UNIVERSIDAD DE EXTREMADURA

TESIS DOCTORAL

Estrategias de aprendizaje y conocimiento léxico: un estudio sobre el alumnado de educación secundaria en programas de Aprendizaje Integrado de Contenidos y Lengua Extranjera

IRENE OLGA CASTELLANO RISCO

PROGRAMA DE DOCTORADO EN LENGUAS Y CULTURAS

2021

TESIS DOCTORAL

Estrategias de aprendizaje y conocimiento léxico: un estudio sobre el alumnado de educación secundaria en programas de Aprendizaje Integrado de Contenidos y Lengua Extranjera

IRENE OLGA CASTELLANO RISCO

PROGRAMA DE DOCTORADO EN LENGUAS Y CULTURAS

La conformidad de los directores de la tesis consta en el original en papel de esta Tesis Doctoral

Rafael Alejo González   Ana María Piquer Píriz

Learning strategies and vocabulary knowledge: a study of secondary-school learners in Content and Language Integrated Learning programmes

Languages define personal identities but are also part of a shared in-heritance. They can serve as a bridge to other people and open access to other countries and cultures, promoting mutual understanding.

European Commission, 2008, p. 3

Consideraciones previas

Este proyecto ha tenido una duración total de más de cuatro años, y sus resultados han sido presentados en diversos congresos tanto nacionales como internacionales, además de haber sido publicados parcialmente en las siguientes publicaciones:

Preliminary considerations

This project has lasted for more than four years, and its results have been presented in several national and international conferences, and partially published in the following papers:

Castellano-Risco, I., Piquer-Píriz, A. M. & Alejo González, R. (2020). The development of receptive vocabulary in CLIL learners: Is it all a matter of exposure? *System*, 91. DOI: 10.1016/j.system.2020.102263

Castellano-Risco, I. (forthcoming). La influencia del enfoque AICLE en el vocabulario académico: el caso de alumnos de Educación Secundaria. Revista Cauce. Revista Internacional de Filología, Comunicación y sus Didácticas, 42.

Castellano-Risco, I. (2019). Las estrategias de aprendizaje de vocabulario: estudio evolutivo de su uso en alumnos AICLE e ILE en Extremadura. *Campo Abierto, Revista de Educación, 38*(1), 19-34.

Castellano-Risco, I. (2019). Vocabulary Learning Strategies: Their relation to learning success. *Lenguajes y Textos 48*, 11-20.

Castellano-Risco, I. (2019). Understanding the selection of vocabulary learning strategies: the impact of the language teaching approach. *Journal of English Studies, 17*, 79-106.

Castellano-Risco, I. (2018). Análisis del uso de estrategias de aprendizaje de vocabulario por alumnos de secciones bilingües de Educación Secundaria. *Campo Abierto, Revista de Educación, 37*(2), 133-144.

Castellano-Risco, I. (2018). The importance of selecting the appropriate instruments when measuring receptive vocabulary size. *Revista Internacional de Lenguas Extranjeras (RILE), 8,* 45-62.

Castellano-Risco, I. (2018). La implementación de AICLE en Extremadura: un estudio preliminar del rol del contexto educativo en la adquisición de vocabulario. In M.E. Gómez-Parra & R. Johnstone (eds.), *Nuevas perspectivas en educación bilingüe: investigación e innovación* (pp. 71-75). Universidad de Granada: Editorial Universidad de Granada.

Acknowledgments

There is an old African proverb which says, "It takes a village to raise a child". The saying stresses that the whole community is the one in charge of the education of the children. This dissertation has been one of the biggest challenges of my life, and, of course, it could not have been accomplished without the support and contribution of several people. In this sense, I am surrounded by a huge community for which I am grateful.

First of all, I would like to express my utmost and most sincere gratitude to my supervisors, Dr Rafael Alejo and Dr Ana María Piquer-Píriz, whose unconditional support, dedication, guidance and criticism have been determining in this long process. You always found the right words to help me to go on.

I must also convey my gratitude to the Joint Action of the University of Extremadura and Banco Santander. The scholar grant I received allowed me to focus on my project and deepen my knowledge of critical aspects.

My acknowledgments should also be extended to Dr. Miguel González Velasco, whose advice on the statistical issues was decisive in the development of this dissertation.

I am also indebted to the Applied Linguistic department of Swansea University, and, in particular to Dr Tess Fitzpatrick and Dr James Milton, for their warm welcome and support. Their work, advice and thorough comments have been a source of inspiration. My time at Swansea will always hold a special place in my heart.

I am also thankful to my fellow colleagues Marta, Margarida, and Raquel, who, in one sense or another, have shared with me this process. It is usually said that PhD studies are a highly demanding experience in one's life, but, for me, they have also been very rewarding, as it has brought into my life beautiful and charming people and friends.

Last but not least, I would like to extend my gratitude to my friends and family and, in particular, to Gabriel. I know it has not been easy to live with an often absent-minded PhD student. Thank you for accompanying me on this journey, especially in the darkest times.

Thank you

Abstract

In recent years, the relevance of lexical competence in SLA has grown in importance (Jiménez Catalán & Terraza Gallego, 2005; Boers & Lindstromberg, 2008; Schmitt, 2008; Milton, 2009; Nation, 2001) together with a developing interest in the strategies students use to learn L2 vocabulary in different educational contexts (Schmitt, 1997; Nation, 2001). In one of these contexts, Content and Language Integrated Learning (CLIL), however, most of the attention has been usually placed on the potential increase of learners' vocabulary as a result of the implementation of the methodologies typically associated to this educational approach (Jiménez Catalán & Ruiz de Zarobe, 2008; Canga Alonso, 2013) and there has been a certain neglect of the analysis of the specific strategies learners use.

This PhD dissertation explores the development of lexical competence (vocabulary knowledge and vocabulary learning strategies [VLS]) in 138 Extremaduran secondary school learners following two educational approaches (CLIL vs mainstream EFL [English as a Foreign Language]). Two sets of tests –the Vocabulary Levels Tests (Schmitt, Schmitt & Clapham; 2001) and the Productive Vocabulary Levels Tests (Laufer & Nation, 1999)– were used to assess the receptive and productive mastery of the 2K and academic vocabulary bands. Finally, to explore the learners' use of strategies, a questionnaire adapting Schmitt's (1997) taxonomy was developed.

Results indicate a clear difference between CLIL and EFL learners as regards both their selection and use of strategies and their vocabulary levels. CLIL learners outperformed EFL learners in the receptive and productive vocabulary tests. Concerning VLSs selection, both groups of learners demonstrated to use a different range of strategies, with CLIL learners selecting significantly more often VLSs related to greater lexical development.

The results of this study may be relevant given the existing gaps regarding how CLIL may (1) influence the way learners face vocabulary learning, and (2) relate to other factors such as Instructed Amount of Exposure (IAoE). The confirmation of the differences between CLIL and mainstream EFL learners not only in general but also in academic vocabulary and the finding that the teaching context affects the way L2 vocabulary is processed in the mind, together with the consideration of the potential influence of IAoE in these findings, may help to shed some light on some of the most contentious CLIL issues questioning the effectiveness of CLIL and its impact on L2 learning.

Resumen

En las últimas décadas, se ha asistido a una mayor consideración de la competencia léxica en el campo de la adquisición de segundas lenguas (Jiménez Catalán & Terraza Gallego, 2005; Boers & Lindstromberg, 2008; Schmitt, 2008; Milton, 2009; Nation, 2001) junto con un creciente interés en la identificación de los mecanismos que los alumnos emplean para aprender léxico de la L2 (Schmitt, 1997; Nation, 2001). Sin embargo, el análisis del impacto de la implementación del enfoque 'Aprendizaje Integrado de Contenido y Lengua Extranjera' (AICLE) se ha concentrado en el potencial incremento del conocimiento léxico del alumnado AICLE (Jiménez Catalán & Ruiz de Zarobe, 2008; Canga Alonso, 2013), dejando en un segundo plano el análisis de cómo este léxico se desarrolla.

Esta tesis doctoral explora el desarrollo de la competencia léxica (entendida como conocimiento de vocabulario y uso de estrategias de aprendizaje) de 138 alumnos de educación secundaria extremeños que aprenden inglés en dos contextos educativos distintos (AICLE vs Inglés como Lengua Extranjera [ILE]). Dos instrumentos —el *Vocabulary Levels Tests* (Schmitt, Schmitt & Clapham, 2001) y el *Productive Vocabulary Levels Tests* (Laufer & Nation, 1999)— se utilizaron para medir el conocimiento léxico receptivo y productivo de las 2.000 palabras más frecuentes y de los términos académicos; mientras que para analizar el uso de estrategias de aprendizaje se desarrolló un cuestionario basado en la taxonomía de Schmitt (1997).

Los resultados de este estudio indican una clara diferencia entre alumnos AICLE e ILE, tanto en el nivel de conocimiento léxico como en el uso de estrategias. Los alumnos AICLE muestran mejores resultados en los cuestionarios de conocimiento pasivo y productivo de las bandas léxicas analizadas. En lo referente a la selección de estrategias de aprendizaje, ambos grupos difieren en su uso de ciertas estrategias, con los alumnos AICLE haciendo un mayor uso de estrategias de análisis léxico.

Estos resultados pueden ser relevantes dado el déficit de investigación en lo referente (1) al impacto de AICLE en el procesamiento del léxico en la lengua extranjera y (2) al análisis de la relación de AICLE con otros factores como la cantidad de exposición a la lengua extranjera. La confirmación de las diferencias entre ambos grupos, no solo en cuanto al vocabulario general sino también al vocabulario académico, y una primera aproximación de cómo el contexto de enseñanza afecta al procesamiento de léxico en la L2, junto con la consideración de la posible influencia de la exposición a la lengua extranjera en los resultados, puede ayudar a esclarecer alguno de los aspectos más conflictivos sobre la efectividad del enfoque AICLE y su impacto en el aprendizaje de lenguas extranjeras.

TABLE OF CONTENTS

PART TWO

THE STUDY

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AoE — Amount of Exposure

AWL — Academic Word List

BICS — Basic Interpersonal Communicative Skills

BNC — British National Corpus

CALP — Cognitive Academic Language Proficiency

CEFRL — Common European Framework of References for Languages

CLIL — Content and Language Integrated Learning

CSE —Compulsory Secondary Education

DOE — *Diario Oficial de Extremadura*

EEC — European Economic Community

EU — European Union

ELP — European Language Portfolio

EFL — English as a Foreign Language

EMI — English as a Medium of Instruction

EST — English for Science and Technology

EU — European Union

FFL — French as a Foreign Language

FDR — False Discovery Rate

FL — Foreign Language

HOTS — Higher-Order Communication Skills

ICTs — Information and Communication Technologies

ID — Individual Factor

L1 — First language

L2 — Second Language

L3 — Third Language

LFP — Lexical Frequency Profile

LOMCE — *Ley Orgánica para la Mejora de la Calidad Educativa*

LOGSE — *Ley Orgánica de Ordenación General del Sistema Educativo*

LOTS — Lower-Order Communication Skills

MCT — Multiple-Choice Test

MEC — Spanish Ministry of Education and Culture

PFL — Portuguese as a Foreign Language

PVLT — Productive Vocabulary Levels Test

SES — Socio Economic Status

SILL — Strategy Inventory for Language Learning

SLA — Second Language Acquisition

SRCVLS — Self-Regulating Capacity in Vocabulary Learning Scale

TPR — Total Physical Response

UN — United Nations

USA — United States of America

UWL — University Word List

VLS — Vocabulary Learning Strategy

VLSQ — Vocabulary Learning Strategies Questionnaire

VLT — Vocabulary Levels Test

ZPD — Zone of Proximal Development

CHAPTER ONE:

INTRODUCTION

In the last 40 years, vocabulary has reached an unforeseen position within the field of Second Language Acquisition (SLA; Boers & Lindstromberg, 2008; Jiménez-Catalán & Terrazas Gallego, 2005; Meara, 1980; Milton, 2009; Milton & Fitzpatrick, 2014; Nation, 2001; Schmitt, 2010). For several decades, vocabulary learning played a secondary role in the most popular approaches to teach foreign languages. It was not until the 1980s that, when the Communicative Language Teaching became the dominant approach in foreign language instruction, vocabulary acquisition became the focus of L2 teaching and learning. From that moment onwards, vocabulary studies area gradually began to gain momentum, showing the importance and the key role that vocabulary plays in foreign language learning (Boers & Lindstromberg, 2008).

In this context, studies on vocabulary learning strategies emerged as a response to the need to understand how L2 learners come to master and process vocabulary. Language learning strategies had already been explored for nearly twenty years from a psycholinguistic perspective (Bialystok, 1978; O'Malley & Chamot, 1990; Oxford, 1990; Rubin, 1975) when they started to be applied to lexical development. This field of study emerged due to the existing concern about how learners' individual characteristics may be affecting the language learning and aimed to identify the actions or behaviours that language learners adopted when learning a new language. The findings in the field were soon applied in a number of taxonomies gathering a varied number of actions learners used when learning an L2 (Stöffer, 1995; Gu & Johnson, 1996; Schmitt, 1997; Nation, 2001; Intaraprasert, 2004; Tseng, Dörnyei & Schmitt, 2006), but the practical view of the construct led to a lower theoretical development of the concept than expected. As a result of this situation, although language, and, consequently,

vocabulary learning strategies have been examined and identified for various decades, there is still much to do in this area of research. For instance, despite the abundant body of literature on language learning strategies, few studies have focussed on secondary-school learners, because it is commonly believed that this kind of learners may have difficulties when reflecting about their own learning process. Moreover, and focussing on vocabulary learning strategies, despite their relevance, research on identifying the effects of using particular learning strategies on lexical development is nearly negligible.

In this PhD dissertation, these two elements —lexical knowledge and vocabulary learning strategies— are conceived as parts of a larger construct: lexical competence. For decades, lexical competence has been understood as a synonym of vocabulary knowledge. However, the term 'competence' entails more than just knowledge, as it implies putting into practice this knowledge. It is in line with this idea that the present study was conceptualised. The main objective of this PhD dissertation is to analyse secondary-school learners' lexical competence development. Specifically, it consists in an analysis of the impact of two elements —language teaching context (CLIL vs mainstream EFL approaches) and instructed amount of exposure (IAoE) to the FL— on two components of lexical competence: vocabulary knowledge and learners' self-regulation capacity for learning vocabulary, in a group made up of 138 students in Extremadura (Spain). Due to the complexity of these concepts and, in order to better conceptualise and explain the research problem and the relationship between the different variables, I will make use of Leow's theoretical framework for L2 learning internal processing (2015). As will be further explained in Chapter 2, this author identifies two external products (*input* and *output*) and five stages that make up the L2 learning internal process. Three of these stages are processes, known as *input processing, intake processing*, and *knowledge processing,* which, in turn, generate two products: *intake* and *L2 knowledge*. The different products and stages are connected as shown in Figure 1.1:
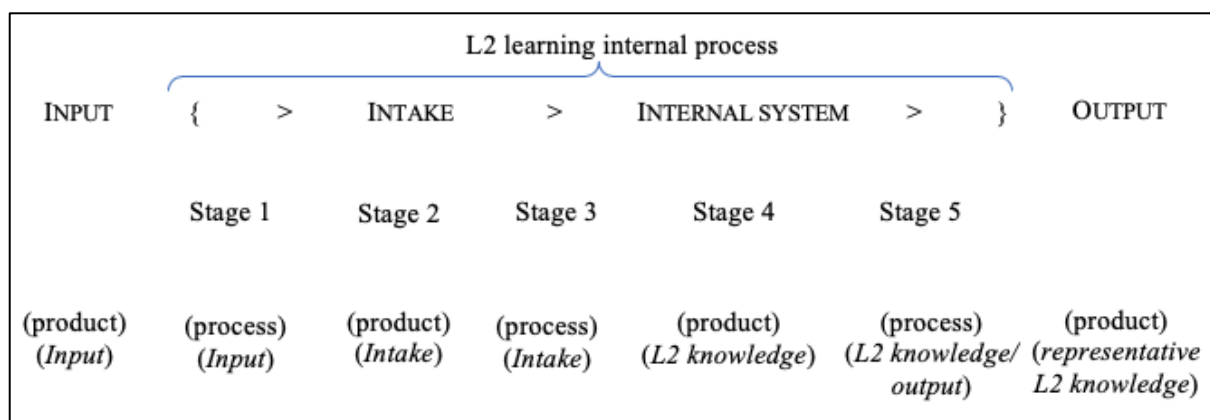
Figure 1.1. Leow's framework for L2 processing.

