Hermann, P. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24(2), 237–270.
- Nagy, W., & Scott, J.A. (2000). Vocabulary processes. In M. L. Kamil; P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research* (pp. 269–284). Mahwah, NJ: Erlbaum.
- Nassaji, H. (2003). L2 vocabulary learning from context: strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, 37(4), 645–670. doi: 10.1017/S0958344008000219.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle and Heinle.
- Nation, I. S. P. (1995). Review of the lexicon in acquisition by E.V. Clark. *Applied Linguistics*, 16(3), 403–405.
- Nation, I. S. P. (1999). Fluency and accuracy. In B. Spolsky (Ed.), *Concise encyclopedia of educational linguistics* (pp. 611–628). Oxford: Elsevier Science.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2011). Vocabulary research into practice. *Language Teaching*, 44(4), 529–539. doi:10.1017/S0261444811000267.
- Nation, I. S. P. (2013). *Learning vocabulary in another language*. (Second edition). Cambridge: Cambridge University Press.
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.

- Nation, I. S. P., & Macalister, J. (2010). *Language curriculum design*. New York: Routledge.
- Nation, I. S. P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12(2), 355–380.
- Nikolov, M. (1999a). Hungarian children's learning strategies. *Strani Jezici Zagreb*, 28(3–4), 225–233.
- Nikolov, M. (1999b). 'Why do you learn English?' 'Because the teacher is short.' A study of Hungarian children's foreign language learning motivation. *Language Teaching Research*, 1999(3), 33–56.
- Nikolov, M. (2003). Hatodikosok stratégiahasználatára olvasott szöveg értését és íráskészséget mérő feladatokon angol nyelvből. [6th graders' strategy use on tasks assessing reading comprehension and writing skills in English]. *Magyar Pedagógia*, 103(1), 5–34.
- Nikolov, M. (2007). A magyarországi nyelvoktatás-fejlesztési politika: Nyelvoktatásunk a nemzetközi trendek tükrében. [Language education development policy in Hungary] In I. Vágó (Ed.), *Fókuszban a nyelvtanulás* [Language learning in focus] (pp. 43–72.) Budapest: Oktatáskutató és Fejlesztő Intézet.
- Nikolov, M. (2008). "Az általános iskola, az módszertan!" Alsó tagozatos angolórák empirikus vizsgálata. ["primary school is real methodology". Empirical investigation of English classes in lower primary school]. *Modern Nyelvoktatás*, 10(1–2), 3–19.
- Nikolov, M. (2011). Az angol nyelvtudás fejlesztésének és értékelésének keretei az általános iskola első hat évfolyamán. [Framework of the development and assessment of English language knowledge in the first six years of primary school] *Modern Nyelvoktatás*, 13(1), 9–32.
- Nikolov, M. (2011). Minőségi nyelvoktatás – a nyelvek európai évében. [Quality language education – in the European year of languages]. *Iskolakultúra*, 11(8), 3–12.
- Nikolov, M., & Józsa, K. (2003). *Idegen nyelvi készségek fejlettsége angol és német nyelvből a 6. és 10. évfolyamon a 2002/2003-as tanévben*. [Foreign language skills in English and German in sixth and tenth grade in the school year of 2002/2003]. Budapest: Országos Közoktatási Értékelési és Vizsgaközpont.
- Nikolov, M., & Mihaljevic Djigunovic (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual Review of Applied Linguistics*, 26, 234–260. doi: 10.1017/S0267190506000122.

- Nikolov, M., & Szabó, G. (2011). Az angol nyelvtudás diagnosztikus mérésének és fejlesztésének lehetőségei az általános iskola 1–6. évfolyamán [Possibilities of developing English diagnostic tests for years 1–6 in the primary school]. In B. Csapó & A. Zsolnai (Eds.), *A kognitív és affektív fejlődés diagnosztikus mérése az iskola kezdő szakaszában* (pp. 13–40). Budapest: Nemzeti Tankönyvkiadó.
- Nisbet, D. L., Tindall, E. R., & Arroyo, A. A. (2005). Language Learning Strategies and English proficiency of Chinese university students. *Foreign Language Annals*, 38(1), 100–107. doi: 10.1111/j.1944-9720.2005.tb02457.
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Orosz, A. (2009). The growth of young learners' English vocabulary size. In M. Nikolov (Ed.), *Early learning of modern foreign languages. Processes and outcomes* (pp. 181–195). Bristol: Multilingual matters, 181–195.
- Oxford, R. (1991). *Language learning strategies: What every teacher should know*. New York: House/Harper and Row
- Paradis, M. (2004). *A neurolinguistics theory of bilingualism*. Amsterdam: John Benjamins.
- Paribakht, T.S., & Wechse, M. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal* 11(1), 929–942.
- Paribakht, T.S., & Wechse, M. (1999). Reading and incidental L2 vocabulary acquisition. An introspective study of lexical inferencing. *Studies in Second Language Acquisition*, 21(2), 195–224.
- Paribakht, T.S., & Wechse, M. (2006). Lexical inferencing in L1 and L2: Implications for vocabulary instruction and learning at advanced levels. In H. Byrnes, D. Weger-Guntharp, & K. A. Sprang (Eds.), *Educating for Advanced Foreign Language Capacities: Constructs, Curriculum, Instruction, Assessment* (pp. 118–135). Washington, DC: Georgetown University Press.
- Pavičić, T. V. (2008). *Vocabulary learning strategies and foreign language acquisition*. Clevedon: Multilingual Matters.
- Poole, A. (2011). *Concordance-based glosses for facilitating semantization and enhancing productive knowledge of academic vocabulary*. PhD dissertation. Montgomery: University of Alabama.