Vocabulary learning is a central component of L2 acquisition, and, consequently, follows the same scheme as that of general L2 learning (see section 3 of Chapter Two for a detailed explanation of the process). Therefore, L2 lexical development is a long process that entails a large number of internal

sub-processes. In this doctoral dissertation, rather than exploring all these different vocabulary learning sub-processes, the focus has been placed on specific aspects —language teaching approach, IAoE, intake processing and learning output— and their relationships, which have been rarely explored in the field of CLIL. Specifically, it aims to examine whether variations in *input* may affect a specific subdomain of the internal processing (*stage 3: intake processing*) and *output* measured as vocabulary knowledge. The following figure outlines the main variables explored in this piece of research and conceptualises them making use of Leow's framework. The reasons for selecting these elements of the vocabulary processing are explained below.
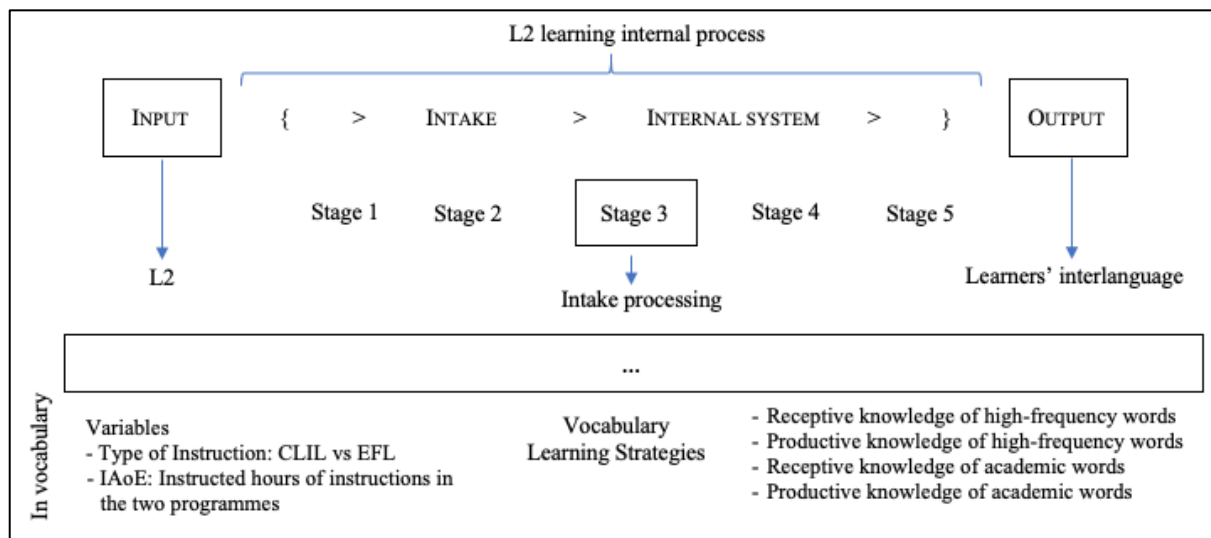
Figure 1.2. Outline of the main variables explored following Leow's framework.

Regarding the first element, i.e., *input*, traditionally, this variable and its impact on L2 development has been related to differences in two main aspects: *quality* and *quantity* of input. However, in a recent paper, Rothman and Guijarro-Fuentes (2010) point to a third factor that affects the impact of the input variable: differences in setting and instructions. These authors carry out a literature review in which input differences are related to *setting of acquisition* —naturalistic or formal. This PhD dissertation aims to go a step further and attempts to prove that not only the setting, but also, the *language teaching approach* (CLIL vs EFL) followed may affect L2 processing and the output of this learning process.

In the last decades, there is a new teaching approach being implemented in Europe —known as Content and Language Integrated Learning (CLIL)— which integrates content and linguistic aims in the development of content subjects. In the case of Spain, CLIL has become, after some years of pilot implementation, an intrinsic element of the Spanish educational system. This approach entails the use of a foreign language as a vehicle of communication in content subjects and since its implementation, many voices have been raised for both, defending its benefits but also for highlighting their main backwards and risks. On the one hand, its main detractors argue that (1) CLIL threatens egalitarianism at school (Bruton, 2011a, 2011b, 2013), (2) that its supposed benefits are not the result of the methodologies employed, but of a greater exposure to the foreign language (FL) and (3) that this

approach is just a weaker copy of other content-based approaches, such as immersion or English Medium Instruction programmes. On the other hand, the main advocates of this approach argue that this new way of addressing the FL, by means of considering it not an objective itself, but a means of communication and learning, presents clear language learning benefits. As a result, there has been a wide body of research exploring the benefits of CLIL in very different areas such as motivation, syntax or vocabulary. Focussing on vocabulary, the vast majority of research studies have analysed differences regarding vocabulary size between CLIL and EFL learners (Agustín-Llach, 2012; Arribas, 2016; Canga Alonso, 2015a, 2015b; Canga Alonso & Arribas, 2014; Fernández Fontecha, 2014), suggesting as the main linguistic benefits in comparison to previous approaches variations in the quality and quantity of input to which learners are exposed and in the way CLIL learners are exposed to this input.

Concerning the variations in the input to which CLIL learners are exposed, two main changes are found. First, CLIL leads to a more varied and rich input, as learners are asked to learn content subjects in a foreign language. Second, CLIL programmes involve an increase in the number of hours of L2 input learners are exposed to, as learners attend FL lessons and, in addition, content subjects taught in this FL. In fact, as mentioned above, some researchers (Bruton, 2011, 2013, Paran, 2013) argue that linguistic benefits attributed to CLIL (see section 4.3.1 of Chapter Two for an overview) are not related to the use of different pedagogical techniques but just to an increase in the exposure to the L2. However, CLIL entails more than just a mere variation of the amount of input, as, for instance, it also involves a change in the kind of input to which learners are exposed as well as the incorporation of a large body of methodologies for the teaching of both, content and language. This leads us to the second idea regarding input to be explored in this doctoral dissertation: is an *increase of IAoE* to the FL the only reason to explain the changes in CLIL learners? To explore this aspect, participants are going to be clustered according to their exposure to the FL and their performance is going to be compared. If, despite the differences in the IAoE, CLIL groups present equivalent results, this may indicate that CLIL results are due to the effect of other elements rather than the IAoE.

As for the latter aspect, i.e., the way input is presented to learners, CLIL is believed to foster incidental L2 learning: given that the L2 is used to work on content, language learning takes place while learners are focussed on fulfilling content tasks (Dallinger et al., 2016; Merikivi & Pietilä, 2014; Surmont, et al., 2016; Vallbona González, 2014). However, this does not mean that explicit language explanations cannot be introduced in the CLIL classroom. Indeed, there is some evidence that the inclusion of these explanations in CLIL benefits L2 development, particularly, vocabulary learning (Scott & Beadle, 2014). In this respect, the CLIL context seems optimal for vocabulary acquisition, as it provides both the explicit and incidental learning opportunities required to enhance lexical knowledge (see Chapter Two for an overview). In this sense, Merikivi and Pietilä (2014, p. 31) states that "CLIL may in fact be ideal, as it combines explicit and implicit learning conditions". It seems, thus, that the combination of the more naturalistic and meaningful environment that CLIL promotes and the explicit

language explanations that CLIL learners find in both CLIL and EFL lessons may benefit L2 vocabulary development.

Taking into consideration these ideas, this study attempts to explore whether variations in input related variables (language learning approach and IAoE) are reflected, and, if so, in what ways, (1) in *L2 intake processing* and (2) in the *learning output*, i.e., in this case, lexical knowledge. However, given the many sub-dimensions the vocabulary knowledge construct has, there is a need to focus on specific aspects within it. There is some agreement that the different sub-dimensions of word-knowledge correlate. Therefore, the exploration of any of them, for instance, written form recognition or word associations, may yield representative results for lexical competence in general. To select the elements to be studied, Nation's taxonomy of word knowledge was used (see p. 46). From the three categories and the nine sub-elements of study, this piece of research focusses on *form*, and more concretely on *written form recognition* and *production*, as they seem to be one of the most widely explored aspects in young teenagers. However, in contrast to previous studies in which the analysis of L2 vocabulary knowledge was only restricted to general vocabulary, this study offers a broader vision of vocabulary knowledge, as it analyses general, but also academic vocabulary. This inclusion is justified by the conceptualization of language in CLIL. Generally, three kinds of languages seem to co-exist in the CLIL classroom —*language of learning, language for learning,* and *language through learning*— and each of them present their own particularities (see p. 91 for an overview or Coyle, 2007 for further information) and are applied for different purposes. In practice, CLIL learners are required, for instance, to discuss, analyse, synthesise or apply ideas using an adequate L2 language and tone. These language demands are a complete novelty in comparison to the traditional EFL classroom and demand specific language structures: the *academic English*, which, in the CLIL context, is closely reltated to *the language for learning*. As this kind of language is specifically worked in CLIL, then, CLIL learners are expected to have a better academic vocabulary knowledge than more traditional EFL learners.

The implementation of CLIL not only seems to modify L2 input, but also seems to affect the way language and content are learnt, i.e., *intake processing* according to Leow's framework. CLIL places new language and content demands on learners which result in cognitive demands. If CLIL learners are to fully achieve these aims, these learners need to develop and put into practice new cognitive processes. This brings us to modifications in the learning process itself. However, despite its importance for the correct language and content development in CLIL, research on the cognitive effects has been a neglected area, as most research has been devoted to identifying linguistic and content implications. This situation is similar in EFL processing research, where studies on language processing are limited to specific dimensions of the construct. Traditionally in SLA, among the different sub-processes encompassed in the learning process, a great deal of attention has been paid to the understanding of the *input processing* in order to identify the attentional processes that make learners focus on some aspects rather than on others (Schmidt, 1995; Godfroid, Boers & Housen, 2013). This has led to a lack of

development in other areas such as *intake* and *knowledge processing*. This doctoral dissertation aims to partly fill this gap by exploring how variations in the input, namely, *differences in the amount of exposure* and *differences in the teaching approaches* used to introduce this input in the classroom resonate in the *intake processing*.

Therefore, this dissertation puts together lexical knowledge and vocabulary learning strategies under the umbrella of Lexical Competence. This is, *per se*, a notable difference with previous research into the implications of CLIL in lexical development. However, it also includes a new aspect in comparison to previous research: implications are explored cross-sectionally and longitudinally. Unlike previous research, this dissertation includes a longitudinal analysis of participants' receptive vocabulary growth and use of vocabulary learning strategies that may help to better define the effects of CLIL on vocabulary learning. To date, most research on CLIL and lexical development has been carried out cross-sectionally (see, for example, Agustín-Llach, 2012; Arribas, 2016; Canga Alonso, 2015b, Castellano-Risco, 2018, Merikivi & Pietilä, 2014), with little research exploring how the differences, or the lack of them, evolve over time. The inclusion of a longitudinal analysis follows current trends in the analysis of CLIL effects (Perez Cañado, 2018, Sylvén, 2019), as the use of this kind of studies is thought to help to have a better control on the variables that may bias the results, such as the IAoE or maturational level.

This dissertation is organised as follows. It is structured around two main parts: Part One comprises the literature review chapters and Part Two deals with the methodological aspects of the study, the results and their discussion, concluding with the drawing up of the conclusions. Table 1.1. describes the contents of each chapter.

Table 1.1

PhD dissertation structure

Part	Chapter	Description
Part 1: Literature Review	Chapter Two	Vocabulary Learning in Second Language Acquisition
	Chapter Three	The Vocabulary Learning Strategies Construct
Part 2: Experimental Study	Chapter Four	Methodology
	Chapter Five	Results
	Chapter Six	Discussion
	Chapter Seven	Conclusions

Part One deals with the theoretical background for this study and comprises two chapters. Chapter Two is devoted to vocabulary in SLA and one of its main exponents, 'lexical competence'. It provides an overview of the lexical competence and lexical knowledge constructs, mainly focussed on four main points of interest: the definition of lexical competence and how it is processed in the mind; how to measure lexical development, and, finally, the issue of lexical knowledge development in school-age learners. Chapter Three addresses the concept of *vocabulary learning strategies*. To elaborate on this issue, this chapter presents the most relevant literature in the field in order to contextualise what learning strategies are and to identify the main factors that influence the selection of strategies. The chapter concludes by exploring how different researchers have tackled the issue of how second language learners acquire vocabulary through vocabulary learning strategies.

As for the second part of this dissertation, it is made up of four chapters. Chapter Four deals with the description of the methodology, in which the research questions are posed, and the context of the study, participants, instruments and data treatment are detailed. Then, in Chapter Five, the results are presented. It follows a general-to-specific approach and the same order of the research questions posed, so it starts with a description of the overall results, moves on then to the differences found between CLIL and mainstream EFL learners in lexical knowledge first, and then in the use of vocabulary learning strategies, and ends up with the differences found within the CLIL subgroups in the same two variables. Then, in Chapter Six, these results are taken up and discussed in relation to the research questions posed in Chapter Four. Finally, Chapter Seven presents the conclusions of the study, together with the analysis of the limitations and the potential implications of this study.

PART ONE

LITERATURE REVIEW

CHAPTER TWO:
LEXICAL COMPETENCE IN SECOND LANGUAGE ACQUISITION

This Chapter deals with the notion of *lexical competence*, as in the last thirty years, this construct has become very relevant in SLA and it is the basic concept of this PhD dissertation. Therefore, I will start by attempting to provide a clear definition of *lexical competence* and exploring the most influential proposals on this topic. Then, I will move on to another issue relevant to the interests of this piece of research, the understanding of how lexical items are processed in L2 learners' minds. For this purpose, I will present some of the most influential models of L2 processing to date. After that, I will focus on vocabulary knowledge as one of the main components of *lexical competence* and will mainly concentrate on two aspects of relevance to the present work: how to measure vocabulary knowledge and how L2 vocabulary is developed in L2 classrooms. I will conclude with an outline of the most important contributions of this Chapter to this PhD dissertation.

1. The Lexical Competence Construct

In the last thirty years, lexical competence has become a central notion in SLA, which makes it necessary to provide a clear definition of the construct. To do so, a revision of some of the most influential papers on this issue will help us in the task. A discussion of what the concept of *lexical competence* actually entails will complete the picture.

Lexical competence started to be used as a construct by vocabulary studies and is described as the ability to recognise and use the words of a specific language in a native-like way (López-Mezquita, 2005). Such is the recognition the term has gained, that it is even included in one of the most relevant

documents in the language policies of the EU: the Common European Framework of Reference for Languages (CEFRL; 2001). Published in 2001 by the European Council, this important document was created with the aim of establishing a common base for national educational authorities to set language educational objectives. It establishes six levels and, for each level, an exhaustive description of the competences is provided. In order to achieve a full development of communicative competence, the CEFRL distinguishes between two types of competences: (1) general competences, which are human knowledge or skills that may contribute to the language development and (2) communicative language competences, which refer to specific linguistic knowledge. It is within the communicative language competences that explicit reference to lexical competence is made (see Table 2.1), situating it at the same level as other competences that had traditionally been prioritised, such as the grammatical or phonological competences.

Table 2.1

Classification of competences the CEFRL.

Competences	Sub-competences	Specific competences
General competences	Declarative knowledge	
	Skills and know-how	
	Existential competence	
	Ability to learn	
Communicative language competences	Linguistic competences	**Lexical competence**
		Grammatical competence
		Semantic competence
		Phonological competence
		Orthographic competence
		Orthoepic competence
	Sociolinguistic competences	Linguistic makers of social relations
		Politeness conventions
		Expressions of folk wisdom
		Register differences
		Dialect and accent
	Pragmatic competences	Discourse competence
		Functional competence

Source: Common European Framework of References for Languages (2001, pp. 101-130). (Emphasis mine).

Lexical competence is defined in the CERFL as the "knowledge of, and ability to use, the vocabulary of a language, consists of lexical elements and grammatical elements" (p. 109). The importance given to it is demonstrated by the fact that, for each level, the CEFRL specifies the richness of vocabulary needed, the topics the language learner should manage and the registers in which the

learner should be able to use this lexis. Thus, the definition given in the CEFRL makes reference only to the knowledge of vocabulary items, leaving psycholinguistic considerations aside. Moreover, the definition provided in the CEFRL is rather descriptive and it does not provide further explanation of its origins or its theoretical roots, therefore, in the following section, I will focus on the construct from which it has emerged and within which it is integrated: Communicative Competence.

1.1. Tracing the origins of Lexical Competence: Communicative Competence

The notion of *lexical competence* has often been encapsulated within the *communicative competence* construct (López-Mezquita, 2005), which is thought to have its origins in Chomsky's differentiation between *linguistic competence* and *linguistic performance* (1965). The American linguist saw language as a structure and linguistic competence as the knowledge of such structure. Performance, on the other hand, was the actual production of linguistic utterances of this knowledge. Chomsky's notion of *linguistic competence* brought about a growing discussion on its actual application to the FL classroom. One of the most fruitful contributions was the one made by Hymes (1972), who argued that learners' linguistic competence could not only be related to language knowledge, as this view ignored the appropriateness of language use which is dependent on the communicative situation. Thus, he replaced the Chomskyan notion of *linguistic competence* with his own concept of *communicative competence*, which comprised both *linguistic* and *sociolinguistic competences*. *Linguistic competence* encompassed, in his view, language knowledge, whereas *sociolinguistic competence* had to do with the knowledge of the appropriateness of the language in context.

Given the rising interest in the notion of communicative competence and the various discussions this issue generated, Canale and Swain (1980) published a seminal article that shed some light on this issue. They presented a thorough, systematic revision of the most influential communicative competence proposals up to that date finding that most of them missed an element that, for them, was essential: the strategies used to handle breakdowns in communication. They presented their own model in which they added a third element to the linguistic and sociolinguistic competences that Hymes (1972) had already proposed: *strategic competence*. Moreover, they provided a definition for each of the elements. Thus, in their view, *communicative competence* consisted of three elements: grammatical competence —which involved the "knowledge of lexical items and of rules of morphology, syntax, sentence-grammar semantics, and phonology" (p. 29)—, *sociolinguistic competence* —that included two sets of rules: sociocultural rules and rules of discourse— and *strategic competence* —that referred to linguistic strategies to be used to compensate for breakdowns in communication.

The importance of this proposal for us lies in that it is one of the first ones including *lexical knowledge* in some way. It is noteworthy to mention that, since that moment, the notion of *lexical knowledge* became a constant in the subsequent communicative competence models. This is not the only Canale and Swains' contribution, as they dealt with several key aspects, such as the terminology.

In this respect, for example, they emphasised the importance of selecting the most appropriate terms, since, in their view, there were different connotations in using the terms linguistic competence or grammatical competence to refer to language knowledge.

Bachman and Palmer's model of Language Ability (1996) is the next milestone in the development of the concept of *communicative competence*. These authors proposed "a model for describing the characteristic of the language users, or potential test takers" (1996, p. 61), placing the emphasis on learners as an active part of the learning process. For them, language was no longer viewed as an abstract structure, but as a tool L2 learners needed to know to communicate in a specific context. For this reason, they identified two main components of language ability: (1) language knowledge, or, in other words, language proficiency, and (2) strategic competence, related to the capacity to put into practice their linguistic knowledge. In their view, language knowledge was "a domain of information in memory that is available for use by the metacognitive strategies in creating and interpreting discourse in language use" (p. 67) and was divided into two main groups: organizational knowledge, that, in their own words, "[was] involved [in] controlling the formal structure of language for producing or comprehending grammatically acceptable utterances or sentences, and for organizing these to form texts, both oral and written" (pp. 67-68), and pragmatic knowledge, that enabled speakers to create discourse. At the same time, within each of the groups, two sub-levels were established: in the case of organizational knowledge, it comprised both grammatical and textual knowledge, whereas pragmatic knowledge was made up of functional and sociolinguistic knowledge. Grammatical knowledge was related to the production or comprehension of accurate sentences or utterances, and included the knowledge of vocabulary, syntax, phonology and graphology. Textual knowledge dealt with the comprehension or production of texts, that were made up of sentences or utterances, and consisted of two kinds of knowledge: knowledge of rhetorical or conversational organization and knowledge of cohesion. As for functional knowledge, it allowed speakers to interpret the relationship between the productions of one user and his or her intention. It included the knowledge of (1) intentional functions, (2) heuristic functions, (3) imaginative functions, and (4) manipulative functions. Finally, sociolinguistic knowledge was related to the speaker's capacity to relate the appropriateness of the language utterance with the possible different settings. Figure 2.1 below shows a schematic representation of Bachman and Palmer's *language ability* concept (1996).
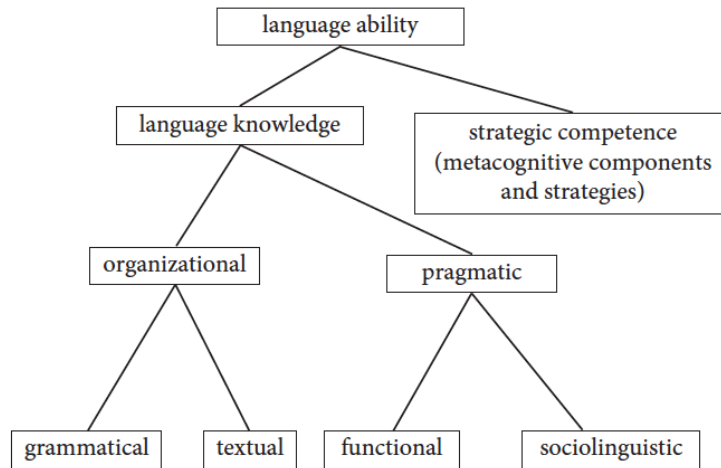
Figure 2.1. Schematic representation of language ability by Bachman & Palmer (1996). Source: Hulstijn, 2015, p. 41.

Bachman and Palmer's model is one of the best-known examples of the different proposals of communicative competence. However, a year before the publication of Bachman and Palmer's book, Celce-Murcia, Dörnyei and Thurrell (1995) had already made their own proposal of Communicative Competence. In Celce-Murcia et al.'s view, communicative competence consisted of five sub-competences and the relationship among them: discourse competence, socio-cultural competence, linguistic competence, actional competence and strategic competence. The most salient feature of their proposal is the consideration that the different competences were interrelated. As can be seen from Figure 2.2, they represented communicative competence as "a pyramid enclosing a circle and surrounded by another circle" (p. 10). In the inner circle, they placed discourse competence, i.e., the capacity concerning the "selection, sequencing and arrangement of words, structures, sentences and utterances to achieve a unified spoken or written text" (p. 13). Each of the points of the triangle represented the linguistic, *actional* and *sociocultural* competences. Finally, the circle surrounding the pyramid was related to strategic competence, which was regarded as a series of skills that could help a speaker to negotiate messages, solve linguistic problems or compensate deficiencies in the other competences.
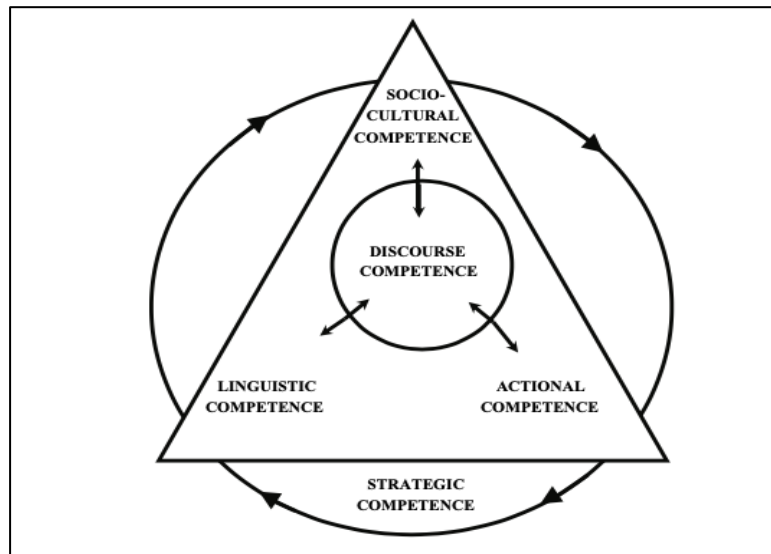
Figure 2.2. Schematic representation of communicative competence by Celce-Murcia et al. (1995).
    Source: Celce-Murcia (2008, p. 44).

Apart from the specification of the interrelationship between the different elements, other changes were proposed to both Canale and Swain's (1980) and Bachman and Palmer's (1996) models. Firstly, Celce-Murcia et al. (1995) added a fifth competence, *actional competence*, that was defined as the capacity or ability to convey and understand "communicative intent, i.e., matching actional intent with linguistic form based on the knowledge of an inventory of verbal schemata that carry illocutionary force (speech acts and speech act sets)" (Celce-Murcia et al, 1995, p. 18). Moreover, they made two terminological changes to Canale and Swain's (1980) proposal: they decided to use the term *linguistic competence* rather than *grammatical competence* and (2) they decided to use the term *sociocultural competence* rather than *sociolinguistic competence*.

Canale and Swain's (1980), Bachman and Palmer's (1996) and Celce-Murcia et al.'s (1995) models of communicative competence placed the lexical element within language knowledge without providing specific orientations about what lexical knowledge really implied. A substantial change in this respect came with Celce-Murcia (2008), in which the author carried out a revision of her previous model (see Figure 2.3). In the case of *linguistic competence*, Celce-Murcia refined and defined its sub-dimensions, and, as a result, for instance, the knowledge of lexical items was defined as "knowledge of both content words (nouns, verbs, adjectives) and [f]unction words (pronouns, determiners, prepositions, verbal auxiliaries, etc.)" (Celce Murcia 2008, p. 47). Similarly, the rest of the competences were also portrayed more and more concretely. For instance, in the characterization of strategic competence, she differentiated between communicative strategies, used to support communication and overcome communicative problems, and learning strategies, used to develop learners' linguistic knowledge. These latter strategies are the ones explored in this thesis dissertation.
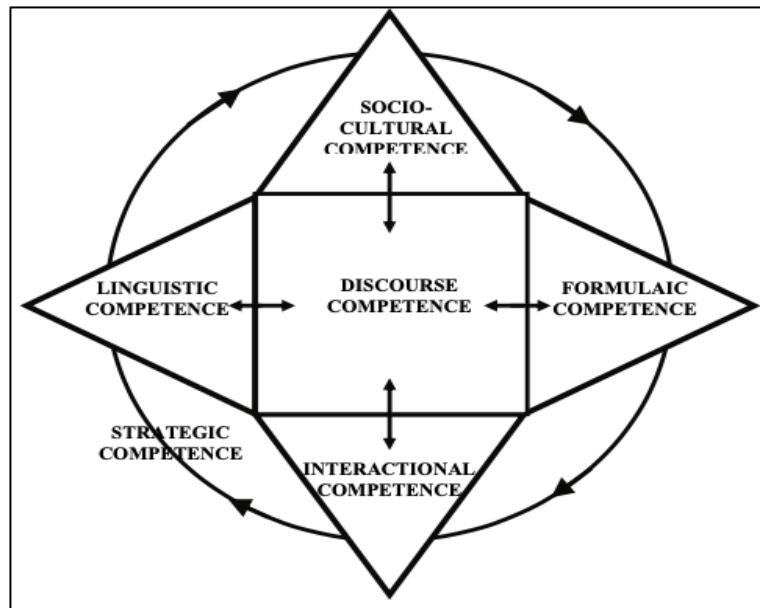
Figure 2.3. Representation of communicative competence by Celce-Murcia (2008). Source: Celce-
Murcia (2008, p. 44).

All in all, as can be seen in this review of different models of *communicative competence*, the notion of *lexical knowledge* has progressively been incorporated within the construct *communicative competence*. It has been included in different dimensions within the linguistic competence and re-defined and re-interpreted several times. This reference to lexis has not been homogeneous, but, depending on the model, it has been denoted in one way or another. For example, Celce-Murcia et al. (1995) talk about 'knowledge of lexical resources', while Celce-Murcia (2008) or Canale & Swain (1980) make use of 'lexical knowledge'. In short, the knowledge of vocabulary has been denoted in different ways, such as vocabulary knowledge, knowledge of lexical items or knowledge of lexical resources. Figure 2.4 offers a summary of the models presented in this section and their consideration of vocabulary knowledge.
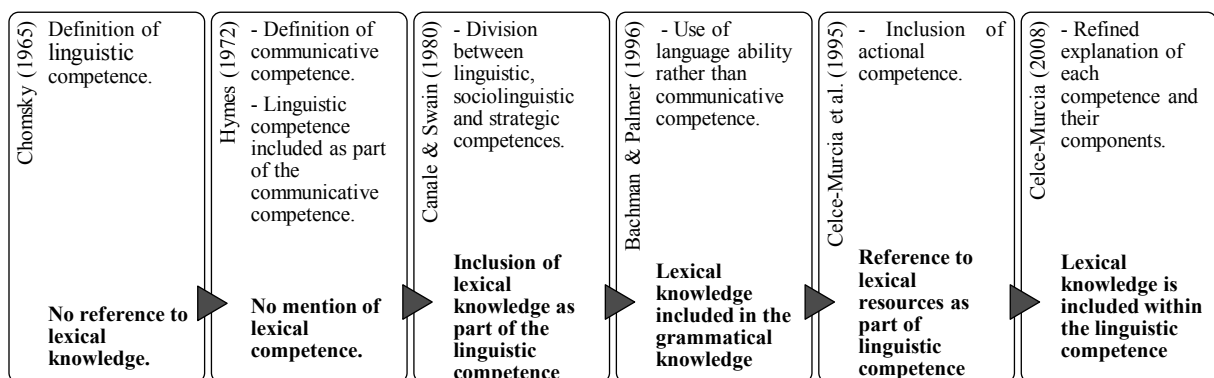
Figure 2.4. A summary of the consideration of lexical knowledge in the main communicative competence models.

However, in none of the models presented above, has vocabulary knowledge been labelled as 'lexical competence'. Therefore, despite the fact that 'lexical competence' appears in official documents, such as the CEFRL, and that, in practice, it is considered part of the Communicative Competence in

vocabulary research (López-Mezquita, 2005), it has not been used in any of the most influential models of Communicative Competence. With this in mind, I would say that, at first sight, it seems that both lexical competence and lexical knowledge refer to the same idea, the knowledge of vocabulary items.

1.2. Approaches to the description of Lexical Competence

The dawn of the interest in lexical knowledge can be situated at the beginning of the 20th century, when some scholars attempted to identify the different elements that make up lexical knowledge (Palmer, 1921). However, the area was soon abandoned, and it was not until the 1970s that it flourished, grounded in the idea that lexical development was central to L2 learning (Meara, 1980; Nation, 1974, 1975; Richards, 1976). Thus, these studies precede the notion of Communicative Competence and have somehow developed alongside it. For decades, the terms *vocabulary* and *word knowledge* have appeared time and again in SLA research to make reference to the lexical development of L2 learners. However, from the 1990s onwards, there was a reconceptualization of the term and the concept *lexical competence* started to replace that of *vocabulary knowledge*. Due to the multifaceted nature that characterises the lexical competence construct, lexical studies have explored vocabulary knowledge or lexical competence from different perspectives and approaches that vary to a large extent from one to another. In an attempt to shed some light on the definition of the construct, Jiménez Catalán (2002) examined the different definitions, descriptions and dimensions of lexical competence that appeared in twenty-two papers published in the most recognised journals and she concluded that, depending on the terms used and where the emphasis was placed, two main approaches to lexical competence could be drawn: linguistic and psycholinguistic approaches. The former conceived language and vocabulary as a compendium of lexical items, whereas the latter saw language as a system, and the description of lexical competence varied considering the language preconceptions. The same distinction has been made recently by Leńko-Szymańska (2020) who distinguished to two main approaches —word-centred and lexicon-centred— in the description of lexical competence. This distinction is going to be used below to provide a definition of what lexical competence is.

1.2.1. Linguistic approach

The research following this perspective focusses on the semantic and grammatical analysis of the lexical elements (Jiménez Catalán, 2002). It proposes thus an analysis of the vocabulary construct at a microscopic level, in which *lexical competence* is often equated to *word knowledge* or even to *vocabulary knowledge*. Studies that follow this approach start from a general term —usually, *vocabulary*— and move on to more particular and concrete elements that identify the different components making up *word knowledge*. Following the same sequence, and. as shown in Figure 2.5, the definition of *vocabulary* will be taken as a starting point, and from it, I will focus on the definition of word and, finally, on the identification of the dimensions needed to know a word.

```
┌─────────────────────────────────┐
│          Vocabulary             │──┐
└─────────────────────────────────┘  │
     ┌──────────────────────────────┐│
     │            word              │◄┘──┐
     └──────────────────────────────┘    │
          ┌──────────────────────────┐    │
          │           word           │◄───┘
          │        knowledge         │
          └──────────────────────────┘
```
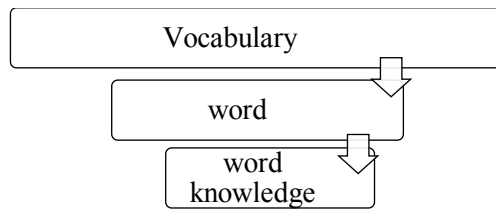
Figure 2.5. Sequence of analysis in a linguistic approach.