- Poplack, S. The care and handling of a mega-corpus. In Fasold, R. & Schiffrin D. (Eds.), *Language Change and Variation* (pp. 411–451). Amsterdam: Benjamins. 1989.
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning* 52(3), 513–536. doi: 10.1111/1467-9922.00193
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Sayer, P., & Ban, R. (2014). Young EFL students' engagements with English outside the classroom. *ELT Journal*, 68(3), 321–329. doi: 10.1093/elt/ccu013.
- Schifko, M. (2001). Prüfungen, Zertifikate, Abschlüsse als Planungskategorien für den Unterricht. In: G. Helbig (Ed.), *Deutsch als Fremdsprache: ein internationales Handbuch. Handbücher zur Sprach- und Kommunikationswissenschaft* (pp. 827–834). Hamburg: Schritte GmbH.
- Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 198–227). Cambridge: Cambridge University Press.
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (2008). Instructed Second Language Vocabulary Learning. *Language Teaching Research*, 12(3), 329–363. doi: 10.1017/S0261444815000075.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. New York: Palgrave Press.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913–951. doi: 10.1111/lang.12077.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503. doi: 10.1017/S0261444812000018.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing* 18(1), 55–88. doi: 10.1177/026553220101800103.
- Schmitt, N., & Zimmerman, C. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, 36(2), 145–171.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503.

- Schmitt, N., Jiang, X., & Grabe, W. (2011). The %age of words known in a text and reading comprehension. *Modern Language Journal*, 95(1), 26–43.
- Schoonen, R., & Verhallen, M. (2008). The knowledge of words and its assessment. *Pedagogische Studiën*, 75(3), 153–168.
- Schouten-Van Parreren, C. (1992). Individual differences in vocabulary acquisition: A qualitative experiment in the first phase of secondary education. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 94–101). Basingstoke: Macmillan.
- Scott, V., & de la Fuente, M. J. (2008). What's the Problem? L2 Learners' use of the L1 during consciousness-raising, form-focused tasks. *The Modern Language Journal*, 92(1), 100–113.
- Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge: Cambridge University Press.
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in EFL. *Studies in Second Language Acquisition*, 31(4), 577–607. doi: 10.1017/s0272263109990039.
- Stoffer, J. (1995). University foreign language students' choice of vocabulary learning strategies as a related to individual difference variables. *Journal of Education of the University of Alabama*, 4, 122–153.
- Szpotowicz, M., & Lindgren, E. (2011). Language achievements: a longitudinal perspective. In J. Enever (Ed.), *ELLiE. Early Language Learning in Europe* (pp. 125–144). London: British Council.
- Szpotowicz, M., & Szulc-Kurpaska, M. (2009). *Teaching English to young learners*. Warszawa: Wydawnictwo Szkolne PWN.
- Szulc-Kurpaska (2001). *Foreign language acquisition in the primary school: teaching and strategy training*. Wrocław: Atut.
- Thékes, I (2016). *Assessing young Hungarian EFL word knowledge and learning strategies*. University of Szeged. Unpublished dissertation.
- Thornbury, S. (2002). *How to teach vocabulary*. London: Pearson.
- Thornbury, S. (2004). Big words, small grammar. *ELT Professional*, 31(1), 10–11
- Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, 27(1), 78–102. doi:10.1093/applin/ami046.

- Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4, 105–122. doi: 10.1080/02643294.2014.882814.
- Vidákovich T., Vígh T., Sominé Hrebik O., & Thékes I. (2013). Az angol és német nyelvi szókincs online diagnosztikus tesztelése a 6. évfolyamon. (Diagnostic assessment of English and German as a foreign language vocabulary amongst 6<sup>th</sup> graders). *Iskolakultúra*, 23(11), 117–131.
- Webb, S., & Chang, A, C-S. (2012). Vocabulary learning through assisted and unassisted repeated reading. *Canadian Modern Language Review*, 68(3), 267–290. doi: 10.3138/cmlr.1204.1.
- Webb, S. (2008). Receptive and productive vocabulary size. *Studies in Second Language Acquisition*, 30(1), 79–95. doi:10.1017/S0272263108080042.
- Webb, S. (2010). A corpus driven study of the potential for vocabulary learning through watching movies. *International Journal of Corpus Linguistics*, 15(4), 497–519. doi: 10.1075/ijcl.15.4.03.
- Webb, S. (2012). Receptive learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 2783–2785). New York: Springer.
- Webb, S., & Sasao, Y. (2013). New directions in vocabulary testing. *RELC Journal*, 44(3), 263–278. doi: 10.1177/1362168816639619.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315–327). New York: Macmillan.
- West, M. (1953). *A general service list of English words*. London: Longman.
- Wood, E., & Attfield, J. (2005). *Play, learning and the early childhood curriculum*. London: Paul Chapman.
- Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication*, 3(2), 215–229. doi: 10.1177/0265532209349471.
- Zareva, A. (2007). Structure of the L2 mental lexicon: How does it compare to native speakers' lexical organization? *Second Language Research*, 23(2), 123–153. doi: 10.1177/0267658307076543,