Vocabulary is often defined as "a set of lexemes which includes single words, compound words and idioms" (Richards, Platt & Plat, 1992, p. 400), i.e., in layman's terms, a group of words that make up a language. This definition does not seem to entail many difficulties, with the exception of understanding technical words such as 'idioms'. Nevertheless, it is in one of the terms used in the definition —*word*— where the real problem lies. The understanding of what a *word* denotes seems to be more difficult than it may appear at first sight: the meaning of the term *word* cannot be established in isolation, but only in relation to other concepts, and this hampers the development of the construct (Trask, 1995; Richards & Schmitd, 2010). Due to this difficulty, there are as many definitions of the term as approaches to the subject. As can be seen in table 2.2 below, the definitions given to the term 'word' vary depending on the perspective adopted.

Table 2.2

Definitions of the term word

Perspective	Authors	Definitions
Orthographic	Bloomfield (1926, p. 156)	"a minimum free form is a word"
		"thus a form which may be uttered alone (with meaning) but cannot be analysed into parts that may (all of them) be uttered alone (with meaning)"
	Carter (1998, p. 20)	"[…] any sequence of letters (and a limited number of other characteristics such as hyphen and apostrophe) bounded on either side by a space or a punctuation mark"
	Richards, Platt & Platt (1992)	"the smallest of the linguistic units which can occur on its own in speech or writing"
	Scheeler & Markley (2000, p. 2)	"a unit formed by sounds or letters that have a meaning".
Semantic	Takac, 2008	The smallest meaningful unit of a language

Source: Own elaboration.

From an orthographic point of view, the term *word* is related to a sequence of letters or sounds, that is usually referred to as 'unit'. It could be argued that these definitions are somewhat partial. For instance, in the case of Carter (1998, p. 20), he only refers to the written language, dismissing the

importance of oral discourse. In other examples, there is no reference to a key element in the definition of *word*: meaning (Takac, 2008, p. 5). That is the case of the definitions provided by Richards et al. (1992) or Carter (1998, p. 20). However, the most common problem is the vague specification of 'unit', a term found regularly when comparing definitions. What are these authors referring to when talking about 'linguistic unit'? Is, for example, a phrasal verb, a word? Or are they considered 'multi-words'?

The same difficulty is found when exploring the definitions of *word* provided from a semantic point of view. In this view, a *word* can be the smallest meaningful unit of language (Takac, 2008). Main arguments against this definition refer, once again, to the complicated nature of the task of analysing vocabulary items such as compound terms, phrasal verbs and other compound forms, because, according to this definition, they should be counted as a single word, and this view is in conflict with the idea of *unit*.

It is thus clear from these definitions that a word seems to be an indivisible lexical unit that denotes its own meaning. Nevertheless, there are still some questions that remain open: can inflected or derived forms from other words be considered new word forms? What happens with compound terms or phrasal verbs? Are they one term, or, on the contrary, do they represent different word forms? To date, there is no consensus on these questions and depending on their purposes, researchers take their own decisions about what they consider as a word. Therefore, when exploring vocabulary knowledge, it is important to clarify the unit used as a reference.

Closely related to this complexity of identifying what a word is, is the difficulty in specifying what *knowing a word* involves, given the many components of the concept. Despite this difficulty, as Sanjuan Álvarez (1991) notes, the identification of those aspects involved in *vocabulary knowledge* is central to teaching practice since it may help discriminate what to present to learners. In line with this idea, a large amount of research has been carried out to identify those dimensions involved in word-knowledge. In general, there is some agreement that knowing a word entails more than just the mere identification of the written and oral form of the word; in order to use it efficiently, the learner has to be familiar with aspects such as the register in which it is used, the words with which it usually appears or whether it appears more frequently as a written or spoken form. It is in those other aspects in which it is complicated to reach an agreement.

The exploration of the components of word knowledge has its origins in the work of ancient Greek philosophers. Aristotle, in the fourth century BC, already approached this issue. In *De Interpretatione*, Aristotle explored the relationship between thought and words and distinguished four main reality components: real world things, impressions (the idea of those things), spoken signs and written signs. This interpretation of reality produced two main ideas that have been taken up in more recent times. First, Aristotle introduced the distinction between the form of a word and the meaning it represents. This idea would be taken up centuries later by Saussure (1916), when he referred to the concepts of

signifier —the form which the sign takes— and *signified* —the concept represented by the sign— to define what a word was. Secondly, Aristotle distinguished between the written and spoken form of a word. This division, spoken and written, is of prime importance as it is one of the first divisions of word acknowledged (Milton & Fitzpatrick, 2014).

Until the twentieth century, Aristotle's conceptualization was neither refuted nor questioned (Milton & Fitzpatrick, 2014). However, the development of modern linguistics as a new branch of research brought in fresh ideas. Palmer (1921) made the distinction between the ability to recognise words and being able to use them, introducing the receptive-productive dichotomy. But Palmer went beyond this division and proposed other qualities that were implied in the term *knowing a word*, such as the knowledge of frequent collocations or affixes, and, thus, taking a further step on this issue.

This first approach to a more complex model of *word knowledge* was soon explored by other linguists, who attempted to identify those qualities that might be included in the definition of *word knowledge*. For example, Cronbach (1942) distinguished between five main dimensions: generalization, application, breadth of meaning, precision of meaning and availability. *Generalization* referred to the definition of the word, and *application* to its appropriate usage. *Application* was closely related to *precision of meaning,* that was linked to register, i.e., the correct use of the word considering the context. As for *breadth of meaning*, it denoted the polysemous nature of words, as the same sign may refer to multiple meanings. Finally, *availability* referred to the capacity of using the word. Decades later, Richards (1976) made his own proposal based on the exploration of previous attempts to characterise *word knowledge* and set out seven competences necessary for mastering a word:

(1) Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words, we also know the sort of words most likely to be found associated with the word.

(2) Knowing a word implies knowing the limitations imposed on the use of the word according to variations of function and situation.

(3) Knowing a word means knowing the syntactic behaviour associated with that word.

(4) Knowing a word entails knowledge of the underlying form of word and the derivatives that can be made from it.

(5) Knowing a word entails knowledge of the network of associations between that word and the other words in language.

(6) Knowing a word means knowing the semantic value of the word.

(7) Knowing a word means knowing many of the different meanings associated with the word.

Richards, 1976, p. 83

Given its large level of specificity, Richard's proposal had the limitation that testing such a number of aspects for each word was extremely time-consuming. In one way or another, most of the subsequent proposals presented the same difficulty. In this sense, Laufer (1990), for instance, considered that *knowing a word* included learning the word form, word structure, syntactic behaviour, meaning and associative relations with other word. This view is consistent with Taylor (1990), who defined *word knowledge* as knowledge of frequency of occurrence, style, register, dialect, semantic style and syntactic collocations, morphology, semantics, polysemy and its translations. Similarly, Coady (1993) proposed that *knowing a word* concerned knowing its syntactic behaviour, derivations, network of associations, its semantic features or the register in which it could be found.

Finally, Nation (2013, p. 49) developed the most comprehensive approach up to that moment, by identifying three main categories of word knowledge, namely, *form, meaning* and *use*; each of which had their specific subcategories. In the case of the category *form*, he identified not only the written and spoken forms but also the word parts that make up the words. As for *meaning*, he distinguished also among three aspects: form and meaning, the concept the word is making reference to and finally the different associations that learners may establish when faced with the word. In relation to the dimension *use*, he identified the understanding of grammatical functions, the knowledge of collocations and the constraints on use as the main aspects. Moreover, after this first categorization, he distinguished two dimensions of vocabulary in each of the aspects already mentioned: the receptive and productive form. Finally, he provided assistance for the identification of those concrete dimensions with the inclusion of some questions. Table 2.3 presents Nation's classification of word knowledge.

Table 2.3

Word knowledge dimensions according to Nation, 2013.

Form	Spoken	Receptive	What does the word sound like?
		Productive	How is the word pronounced?
	Written	Receptive	What does the word look like?
		Productive	How is the word spelled?
	Word parts	Receptive	What parts can we recognize in this word?
		Productive	What word parts are needed to express meaning?
Meaning	Form and meaning	Receptive	What meaning does this form signal?
		Productive	What word form can be used to express this meaning?
	Concepts and references	Receptive	What is included in this concept?
		Productive	What items does the concept refer to?
	Associations	Receptive	What other words does this make us think of?
		Productive	What other words are possible to use instead of this one?
Use	Grammatical functions	Receptive	In what patterns does this word occur?
		Productive	In what patterns is this word required to use?
	Collocations	Receptive	What other words or types of words occur with this one?
		Productive	What words or types of words must we use with this one?
	Constraints on use (register, frequency, etc.)	Receptive	Where, when, and how often would we expect to encounter this word?
		Productive	Where, when, and how often can we use this word?

Source: Nation, 2013, p. 49

Nowadays, Nation's proposal is the most widely accepted. Its main advantages are often related to the inclusion of the previous models, the incorporation of questions that facilitate the understanding of what each dimension refers to and the presence of the receptive-productive dichotomy that had been ignored in other approaches. For this reason, it is used as a framework of analysis and structural axis of many vocabulary research studies. As for its main drawbacks, Schmitt notes that this classification lists the aspects learners need to know, but it does not establish any "hierarchical ordering" (2019, p. 262), i.e., it does not sequence which components should be learnt first. In addition, the large number of components may make it difficult to apply to the analysis of vocabulary knowledge in a holistic way. This latter difficulty has been mainly overcome in two ways: on the one hand, some researchers have focussed exclusively on specific elements rather than explored the whole dimension since there is evidence that the different dimensions of word knowledge correlate (Milton & Fitzpatrick, 2014). On the other hand, other authors have grouped different elements in what is given the name of dimension. This view is related to the psycholinguistic approach identified by Jiménez Catalán (2002) and will be explained in greater detail in the following section.

1.2.2. *Psycholinguistic approach*

The second approach to the description of *lexical competence* encompasses a wide set of studies adopting a psycholinguistic perspective to define and explain the concept (Jiménez Catalán, 2002). The emphasis is placed on the learners' difficulty to develop their *lexical competence* and on the identification of the kind of knowledge and skills activated in the learning process. Therefore, from this perspective, *lexical competence* is analysed from a global viewpoint in order to reduce to workable proportions the list of elements that make up word knowledge (Milton, 2010).

A pioneering study in this sense came from Anderson and Freebody (1981), who, in a discussion about how vocabulary knowledge influenced reading comprehension, distinguished two main dimensions of vocabulary knowledge (*breadth* and *depth*):

> It is useful to distinguish between two aspects of an individual's vocabulary knowledge. The first may be called "breadth" of knowledge, by which we mean the number of words for which the person knows at least some of the significant aspects of meaning... [There] is a second dimension of vocabulary knowledge, namely the quality or "depth" of understanding. We shall assume that, for most purposes, a person has a sufficiently deep understanding of a word if it conveys to him or her all of the distinctions that would be understood by an ordinary adult under normal circumstances.

> Anderson & Freebody, 1981, pp. 92–93

Meara (1996a) also pointed to the existence of these two critical dimensions but giving them a different label: *vocabulary size* and *organisation*. According to Meara, the former was the basic dimension of lexical competence, and referred to the amount of vocabulary that a learner knows, whereas the latter referred to how well these words were known. Moreover, he went a step further and also reflected on the importance of the interaction between both dimensions, concluding that while in the first case, as vocabulary known by learners became larger, the importance of size decreased, the opposite happened in the second case: the larger the vocabulary size became, the greater importance to the *depth* or *organisation* of this vocabulary was given.

Since Meara's contribution, subsequent proposals included these two dimensions in some way or the other. For example, Chapelle (1998) developed a four-dimensional framework in which *vocabulary size, knowledge of word characteristics, lexicon organization* and the *processes of lexical access* defined lexical competence. The first two dimensions matched up with Meara's proposal of *organization* and *size* of vocabulary, whereas the remaining dimensions referred, respectively, to the way lexical knowledge was stored in the mental lexicon and to the correspondence between the word stimulus and its mental representation (Thoma, 2009). Likewise, Qian (2002) considered vocabulary knowledge as a four-dimensional framework made up of *vocabulary size, depth of vocabulary*

knowledge, lexical organization and *automaticity of receptive-productive knowledge*. Again, *size* and *depth* were also included. In her model of lexical competence, Henriksen (1999) reduced the dimensions to three, distinguishing between: *partial-to-precise knowledge* (*breadth* of vocabulary knowledge), *depth of knowledge* (relationship of a word to other words) and *receptive-to-productive dimension* (level of mastery of vocabulary knowledge, reflected in learners' comprehension and production abilities). Meara (2005) added a third element to his previous proposal: *vocabulary accessibility*. With this new dimension, he made reference to the automaticity with which a learner could access the knowledge of a word. Daller, Milton and Treffers-Daller (2007) took up Meara's ideas and put forward a three-dimensional framework consisting of *lexical breadth, lexical depth,* and *lexical fluency* (see Figure 2.6 below).
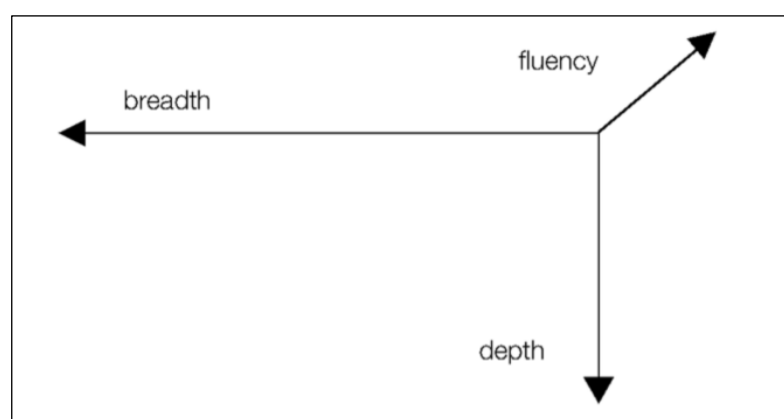
Figure 2.6. The lexical space (Daller et al., 2007, p. 8).

In Daller et al.'s model, the horizontal axis represents the concept of *lexical breadth*, which is related to Meara's and Anderson and Freebody's *breadth* and refers to the number of words a learner knows. The vertical axis represents the concept of *lexical depth*, or, in other words, how much the learner knows about the words. Finally, the third dimension refers to the concept of automaticity when using the words the learner knows. The novelty of this proposal lies in the relation of these three dimensions with the different elements of word knowledge proposed by Nation (2001). In their opinion, two dimensions are closely related to some of the elements proposed by Nation: *lexical breadth* would include *form* and *form and meaning*, whereas *vocabulary depth* would comprise *concepts and referents, associations, grammatical functions, collocations* and *constraints on use*. However, the third dimension would not be related to Nation's proposal, but to L2 word networking, that is, the way lexical items are organised in the mind. Lexical knowledge is seen as a network with elements connected by semantic relations (Azizi et al. 2012; Meara, 2007a, 2007b; Wilks & Meara, 2002). The stronger these relations are established in the mental lexicon, the faster and easier the access. The inclusion of this third dimension gives a new perspective to the construct of *lexical competence*: it does not only entail 'word meaning', in the more traditional view of the concept, but it sheds new light on lexical competence, as it incorporates learners and their capacity to access vocabulary knowledge to the equation.

In short, the *lexical competence* construct has been widely explored in the last decades of the twentieth century, and, being the large number of proposals, the comprehension of lexical competence seems to remain constant, given that vocabulary *size* and *depth* are recognised as the two main dimensions. In fact, Milton (2010) states that these two dimensions summarise and comprise most of the proposals regardless the kind of approaches examined. Given the relevance of both concepts, in the following sections, they will be thoroughly analysed, paying particular attention to how their conceptions are reflected on second language acquisition research.

1.2.2.1. Vocabulary size

Vocabulary size, the first dimension of vocabulary, is related to the number of words a learner knows, even if rarely (Anderson & Freebody, 1981; Hatami & Tavakoli, 2012; Meara, 1996a). Milton (2010) relates this dimension to *word form* and *form and meaning* included in Nation's proposal.

Studies on this area have explored native and non-native speakers' vocabulary size with different purposes. In the case of native speakers, the aim has been to measure the number of words in some absolute terms. In general, educated adult native speakers of English have been shown to master a number ranging between 16,000 (Goulden et al., 1990) and 20,000 word families (Schmitt, 2010), which in practice means that native speakers incorporate word families to their lexicon at a rate of approximately 1,000 per year up to young adulthood.

In the case of L2 learners' vocabulary size, the research has dealt with two different but complementary purposes (Vilkaite-Lozdiene & Schmitt, 2020). On the one hand, a strand of these studies has focussed on identifying the percentage of words needed to gain adequate comprehension. On the other hand, certain authors have explored the frequency profile of specific texts. The research has been conceived to help language teachers in their main tasks. Thus, given the impossibility of teaching all the lexicon of a language, language teachers need to prioritise the kind of vocabulary learners had to acquire in order to be functionally efficient in the foreign language, and, at the same time, they need to know the lexical difficulty of the texts with which learners are faced.

The first step in this identification is the estimation of the lexical thresholds for comprehension in written and spoken texts (Vilkaite-Lozdiene & Schmitt, 2020), i.e., the percentage of words needed to fully comprehend a text. At first, Laufer (1989) claimed that, in order to understand real second language spoken and written texts, a coverage of between 95% (one unknown word in every 50) and 98% was required. Later, she suggested that the knowledge of the most frequent 3,000 word families would be required to comprehend real written texts (Laufer, 1992). However, these figures are now being refined, and that the idea that there are different thresholds is becoming more and more accepted. In general, Vilkaite-Lozdiene & Schmitt (2020) find that there are two main lexical thresholds depending on the coverage aimed at: an optimal one, that places the percentage needed at 98% (Hsuch-Chao & Nation, 2000; Nation, 2006; Schmitt, Jiang & Grabe, 2011), and a minimal one, in which the

percentage of understanding decreases to 95% (Laufer & Ravenhorst-Kalovski, 2010). Although the difference between both coverages is only of 3%, this apparently small variation results in a very important difference in number of family words L2 learners have to recognise to understand a text. Therefore, setting an appropriate and realistic threshold consistent with learners' needs is essential and more research should be carried out to clarify whether an understanding of 95% of any piece of text is enough or, on the contrary, a 98% is required.

Once the threshold is established, the second step is the estimation of the number of words needed to achieve the threshold, and this number will depend on the kind of discourse, written or oral, examined. In the case of oral English, reaching conclusions is not as easy, and researchers enter the realm of estimations. If we take 95% as the level of vocabulary needed for everyday communication, results obtained suggest that the mastery of between 2,000 to 3,000 word families (Adolphs & Schmitt, 2003) could suffice. If 98% coverage is aimed, the number of word families needed rises to a range of between 6,000 and 7,000 word families (Nation, 2006). In the case of written texts, the figures are higher and may vary depending on the genre. In the case of novels, Hirsh and Nation (1992) found that a knowledge of about 5,000 word families was needed to achieve the optimal threshold. Nation (2006) concluded that, in order to reach 95% coverage, a knowledge of 3,000 word-families was needed. Moreover, he emphasised that these word families were not randomly chosen, but they were specifically the most frequent 3,000 words.

The classification of word families into frequency-based categories was central in order to facilitate the study of learners' development. One of the first and most influential classifications was developed by Nation (2001). Given its relevance to vocabulary research, this classification has been taken up by other authors, such as Schmitt (2010) or Schmitt & Schmitt (2014). Nation divided vocabulary into different levels on the basis of frequency and range: high frequency words, academic words, technical words and low frequency words. More specifically, these levels can be defined as follows:

- High frequency words: are those words occurring very frequently in all kinds of registers. They are usually the 2,000 most frequent words of the English language, covering approximately 80% of the running words in any academic text (Schmitt, 2010).

- Academic words: are those words frequently found in all kinds of academic subject areas. The best-known lists are the University Word List (UWL; Xue & Nation, 1984) and the Academic Word List (AWL; Coxhead, 2000). In academic texts, Schmitt (2010) notes that these words account for about 9% of the running words.

- Technical words: are the lexical items that are very common in one particular area. They make up about 5% of the running words in academic texts.

- Low frequency words: are those occurring infrequently in the different registers of a language. In academic texts, they account for over 5% of the words.

However, this classification can become, at some point, somewhat problematic. First, it is seemingly based on frequency, but this is only clear in the case of the first group, which is set at the 2,000 frequency level. In the case of low frequency words, there is a lack of consensus, as it has been characterised in different ways, ranging from those words not included in the high-frequency word families up to all the word families beyond the 10K frequency level. Moreover, Schmitt and Schmitt (2014) point out an additional problem: whereas general vocabulary is usually divided into 1K frequency ordered categories, academic and technical vocabulary are composed of word families belonging to different frequency word-bands, which means using a different criterium to establish the categories. Given the problems that the use of Nation's classification entailed, Schmitt & Schmitt (2014) made a new proposal based on three levels of word frequency: high-frequency, mid-frequency and low-frequency words.

- High-frequency words: despite the fact that traditionally the most frequent 2K word families constituted the cut-off point for high-frequency vocabulary, in this classification, Schmitt and Schmitt decided that it was more accurate to expand it to the most 3,000 word families. They supported their decision with empirical evidence from corpus linguistics which seemed to indicate that the knowledge of the 3K most frequent word families facilitates 95% coverage in spoken discourse (Adolphs & Schmitt, 2003; Nation, 2006; Van Zeeland & Schmitt, 2013) and 98% of the words in most graded reading materials (Webb & Nation, 2017) Thus, it may be considered that the 3K band represents the basic learner lexicon (Cobb, 2007). In relation to the kind of words that make up this group, both Nation (2006; 2001) and Schmitt and Schmitt (2014) show that part of those most frequent words correspond to grammatical items.

- Mid-frequency words: the authors include this new level of frequency which encompasses word families belonging to the bands situated between the 3K+ and 9K. The importance of this band lies in the capacity it gives leaners to participate in a wider range of activities across a range of topics and situations. Its knowledge allows learners to go beyond 95% coverage and facilitates the engaging with English for authentic purposes.

- Low-frequency words: this band is made up of word families which do not occur frequently in English. Schmitt and Schmitt (2014) suggest situating the boundary of this band in the 9,000 word families, based on Nation's (2006) and on analyses of the Corpus of Contemporary American English (COCA; Davies, 2008).

Coverage and corpus studies, and, particularly, the classification of vocabulary into frequency levels, have had strong implications for teaching practice (Schmitt, 2010). They are essential for understanding language learning and they show that that the acquisition of the first bands is a requirement to be functional in English. Moreover, it also clarifies that mastering key vocabulary will definitely be beneficial for learners, as they will be able to understand a wide range of texts, in which the same kind of words appear. Supporting this idea, Milton states that "the knowledge of the most frequent words in

the foreign language appears crucial to successful performance" (2010, p. 218). These are some of the reasons why this PhD dissertation focusses only on the analysis of the mastery of those most frequent words (2K) in order to explore the *lexical competence* of secondary school learners.

1.2.2.2. Depth of vocabulary

The second dimension of vocabulary knowledge —vocabulary depth— is not defined as easily as breadth. As already pointed out, depth as a dimension of vocabulary learning is common to most of the studies on vocabulary knowledge. Anderson and Freebody (1981, pp. 92-93) were the first to point out to the existence of what they called "depth" of vocabulary, when saying that "[there] is a second dimension of vocabulary knowledge, namely the quality or "depth" of understanding. We shall assume that, for most purposes, a person has a sufficiently deep understanding of a word if it conveys to him or her all of the distinctions that would be understood by an ordinary adult under normal circumstances". According to their view, *vocabulary depth* refers to the quality of the knowledge. However, this definition seemed to be rather vague due to the difficulty of measuring the so-called 'quality' of understanding, and this explains why other authors have attempted to narrow down the concept. For example, Meara and Wolter (2004) give a different view on the matter when defining *depth* in terms of the number of links between words and the networks words can create and, similarly, Milton (2013) explains the term by relating it with Nation's proposal of dimensions of word knowledge, specifically, to the *associational knowledge, collocational knowledge, inflectional and derivational knowledge, knowledge of concepts and referents,* and *knowledge of constraints on use*.

In other words, whereas in the case of vocabulary size, it is clear that it refers to a superficial knowledge of a word, in the case of vocabulary depth, there is a lack of agreement on what it clearly refers to. In 2004, Read re-elaborated the notion of vocabulary depth and incorporated two new dimensions to the ones presented by Anderson and Freebody (1981). In his view, vocabulary depth is clearly related to three different dimensions: *precision of meaning, comprehensive word knowledge* and *network knowledge*. The first dimension corresponds to the same idea presented by Anderson and Freebody (1981). The second dimension, the so-called *comprehensive word knowledge*, refers to the word knowledge dimensions already explored in this chapter, and encompasses the dimensions related by Milton (2013). As for the third component, known as *network knowledge*, it refers to the capacity of incorporating those new words to a network of already-known words. These concepts are defined in Table 2.4 below.

Table 2.4

Lines of development in the application of depth to L2 vocabulary acquisition

Lines of development	Definition
Precision of meaning	"The difference between having a limited, vague idea of what a word means and having much elaborated and specific knowledge of its meaning" (2004, p. 211).
Comprehensive word knowledge	"Knowledge of a word which includes not only its semantic features but also its orthographic, phonological, morphological, syntactic, collocational and pragmatic characteristics" (2004, p. 211).
Network knowledge	"The incorporation of the word into a lexical network in the mental lexicon, together with the ability to link it to —and distinguish it from— related words" (2004, p. 212).

Source: own elaboration, adapted from Read, 2004, p. 211-212.

With the inclusion of these two new elements, this author, on the one hand, incorporated the ideas proposed by Meara and Wolter (2004) and anticipated the ideas later developed by Milton (2013), and, on the other hand, embraced the two approaches to *lexical competence* (linguistic and psycholinguistic) as he considers the micro-analysis, i.e., word knowledge, as one sub-dimension of vocabulary depth while maintaining the global vision of lexical competence. Given that it is the most inclusive proposal, and, in the absence of more concrete and accepted views, this thesis dissertation will adopt Read's vision of vocabulary depth, which in conjunction with Nation's proposal of word knowledge, which will help us to start from an accurate and complete definition of lexical competence.

In short, according to the different definitions analysed in this section, it could be said that lexical competence may be defined as the ability to recognise and use words, which implies the knowledge of a number of more concrete aspects, such as the word form, meaning and use. However, most definitions of lexical competence here presented are clearly related to vocabulary knowledge. In this sense, terms such as lexical competence, vocabulary knowledge, vocabulary size, or word knowledge are used indistinctly. This may lead to a misunderstanding of what lexical competence is and what it involves. As stated in Chapter One, from a psycholinguistic point of view, competence cannot be used as a synonym for knowledge, as it encompasses not only knowledge, but also applying this knowledge. One of the main objectives of this PhD dissertation is to analyse secondary-school learners' lexical competence and, therefore, it is key to provide a clear understanding of what it encompasses. For this reason, the following section is devoted to clarifying the difference between lexical competence and knowledge.

1.3. Lexical Competence and Lexical Proficiency: two faces of the same coin

Throughout this section, *lexical competence* has been equated time and again to lexical knowledge and lexical proficiency, and how most of its descriptions rather refer to lexical proficiency. Once the origins of the construct as well as its description have been explored, it is time now to turn back to the main objective of this section: providing a clear definition of *lexical competence*.

In this sense, the work by Bulté, Housen, Pierrard, and Van Daele (2008) may help the differentiation of both concepts. In their view, the main difference between *lexical competence* and lexical proficiency lies in the nature of both constructs. For them, *lexical competence* is a theoretical construct which consists of two components: declarative and procedural. The former has to do with lexical "knowledge", whereas the latter is related to how lexical information is stored in the mind. However, this competence cannot be directly observed or measured, so they propose a second level of knowledge, known as Lexical Proficiency, in which behavioural manifestations resulting from the underlying construct are observed. From my point of view, this differentiation of the Lexical Competence as a theoretical construct and Lexical Proficiency as its observable manifestation presents clear advantages: (1) it prevents the different problems that may arise from the use of misleading terminology and (2) it helps to delimit the coverage and areas of study.

Having made this point clear, in my view, this distinction is not enough to provide a complete definition of *lexical competence* and the construct needs to be specified. In my opinion, the term 'lexical competence' needs to evolve in parallel to the concept of *communicative competence*. The chronological analysis of the different communicative competence models has shown an evolution from a first stage in which competence is equated with proficiency, and a second phase, in which competence is seen as a more inclusive concept. This process should also occur in the conceptualization of lexical competence: it should be broadened to more than just the knowledge of words, and it should be turned into a more inclusive concept, in which other elements, such as the actions that L2 learners take or the abilities they have, could have their own space. This idea is also illustrated by Caro and Rosado-Mendinueta (2017), who note the importance of considering lexical competence as "a cluster of knowledge (form, meaning and use of a lexical item), abilities and skills that a person develops and deploys in different contexts of communication" (p. 207).

For this reason, from here on, I will differentiate between lexical knowledge and lexical competence. More concretely, I propose, as a first approach to a broader conceptualization of lexical competence, to adapt Bachman & Palmer's model of language ability. Following this model, lexical competence would be made up of two elements: lexical knowledge and lexical strategic knowledge. Thus, within the lexical knowledge, Reads' model of vocabulary depth (2004) and Nation's elements of word knowledge (2001, 2013) would be included, whereas in the case of strategic knowledge, a difference between communicative strategies, mainly used to recall the lexical items and communicate

in the L2, and vocabulary learning strategies, used by L2 learners to understand and retain new L2 items, would be established.
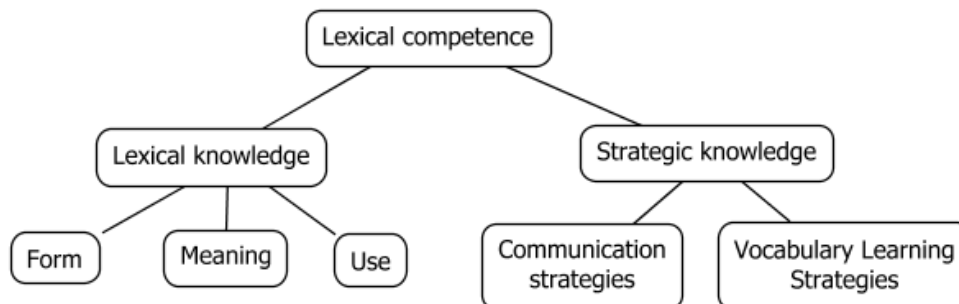
Figure 2.7. Model of lexical competence. Own elaboration.

Once the terminology has been clarified, in the following section I will move on to explain and discuss how lexical items are processed by L2 learners. To do so, I will start with a general explanation of the language learning process mostly following Leow (2015) to focus then on the two most relevant aspects of his model for the purposes of this PhD dissertation: intake processing and language storing.

2. Lexical Processing

Given that this PhD dissertation aims to explore L2 learners' lexical competence development in CLIL settings, by paying attention to both, learners' vocabulary knowledge and their learning process, it will need to focus on the analysis of the vocabulary knowledge these learners present, but also on exploring how they go about learning this vocabulary. Therefore, it is of paramount importance to understand the mechanisms used by learners to develop L2 lexical knowledge. In the last decades, given the emphasis placed on lexis and its relationship with language proficiency, a question regarding how L2 learners develop their L2 lexical competence has been raised: the issue of how lexical items are processed and stored in learners' minds.

As for the processing of lexical items in L2 learners' minds, one of the latest and most inclusive proposals of L2 processing comes from Leow (2015). This author conceives language learning as a chain consisting of several processes. Concretely, he identifies two external products (*input* and *output*) and five stages that make up the L2 learning internal process. Three of these stages are processes, known as *input processing, intake processing*, and *knowledge processing,* that generate two products: *intake* and *L2 knowledge.* The different products and stages are connected as shown in Figure 2.8 below:

INPUT	{ >	INTAKE >		INTERNAL SYSTEM	> }	OUTPUT
	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	
(Product)	(process)	(Product)	(process)	(product)	(process)	(product)
(input)	(input)	(intake)	(intake)	(L2 knowledge)	(L2 knowledge/output)	(representative L2 knowledge)
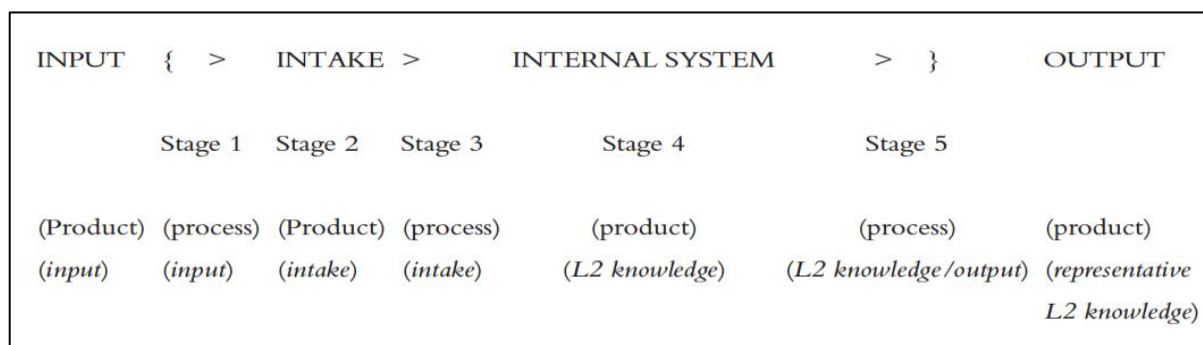
Figure 2.8. Leow's model of the L2 learning process in instructed SLA. Source: Leow (2015, p. 242)

According to his proposal, *L2 Input* refers to what "foreign language learners are exposed to, be it aural or written" (p. 17). This input is received by the learners, who need to process both the extralinguistic and linguistic data. This stage is thought to represent the initial phase of the learning process and, in it, L2 learners need to pay conscious attention to process the linguistic information in the input. Learners are thought to have a limited capacity for processing all the information they receive and thus, it prevents the acquisition of all the input received, therefore, *intake* is that specific input the learner has paid attention to. This leads to the second stage: intake as a product, in which part of the *input*, after the attention given by the learner, becomes *intake*. This *intake* is processed for its incorporation in learners' internal systems at stage 3, so that it becomes *L2 knowledge* and is stored in the internal system at stage 4. Finally, the last stage of the internal processing is *knowledge processing*. In Leow's words, "this stage deals with learners' manipulation of the L2 linguistic knowledge, together with other knowledge bases that govern, for example, phonological, syntactic, semantic, cultural, pragmatic, and discourse features that register aspects of the L2 language that are employed to produce the L2" (2015, p. 20). The final product is the *output*, that reflects the learners' interlanguage or internal systems, which, in the context of this piece of research, refers to vocabulary items. Therefore, this model conceives language learning, and, in consequence, vocabulary knowledge, as an overall process consisting of several mental sub-processes in which different elements are involved. Two of these elements –intake processing and learner's storing– are central to this PhD dissertation. For this reason, in the following subsections they will be explained in detail, and they will be used as a framework to conceptualise the different elements explored in this dissertation.

2.1. Intake processing

Intake processing corresponds to the third stage of Leow's model and refers to the process in which intake is treated in L2 learners' minds to finally become L2 knowledge. It is, therefore, a mental process in which different sub-processes intervene and different factors determine the product, in this case, the L2 vocabulary knowledge. In order to provide a model that fully explains the processing of L2 input, Leow developed a more detailed schematic representation (see Figure 2.9 below) of his model, in which he included the different mental processes that intervene in each stage.
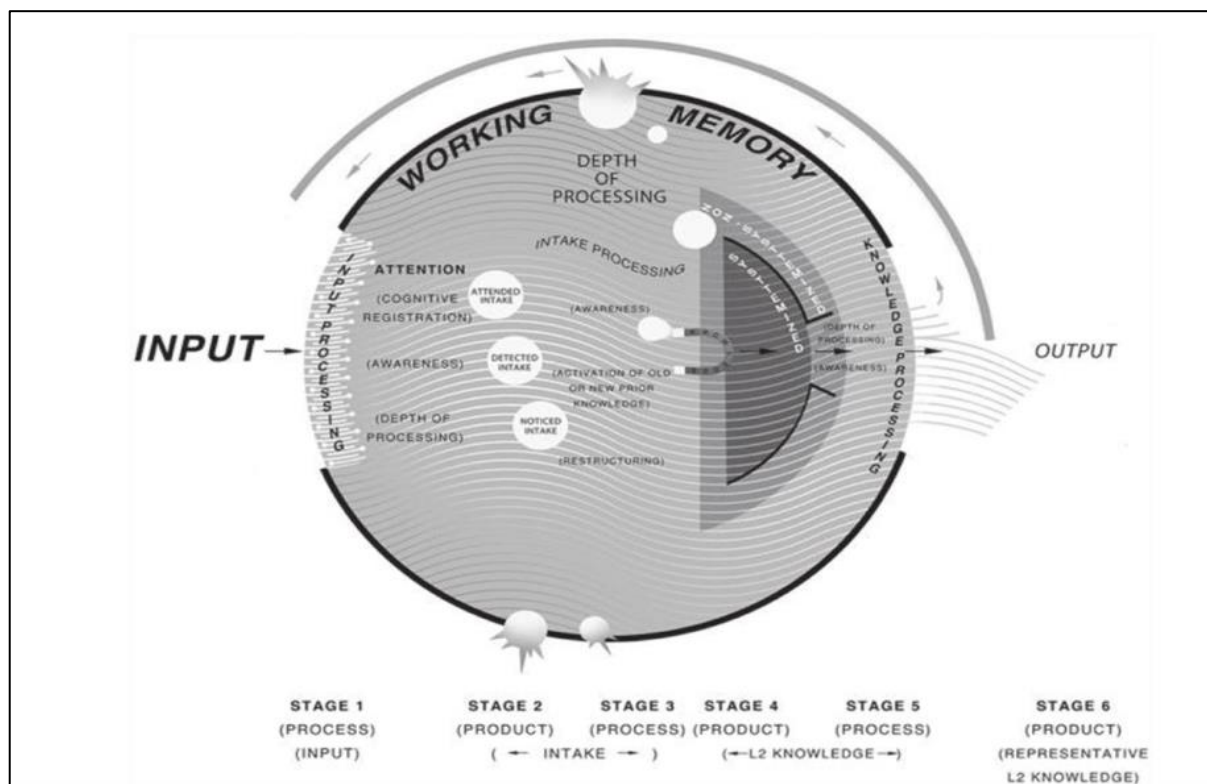
Figure 2.9. Leow's detailed model of the L2 learning process in instructed SLA. Source: Leow (2015, p. 242).

As can be seen in this figure, whereas in the first stage –*input processing*–, attention or noticing (Schmidt, 1990) drives the process, in the case of *intake processing*, there is also a key issue –depth of processing– and other more concrete elements that are somehow related to it and also have their role in the process: learners' awareness, activation of old or new prior knowledge and information restructuring.

There are two concepts whose relationship should be first clarified to fully understand the intake processing stage: depth of processing and awareness. As for awareness, one of the most widely accepted definitions is that of Schachter (1989), who states that awareness "refers to a state of mind in which one has become cognizant of the regularities underlying the data" (p. 577). Leow (2001) gives a more pragmatic definition of 'awareness' by establishing some criteria to determine the presence of awareness. In his view, awareness is present if there is, at least, "(a) some resulting behavioral or cognitive change, (b) a meta-report of the experience but without any metalinguistic description of a targeted underlying rule, or (c) a metalinguistic description of a targeted underlying rule" (Leow, 2015, p. 185). Moreover, in later work, he presented a systematic and thorough revision of the awareness literature, from which he drew six important conclusions for language teaching:

(1) awareness at the level of noticing and understanding contributed substantially to a significant increase in learners' ability to take in the targeted form or structure (Leow, 2000, 2001; Medina, 2015; Rosa & Leow, 2004; Rosa & O'Neill, 1999) and to produce in

writing the targeted form or structure (de la Fuente, 2015; Leow, 2001; Medina, 2015; Rosa & Leow, 2004), including novel exemplars (Rosa & Leow, 2004); (2) awareness at the level of understanding led to significantly more intake when compared to awareness at the level of noticing (Leow, 2001; Medina, 2015; Rosa & Leow, 2004; Rosa & O'Neill, 1999); (3) there is a correlation between awareness at the level of understanding and usage of hypothesis testing / rule formation (Hsieh et al., 2015; Leow, 2000, 2001; Rosa & Leow, 2004; Rosa & O'Neill, 1999); (4) there is a correlation between level of awareness and formal instruction and directions to search for a rule (Rosa & O'Neill, 1999); (5) there is a correlation between awareness at the level of understanding and learning conditions providing an explicit pre-task (with grammatical explanation), as well as implicit or explicit concurrent feedback (Rosa & Leow, 2004); and (6) there is a strong correlation between reported awareness and comprehension and production scores (de la Fuente, 2015).

On the other hand, Depth of Processing is defined as "the relative amount of cognitive effort, level of analysis, and elaboration of intake, together with the usage of prior knowledge, hypothesis testing, and rule formation employed in decoding and encoding the same grammatical or lexical item in the input" (Leow, 2015, p. 204). This concept seems to have its origins in the Cognitive Psychology field, especially in Craik and Lockhart's model of L1 processing (1972). In their view, remembering information depended on several factors such as the degree of attention during its occurrence, the number of times of rehearsal and the way the information was processed. In this sense, they distinguished two types of processing: shallow processing and deep processing. Shallow processing was related to either structural or phonemic processing and did not ensure a strong retention of the item. On the other hand, deep processing had to do with the understanding of the item in relation to its context and meaning and the potential relationship with other items. After a series of experiments, they concluded that a deeper level of information processing resulted in greater chances of remembering this information. This concept has been also applied to L2 learning, and several studies have explored Depth of Processing in relation to other elements such as mental effort (Calderón, 2013; Hsieh et al., 2015) or levels of awareness (Leow, 2012; Hsieh et al., 2015).

Regarding the relationship between Depth of Processing and levels of awareness, after this brief overview about awareness and processing, it seems clear that awareness is related to understanding and the different levels that it may contain, whereas Depth of Processing is related to the cognitive demand and effort used to elaborate knowledge. However, although there are differences between both concepts, there is some evidence (Leow, 2012) that supports the existence of a correlation between depth of processing and level of awareness, i.e., a higher depth of processing usually comes with a higher level of awareness, whereas in those processes where a lower depth of processing is observed, a lower level of awareness is also shown.

Once the distinction between the two key concepts has been clarified, I can turn back to the explanation of L2 language processing based on Leow (2015). He distinguishes two processes by which the data is internalised in the L2 language system depending on the degree of processing and the cognitive effort. Moreover, he also takes into account the number of times that the linguistic data has been presented to the L2 learner. In his view, there is a difference when the data has been already processed by the L2 learners in comparison to the first exposure to the element. If it is the first time the exemplar is presented to the L2 learner this intake may be encoded in the L2 system as a non-systemized chunk of language when there is a minimal data-driven processing (Leow, 2015) awaiting further exposure and information. This is, in Leow's conception, data-driven processing. On the other hand, other elements would intervene in the process if higher levels of processing and awareness take place. At the same time, this higher processing may occur in one of the following ways: by the activation of prior knowledge, i.e., a conceptually-driven processing, or by the connection to new knowledge. As for the activation of prior knowledge, it seeks to facilitate the process of encoding and decoding of linguistic information. Leow represents it as a magnet in his model (see Figure 2.9 above), he sees this activation of prior knowledge as a way of attracting the understanding and processing of new knowledge. Regarding the activation of new knowledge, the un-systemized data previously stored in the L2 developing system may be reactivated (Leow, 2015; Leow, 1998) with further exposure to the linguistic element and after some processing, it can become part of the L2 mental system. In both cases, this activation of prior or new knowledge can lead to what he calls 'restructuration', a concept that corresponds to McLaughlin's second phase (1987). In McLaughlin's view, the L2 learner needs to impose an organization on and structure the information that is being learnt. Thus, once the L2 learner activates his or her previous or new knowledge, he or she can then re-organise or broaden his or her idea of the concept. The activation and restructuration of information can occur in two ways: either explicitly or implicitly, depending on the depth of processing and level of awareness. If it occurs through a low level of processing, this may lead to an implicit restructuring and systemized learning, a less cognitively-demanding process. On the other hand, if this restructuration takes places through explicit learning, a higher depth of processing, together with an increased level of awareness is needed. The different possibilities of intake processing taking into account the level of exposure are summarised in Figure 2.10:
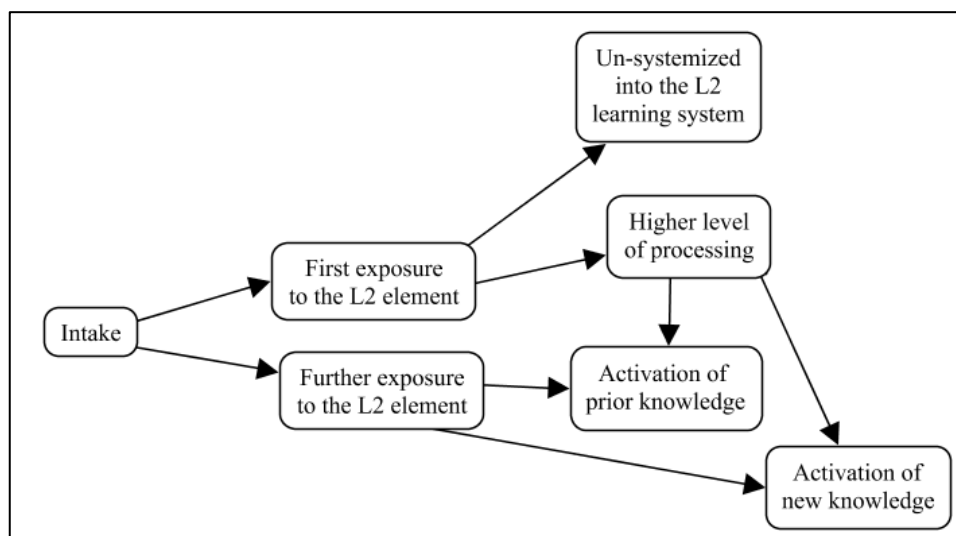
Figure 2.10. Intake processing diagram. Source: own elaboration based on Leow (2015).

In short, intake processing seems to be a complex concept that involves a series of mental sub-processes, being the activation of prior and new knowledge and the restructuring of knowledge the main ones. The activation of the two processes depends on various aspects, such as the number of exposures to the L2 item and the degree of processing and awareness. In general, it is thought that the higher the depth of processing and the level of awareness, the better the understanding and knowledge of the L2 term. Therefore, L2 teaching approaches should seek learners' cognitive engagement in order to facilitate the learning process.

These mental processes have their observable manifestation and materialization in L2 learners' actions. Actions taken by L2 learners started to be explored in the late 1970s to identify the most used actions by good language learners in their process of L2 acquisition. This concern about the performance of good language learners would result in a specific area of research —Language Learning Strategies—, in which learners' actions would be explored in relation to other aspects, such as language performance, learning styles, language skills or learning tasks. Such is the connection between Learning Strategies and mental processes, that, for example, in the case of vocabulary learning strategies, one of the best known proposals, Schmitt's taxonomy (1997), follows the same distinction between the reaction to first exposure to an item and the processing of the following exposures to the same item when distinguishing between strategies used to face the first encounter with an unknown L2 lexical item (discovery strategies), and strategies used to retain and expand the knowledge of a given L2 lexical item (consolidation strategies). As stated in the introduction, this PhD dissertation aims to explore L2 learners' lexical development in CLIL settings. With this aim, L2 lexical items processing is going to be explored by means of the analysis of the selection of vocabulary learning strategies. Given the relevance of vocabulary learning strategies in this PhD dissertation, this notion will be taken up again and discussed in depth in the next Chapter.

2.2. Lexical storing in the multilingual mind

Our next section deals with how vocabulary is stored in the learners' minds, which would correspond to *'stage 4: product L2 knowledge'* or *'internal system'* in Leow's model. Traditionally, this knowledge is usually thought to be located in the so-called mental lexicon (Aitchison, 2012; Dóczi, 2020; Elman, 2004; Pavlenko, 2009) which is where we find the information about the phonological, semantic, morphological or syntactic features of every single word a person knows (Field, 2003). Moreover, there seems to be a consensus that all this information is not randomly stored and organised, but, on the contrary, structured in "a logical way" (Dozci, 2020). However, conclusions in this swampy terrain are far from definitive. Firstly, given the nature of the mental lexicon, scholars find it difficult to agree on a model. The proposals made (see Table 2.5 below for an overview, in which their stronger and weaker points are highlighted) show a progression in their development. Most models are based on their predecessors and attempt to provide solutions for the weaknesses of previous designs. This is clearly observed, for example, in the case of the Semantic Features Model (SFM; Smith et al. 1974). When the Hierarchical Network Model, one of the first contributions, expressed two main concerns related to the impossibility of accounting for familiarity and typicality effects, the authors of the SFM provide a solution to the problem of the typicality effect: the storing of concepts as groups of attributes, which could be either defining and characteristic features. Similarly, the Revised Spreading Activation Model (RSAM) solved some of the problems of its parent model, the Spreading Activation Model (SAM) by adding a new comprehension of the term concept: in this model, apart from the conceptual connections, various levels of an entry are marked for syntax as well as phonology.

Table 2.5

A summary of the L1 mental lexicon models

Model	Description	Strengths	Weaknesses
Hierarchical Network Model (Collins & Quillian, 1969)	Information is organised in categories that are linked to one another and organised hierarchically.	- There is a featured orientation of the concepts.	- Familiarity effects. - Typicality effects. - It fails to predict reaction times.
The Semantic Features Model (Smith et al., 1974)	Concepts are stored as groups of attributes, but there exist two types of attributes: defining and characteristic features.	- It solves the problem of typicality.	- It fails to predict reaction times. - Familiarity effects.
The Spreading Activation Model (SAM; Collins & Loftus, 1975)	"[It] assumes a complex network of concepts connected by various types of relationships with varying degrees of strength". (Monaikul, 2015, p. 9)	- It presents mental lexicon as a network of associations.	- It assumes that people store concepts, making it difficult to consider other aspects of word knowledge. - It assumes that every single person presents a unique mental lexicon.
The Revised Spreading Activation Model (RSAM; Bock & Levelt, 1994)	Similar to the SAM, the RSAM presents three levels of nodes: conceptual, lemma, and lexeme levels.	- It solves the SAM problems. - Links between words can be made according to syntax phonology, semantics, or orthography.	

Nowadays, one of the most accepted models is the RSAM (Bock & Levelt, 1994). This model brings together most of the previous proposals and considers the mental lexicon as a network of associations, which implies that each word may be connected to other lexical items. As the SAM, it posits three levels in a lexical entry —the conceptual, lemma and lexeme levels— with links to syntactic and phonological information. This model also hypothesises that lexical items may be connected to other items through this linguistic information. The following figure is an example of a subnetwork of the RSAM model. As can be seen, there are two words, 'sheep' and 'goat', and they are defined and

connected at the conceptual, lemma and lexeme levels. The conceptual level includes some key concepts that both terms share, but also some aspects that differentiate one from the other. As can be seen, 'sheep' is connected with 'wool', 'milk' and 'animal', whereas 'goat' is only linked to 'milk' and 'animal'. According to this view, the meaning of a term is constructed in opposition to another term, with which it shares some common senses. At the lemma level, syntactic information, such as gender and part of speech, is found. In this case, both terms coincide in part of speech, as both of them are nouns, but they differ in gender, as 'goat' is used to refer to females, whereas 'sheep' is used for males. Finally, at the lexeme level, a phonological and written representation of both lexical items shows, once again, the difference between them.
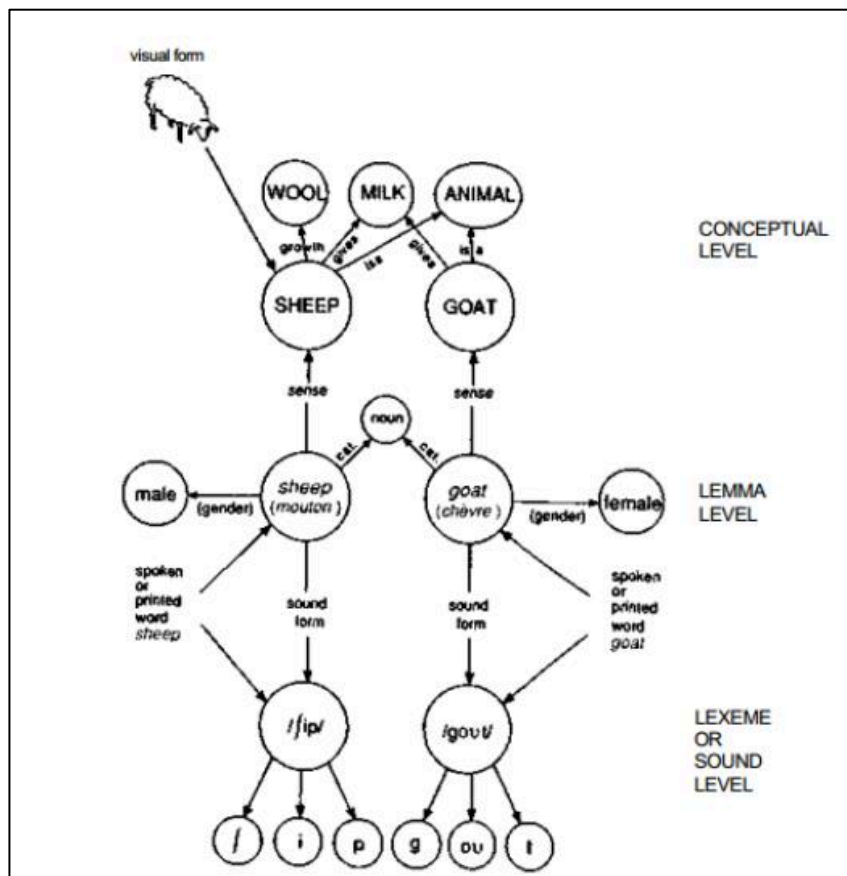
Figure 2.10. A sample multi-level subnetwork of the RSAM reproduced from Bock and Levelt. Source: Monaikul, 2015, p. 11.

The second major concern in the conceptualization of the mental lexicon is the understanding of how bilingual and multilingual networks work. Whereas organizing and representing the lexical network of a single language has been proven to be a demanding issue, the understanding of how lexical networks of different languages interact and co-habit in the same lexicon constitutes a much more challenging task. There are two opposing views on this issue. On the one hand, there are some researchers (Meara, 1982, 1984; Schmid, 2002; Singleton, 2007) who consider that the different languages are stored separately. To support their view, they give different arguments: for example,

Meara (1984) justifies the existence of different lexicons for the different languages by examining answers to word association tests, which shows that semantic links are "qualitatively different" (Dóczi, 2020, p. 54). Similarly, Singleton (2007) argues that multilingual speakers manage languages that present differences in morpho-syntactic and phonological structures. In his view, the use of the different languages implies different routes for lexical access and activation. Finally, there are also some cases of multilingual aphasic patients who only lost one of their languages, which supports the idea that both systems are completely autonomous (Schmid, 2002).

On the other hand, there is a larger proportion of researchers who conceive the relation of the different linguistic systems as an interactive network system, in which there is a shared lexicon (Dóczi, 2020). To support this common storage view, different arguments have been given, such as the proven existence of cross-linguistic influence between different languages, or the faster translation and activation of the L2 item if the morphological systems of the languages are similar. Within this conception of the mental lexicon, the problem is to determine the degree of interaction between the different systems. For this reason, different models have been developed taking as a basis two main ideas: (1) the models and theories generated for L1 mental lexicon and (2) the problems that arose in other L2 mental lexicon models.

One of the latest models is the *modified hierarchical model* (Pavlenko, 2009). From my point of view, it is one of the most comprehensive proposals, as, after a detailed identification of the strengths and weaknesses of the different models, Pavlenko articulates, in a single model, their main conclusions. She keeps (1) the idea that L1 and L2 networks are connected to each other, (2) the mutual transfer between the L1 and L2 (detailed in the Revised Hierarchical Model), and (3) the notion of shared and partially shared categories (conceived in the DFM and the SAM). However, at the same time, there are new elements as it differs from the previous models in the distinction between language-specific, partially overlapping, and completely shared conceptual representations. In this regard, she argues that the complexity of this model lies in the representation of L2 language-specific words, as it may be difficult for the L2 learner to develop a conceptual representation of a word if the term does not exist in his or her L1. Later revisions of the model (Dóczi & Kormos, 2016) have also argued that its complexity resides in the restructuring of the lexicon in those cases where the L2 concept does not correspond to any L1 lexical entry. However, at the same time, the same authors point to this categorization of concepts as one of the salient advantages of the model in comparison to the previous ones.
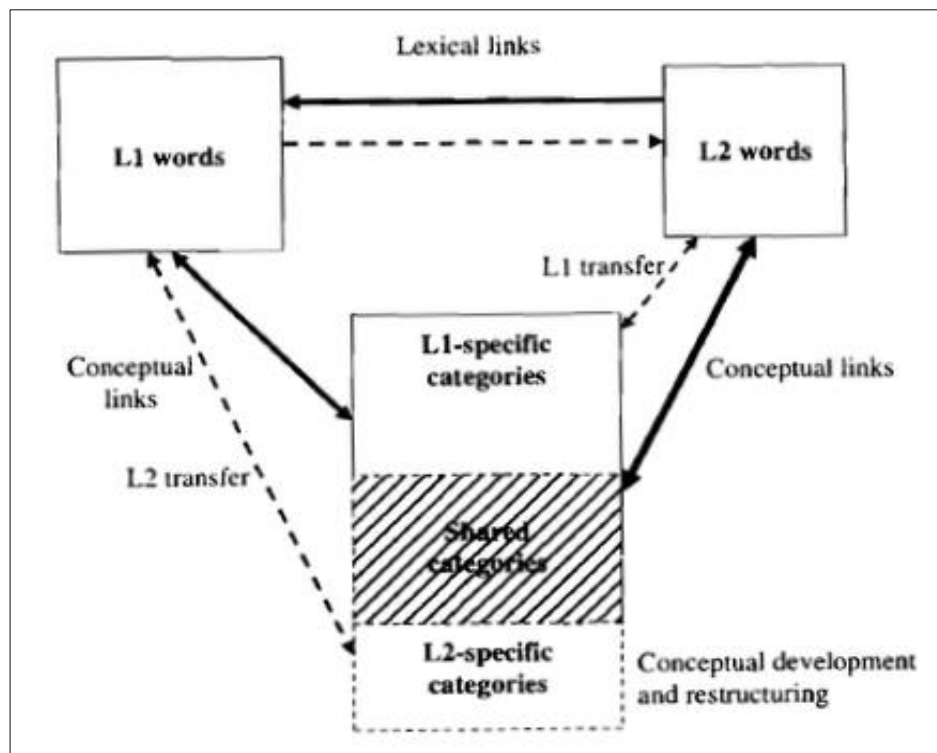
Figure 2.11. The modified hierarchical model. Source: Pavlenko, 2009, p. 147.

In general, despite the fact that there is still much in debate, there are some issues that seems to be clear: (1) the mental lexicon is flexible and in constant change; (2) L1 and L2 lexical items are connected with each other and they influence each other as well; (3) these items are also connected to other elements and (4) the L2 network is less stable than the L1 one. These ideas have clear implications for this PhD dissertation: as explained in the introduction, this piece of research aims to examine lexical development in CLIL settings, and, more concretely, lexical processing and its impact on the output along time. In this respect, it is important to stress that the mental lexicon and its component language systems, are, according to Leow's model, products resulting from mental processes and their observable manifestations: learning actions. In the 1980s, new language theories in which the language learning were examined emerged and were applied to language and vocabulary learning, resulting in a new research area, i.e., Language Learning Strategies. Language learning strategies are usually defined as actions that learners take to develop a foreign language, and it is commonly argued that these actions are merely reflections of mental processes. In general, these strategies relate to cognitive and metacognitive actions, in which some kind of manipulation of the language, such as linking between languages, is required. In this regard, these strategies echo the concept of strategic competence proposed by Celce-Murcia (2008), which I included in the discussion of lexical competence. Given their relevance for this PhD dissertation, the following Chapter will be devoted to language learning strategies and their application to vocabulary teaching, i.e., vocabulary learning strategies.

This section has attempted to explain how lexical knowledge is constructed and stored in the mind of L2 learners. Different models for processing and storing the information have been presented and analysed. I have concluded the section referring back to section one and the notion of lexical competence. After this introduction about what Lexical Competence is and how it works in learners' minds, the remaining sections of this Chapter will focus exclusively on lexical knowledge, one of the main components of Lexical Competence, whereas the second main component, strategic competence, will be taken up in the following Chapter. The following section deals with the issue of how this lexical processing is prompted in the EFL classroom.

3. Measuring vocabulary knowledge

Throughout this chapter, the complexity of the word knowledge construct has been shown time and again. This complexity has resulted in difficulties (1) when aiming to conceptualise word knowledge as presented in the first section of this chapter, (2) when modelling mental representation of the lexicon, as shown in section 2 or (3) when attempting to measure vocabulary knowledge. Focussing now on this latter difficulty, in section 1, I concluded that, from the different word knowledge proposals, Nation's model which defines nine word-knowledge dimensions (2001) seems to have certain advantages. However, measuring the development of each dimension in each word would be virtually impossible. Given the evidence that the different dimensions of word knowledge correlate (see, for example, Chen & Truscott, 2010; González-Fernández & Schmitt, 2019; Milton & Fitzpatrick, 2014; Schmitt, 1998; Webb, 2005), research on vocabulary measurement has focussed on specific dimensions, usually vocabulary size, rather than carrying out multicomponent studies. In other words, if the development of the different dimensions of vocabulary knowledge are related, exploring one of them, in this case, written recognition or production (i.e., vocabulary size), could be used as an indicator of general vocabulary knowledge.

For this reason, considering the purposes of this dissertation, only the main methods and tools to measure vocabulary size will be considered, leaving the other dimensions aside. To do so, I will start with an introduction on the importance of language testing, moving on then to the specification of the criteria to select vocabulary tests. The analysis of the appropriateness of the different instruments for the purposes of this study will be presented in Chapter Four.

Language testing is critical in L2 learning as a way to analyse the development of a foreign language. This is why it has attracted the attention of many scholars, who have mainly explored two areas: the development of L2 tests and the validation of the tools. Starting with the latter, there is some current debate on the need of developing more validation studies that ensure the suitability of the instruments for participants with different learning backgrounds (Schmitt et al., 2019). Schmitt et al. (2019) provide a critical revision on the topic of vocabulary assessment and note this lack of validation

studies when saying "we note that most tests, once launched, are not revised in any systematic way (or if they are, this is often hardly visible to the users)" (p. 2). In their view, every validation or revision must include the following points: in the first place, a score interpretation revision should be included, i.e., an interpretation of what the scores actually mean. These authors suggest that vocabulary test performance should be related to language use, for example, to skills such as speaking, reading, writing or listening. Secondly, they note that test developers should be careful when revising previous tests, as some of them have not been adequately validated. Moreover, the reliability, i.e., the consistency of the test score should be a third point of evaluation. This reliability implies that if a participant gets to know a specific number of word families on a particular day, this result should be similar in subsequent days. The fourth point of discussion is related to validation and piloting of the test in relation to the audience of the test. In the majority of the cases, tests are validated and piloted with adult university participants as these subjects are more readily accessed by the researcher. However, these instruments are not always used with this kind of participants, but other students, usually administered also to secondary or even primary-school learners. These participants may differ in their learning backgrounds or characteristics; therefore, it is important to check out whether these instruments are valid for them. Finally, they suggest that the launching of new tests should be accompanied by a substantial amount of validation evidence.
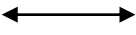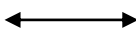
Focussing now on the former aspect, i.e., the development of tests, the first, modern, objective vocabulary test appeared back in the context of the Great War (Spolsky, 1995). Developed by Starch in 1916, the objective of the test was to measure vocabulary knowledge by asking test-takers to match a list of foreign words to their English translations (recognition knowledge). Years later, in 1964, the Test of English as a Foreign Language (TOELF) was launched, and it included a vocabulary section. Nevertheless, these examples were mere exceptions in an area dominated by grammar teaching. In the 1980s, as interest in vocabulary grew, so did the need of developing a number of tools designed for measuring it, and studies on frequency also served this purpose. These new tests were mainly based on the study on word frequency; thus, most tests measured specific bands of vocabulary.

A way to understand the difficulty of developing a test could be glimpsed by having a look at the different types of tests. Due to its complexity, there is not a straightforward and single way of measuring vocabulary, but different procedures with different aims and formats are found. For this reason, it becomes essential to explore the characteristics of each instrument in order to select the most suitable one. To do so, Read (2000) provides one of the clearest classifications of vocabulary tests up to date. He proposes three dimensions in order to identify the kind of test. The first dimension aims to distinguish between those instruments that focus on vocabulary as an independent construct (discrete tests), and those tools that are part of a larger measurement method aiming at measuring language competence in all their variables (embedded tests). The second dimension deals with the target vocabulary. Some tests (selective tests) focus on measuring specific target vocabulary, whereas others (comprehensive tests) explore all the items. The last dimension relates to the way the items are

presented, i.e., whether they are presented in a context that serves as a necessary clue for answering the test and that must be considered (context-dependent tests) or if examinees can produce their own response without considering any given context (context-independent tests). Table 2.6 below shows Read's classification of language tests.

Table 2.6

Read's classification of language tests

Discrete ←→ Embedded	
A measure of vocabulary knowledge or use as an independent construct	A measure of vocabulary which forms part of the assessment of some other, larger construct
Selective ←→ Comprehensive	
A measure in which specific vocabulary items are the focus of the assessment	A measure which takes account of the whole vocabulary content of the input material (reading/ listening tasks) or the test-takers' response (writing/speaking tasks)
Context-independent ←→ Context-dependent	
A vocabulary measure in which the test-taker can produce the expected response without referring to any context	A vocabulary measure which assesses the test-taker's ability to take account of the contextual information in order to produce the expected response

Source: Read, 2000, p. 17.

Although this classification is comprehensive and complete, other criteria have to be taken into consideration when selecting and interpreting vocabulary tests. First, it is important to identify the units that are counted in the test. The difficulty of defining the *word* as a unit led to a movement towards exploring lexical competence and the adoption of other units. For example, when it comes to measuring lexical development in written texts, one of the most widely used instruments is the token-type ratio. It is used for exploring the richness of a text and relates to different measures: tokens —i.e., the total number of word forms that make up a text, no matter their part of speech or meaning— and types — the number of different word forms. More advanced learners would avoid repetition by using synonyms or other related words. Thus, the higher the token-type ratio is, the higher level of lexical competence the learner presents.

Another question relates to the items to be counted as vocabulary, in particular in what regards to elements such as *the*, *a*, *in* or *that*. In general, articles, prepositions, pronouns or conjunctions have little meaning in isolation. They are more related to grammar than to vocabulary, which would explain why

they are known as function words. The difference between content words —those word forms with a 'content meaning'— and function words is of prime importance for vocabulary tests, as they mainly evaluate the knowledge of content words (nouns, adjectives, and adverbs). However, a different kind of problems arises when analysing only content words: what happens with the inflected and derived forms of a *word form*? Are they to be considered different *word forms*, or, on the other hand, should they be counted as one? As a solution, many scholars (Nation, 2001; cited in Schmitt, 2010) have opted for using the term *word families*, which are the set of word forms, closely related in form, with a common meaning. This distinction between word form and word families has strong repercussions on vocabulary measuring, but this is not the only decision researchers have to make, additionally, they also have to decide on how to deal with other aspects such as abbreviations, proper nouns, idioms and other multi-word units.

Third, the format of the test should also be considered. There are different test types for measuring the knowledge of vocabulary items, such as multiple-choice tasks, translation, matching or gap filling (Madsen, 1983; Read, 2000; Thornbury, 2002). Multiple-Choice Tests (MCTs) consist of a set of clusters, each of them containing "an item stem with a target word and a set of response options" (Gyllstad et al., 2015, p. 277). To complete them, test-takers are asked to relate the target word with one of the response options. They are one of the most widely used, because they are easily administered and scored (Gyllstad et al., 2015; Intaraprasert, 2004; Read, 2000). However, since they started to be used in SLA research, some criticism has also been levelled at them. Among others, Wesche & Paribakht (1996) pointed out as their main drawbacks the time devoted to their construction, their format and the bias produced by guessing. Indeed, developing such a test based on this format can be extremely time-consuming, as it requires a thoughtful and thorough analysis of the vocabulary to be tested. Besides, their format severely limits their sampling of the learners' total vocabulary. Finally, the odds of guessing (up to 25% depending in the number of options) the correct word without knowing it are large, and the learner may also be familiar with a different meaning for the word. Two of the most widely used tests —the *Vocabulary Levels* (Schmitt et al., 2001) and the *Vocabulary Size Tests* (Nation & Beglar, 2007)— use this format.

Similarly, the *checklist* format has widespread use for measuring vocabulary knowledge. In this format, test-takers have to tick those words they think they know, with the aim of measuring learners' development of the *written form* dimension of word knowledge. It is considered the simplest and oldest way of measuring vocabulary size (Read, 2000), and researchers have come to different conclusions about the benefits of this simplicity. On the one hand, some scholars consider that it may help test-takers as "it strips away irrelevant task demands that may make it difficult for young readers and poor readers to show what they know" (Anderson & Freebody, 1983, p. 235). On the other hand, other scholars consider that it may not be a reliable measure, as there is no way of knowing if test-takers are honest and they are reporting their knowledge of words. Studies on the reliability of checklist tests yield

conflicting results, as some of them find that this format do not correlate with other testing methods (Sims, 1929), whereas others (Tilley, 1936) obtain high correlations between checklist and multiple-choice tests. To solve this problem of reliability, some authors (Anderson & Freebody, 1983; Meara, 1992, 2010) have decided to add a proportion of items that are not real words (called pseudo-words), in such a way that, if the test-taker ticks a high number of those words, it can be considered as evidence that there is an overestimation of the vocabulary. The most representative example using this format is the Checklist test (Meara, 1992).

A third format used to measure vocabulary knowledge is the *matching of words with synonyms or definitions*. This format does not require a high level of ability like the ones already mentioned. Test-takers are asked to choose from a group of words or sentences, the one that represents the target word. In other words, these tests target the associative knowledge, i.e., the learners' ability to connect the item to other lexical items in the foreign language. They are similar in format to *translation* tests, as the same kind of ability is required. However, in translation tests, test-takers are asked to give a translation of the target word in their L1, or to choose, from a number of options, the most accurate translation. They are used to check the development of the *form and meaning* dimension pointed out by Nation (2001).

Finally, there is another format, the *completion* format, in which examinees are required to complete a prompt word taking into consideration the context. When developing completion tests, it is essential to have some control over the clues given in the prompt, by, for example, providing` the same number of letters or a clear context from which the word can be guessed. They measure the productive knowledge of the *form* dimension according to Nation's proposal of word-knowledge. This format is used in the *Productive Vocabulary Levels Test* (Laufer & Nation, 1999).

As can be seen, the kind of test and format will mainly depend on the dimension aimed to be measured. Therefore, when selecting a test for measuring vocabulary knowledge, the first step is to establish the target dimension. An inadequate selection of the test may result in misleading conclusions. For example, as will be explained in the following sections, the VLT was designed as diagnostic test examining receptive knowledge at specific levels. However, for decades, it has been used in multiple studies as a tool to measure vocabulary size. An analysis of vocabulary size comprises a deeper process than just the exploration of specific knowledge bands. if the real aim is to measure vocabulary size, it may not be appropriate to establish conclusions from the results obtained from the VLT. Instead, another test, the VST, may be better used for this purpose. Accuracy in the establishment of the target dimension is the basis of an appropriate selection of the tests. A second aspect to consider when selecting measuring tools is the kind of knowledge —receptive or productive— that is aimed to be measured. For example, the VLT and the PVLT measure the same dimensions —written form and meaning— but the VLT focusses on recognition and the PVLT measures the capacity of production. The following table organises the information given above about the different tests format in relation to the target dimension and gives information about several tests in which the different formats are used.

Table 2.7

A summary of the different test formats

Format	Target dimension	Tests
MCT	- Written form recognition - Form and Meaning recognition	- *Vocabulary Levels Tests* (Schmitt et al., 2001) - *Vocabulary Size Tests* (Nation & Beglar, 2007)
Checklist	- Written form recognition	- Checklist test (Meara, 1992). - Dialang Tests (2003)
Matching with synonyms	- Association	
Translations	- Form and meaning	- Vocabulary Size Tests (Bilingual versions)
Completion	- Written form production - Form and Meaning production	- Productive Vocabulary Levels Test (Laufer & Nation, 1999)

Lastly, another major problem researchers have to face relates to how to measure vocabulary in its entirely. Given vocabulary breadth, most vocabulary tests focus exclusively on learners' receptive or productive knowledge. This division —already presented in Palmer's (1921) and Nation's (2001) proposals of word knowledge— is grounded on the experience of many teachers, who not uncommonly observe that learners are able to understand a word but feel unable to reproduce it.

Receptive vocabulary, also known as passive vocabulary, has commonly been defined as "words learners need to recognise only" (López Campillo, 1995, p. 36). In other words, it refers to the language input learners understand when reading or listening. On the other hand, productive vocabulary, also known as active, can be defined as "the words learners need to be able to use and understand" (López Campillo, 1995, p. 36). To put it in simpler terms, it is the language produced by learners when attempting to communicate with others. Although at first sight, both terms seem easy to define and distinguish, there is a lack of consensus about them. For some authors, such as Laufer and Goldstein (2004) or Teichroew, cited in Nation (2001), receptive and productive knowledge are placed on a continuum, so, when learning a new word, as this word is presented several times to the student, receptive knowledge becomes active knowledge. Read (2000) argues that the problem lies in determining the particular moment in which receptive vocabulary turns into vocabulary knowledge. For other authors, such as Meara (1997) or Corson (1995), cited in Nation (2001), receptive and productive vocabulary differences may not lie on the natural progression of learning, but on the kind of connections between lexical items. From this point of view, if an item that is related to another productive item

becomes active, it becomes part of the productive vocabulary. On the other hand, receptively-known items have no incoming links from the lexicon, so they cannot be recalled unless activated by some outside stimulus.

To sum up, whether both dimensions may be seen as a continuum or not, what seems to be clear is that learners' receptive knowledge is larger than their productive one. Nevertheless, the relationship between both magnitudes remains unknown. The attempts to establish an exact relationship have been inconclusive. There are some researchers that note that both dimensions are closely related. To illustrate this, Melka (1997) estimated that 92% of receptive vocabulary is known productively, a figure that became even larger in Takala's study (1984). In contrast, other authors have found that they are not as closely related as it may seem. For example, Laufer (1998) concluded that at the 5K level, only 16 % of the receptive knowledge was known productively. Finally, other studies (Fan, 2000; Laufer & Paribakht, 1998) find that the percentage of receptive vocabulary known productively ranges from 50 to 75%.

There are different reasons that may explain why these figures differ to such a large extent. One reason may be the lack of consistency in determining what is considered receptive and productive knowledge. Another reason may be related to the number of instruments researchers have at their disposal for measuring vocabulary size. The use of different instruments can lead to different results. Therefore, it is essential to know the measurement instruments at our disposal and to establish clear criteria about how to select the most appropriate one.

As can be observed, there is a large number of ways to measure vocabulary knowledge. Depending on the choices researchers make, results can vary widely, resulting in inconsistent research findings. For this reason, it is of importance to select the instruments considering the aims and characteristics of the study and the sample. This discussion will be taken up again in Chapter Four as a criterion to select the most appropriate tools for the objectives of this PhD dissertation.

4. Vocabulary Development in the EFL classroom

This Chapter is devoted to the understanding of Lexical Competence in the broadest sense of the term. Thus, in section 1, the *Lexical Competence* construct was defined, and its main features were explained. In section 2, I presented the most recent theories about how lexical items are processed and stored in L2 learners' minds. After these theoretical clarifications, then, in section 3, I focussed on a specific aspect within lexical competence, vocabulary knowledge, and I presented several instruments to measure different dimensions of the construct. This last section deals, again, with lexical knowledge, but from a different perspective: the educational context. Thus, the aim of this section is to explore how lexis is developed in a specific learning setting: the EFL classroom. To start with, I present the main differences between natural contexts —i.e., those where language learning happens in a milieu where

that language is spoken—, and formal settings —where learning is restricted to specific moments and the target language is not the language of the context. Then, in section 4.1, I present an overview of the treatment vocabulary teaching has received in the main language teaching approaches. After that, in section 4.2, I focus on the features of school-age learners. Finally, in section 4.3, I explain how lexical development is fostered in CLIL settings.

Language learning, and, consequently, vocabulary learning, is a complex process in which a large variety of variables affect the final result. In this section, I am going to pay specific attention to one of these variables: the context, as different studies have shown its great impact on learning outcomes. Traditionally, two types of learning settings or contexts have been identified in SLA: natural and instructional settings. On the one hand, in natural settings, the target language (TL) is acquired in a real context, i.e., learners are surrounded by the target language in most of their interactions. According to Lightbown and Spada "natural acquisition contexts should be understood as those in which the learner is exposed to the language at work or in social interaction or, if the learner is a child, in a school situation where most of the other children are native speakers of the target language and where the instruction is directed toward native speakers rather than toward learners of the language" (2004, p. 123). On the other hand, in instructional settings, the target language is learned in a context in which this language is not widely used in everyday life. Given the widespread importance of the latter context, Lightbown and Spada (2014) distinguish different types of environments, i.e., structure-based instruction and communicative instruction (see Table 2.8 to compare the main features of both types of instructions). In structure-based instructional settings, the aim and objective of the learning is the language itself. Thus, the teacher provides learners with a wide range of activities and tasks to develop language learning. In contrast, in communicative environments, the emphasis is placed on the communication of meaning. In my view, as will be seen later on, CLIL, and, in general, content-based approaches, are a kind of extension of communicative instruction. In these approaches the communicative component is central, as learners are prompted to use the language to learn disciplinary subjects. For this reason, they share an important number of features with communicative environments. However, they add a new component: the development of the language of schooling (or academic language), which is specifically promoted through the teaching/learning of disciplinary subjects in English.

Table 2.8

Main features of the different language teaching settings

	Natural settings	Instructional/ formal settings	
		Structure-based instruction	Communicative instruction
Kind of input	Learners are exposed to a large variety of language structures.	Learners usually practices the language structures in isolation, from the simplest to the most complex structure.	Learners are exposed to modified input. Input is simplified to help understanding.
Amount of exposure	Learners are constantly exposed to the language and can find a lot of people who use that target language.	Learners are not constantly exposed to the TL and exposure is limited to some hours per week.	Usually, the only proficient speaker is the teacher, but they are exposed to other learners' output.
Discourse types	Learners uses the target language in different communicative events.	Learners are usually exposed to a limited range of language discourse types.	A greater variety of discourse types is presented to learners.
Pressure	There is no pressure placed on learners to use the language.	There is some pressure to communicate in the foreign language.	Little pressure for production and great emphasis on understanding.
Error treatment	Learner's errors are not usually corrected.	Errors are usually corrected.	There is limited amount of error correction.

Source: own elaboration based on Lightbown & Spada (2014).

This PhD dissertation focusses on lexical learning in instructional settings. Therefore, after this brief contextualization of what an instructional setting is and involves, the following sections deal with several aspects regarding lexical development in instructional settings. In the first section, a review of the attention vocabulary learning has received in the different language teaching methods is provided. Then, in section 3.2, I move on to lexical development in school-age learners, focussing on the particularities these learners present. Finally, section 3.3 deals with lexical development in formal settings, with special emphasis on the relationship between vocabulary knowledge and the CLIL approach.

4.1. Vocabulary in Language Teaching

Words are the basic tools people use to communicate and think (Bowen, Madsen, & Hilferty, 1985; Intaraprasert, 2004; McCarthy, 1990; Nation, 1990). That is to say, no one can communicate in any meaningful and effective way without managing a certain range of vocabulary. This view has been

postulated by prominent scholars in the SLA field such as McCarthy (1990, p. iix) when he pointed out that "no matter how well the student learns grammar, no matter how successfully the sounds of L2 are mastered, without words to express a wide range of meanings, communication in an L2 just cannot happen in any meaningful ways"; or, more recently, Schmitt, Cobb, Horst and Schmitt (2017, p. 213) when they said that the "knowledge of vocabulary is fundamental to all language use, and so must be learned in some manner in order for learners to become communicative in a new language". However, as has been explained before, the relevance given to vocabulary in L2 learning is relatively new. Traditionally, language teaching approaches and theories have ignored vocabulary. For a long time, there was no explicit mention to vocabulary in L2 syllabi or curricula and teaching training materials and books often omitted vocabulary teaching (Lightbown & Spada, 1999 or Mitchell, Myles & Marsden, 2004). Moreover, until fairly recently, as O'Dell (1997) notes, studies on vocabulary were scarce and the few existing focussed on methodological aspects of vocabulary instruction, rather than on vocabulary itself (Laufer, 1990).

This tendency is clearly related to a variation in the conception of language learning. In earlier approaches to language, such as the structuralist or the generative ones, language was conceived of as set of divisible units, among which, grammar was the most relevant. This view was directly spread to the teaching methods. For example, one of the most commonly used methods to teach foreign languages in the 19th century and first half of the 20th century was the *grammar-translation approach* that emerged as an application of the approach used to teach classical languages to the teaching of modern languages. Latin and Greek had been taught for several centuries throughout Europe, therefore, there were some methods used to facilitate the mastery of those languages: classical languages were taught by focussing on grammatical rules and declinations, memorizing vocabulary items and doing translation and written exercises (Brown, 2000). The need for developing competence in modern languages in the 18th and 19th centuries resulted in the direct application of these methods to the teaching of modern languages (Boers & Lindstromberg, 2008; Howatt & Widdowson, 2004; Marqués-Aguado & Solís-Becerra, 2013). However, given the nature and use of classical languages, this approach focussed exclusively on morphology and syntax, especially in written skills, while vocabulary learning[1] was relegated to the mere memorization of lists of items in order to help translations. Thus, vocabulary was simply considered a tool at the service of translational purposes and, as Boers and Lindstromberg note "little support is given to help learners retain new lexis for active usage" (2008, p. 2).

These two authors also point out that understanding oral communication as part of the L2 proficiency resulted in calling into question the convenience of using the *grammar-translation (G-T) approach*. Francis Gouin (1880) was one of the main precursors of the revision. Based on his own

---

[1] In this doctoral dissertation, the concepts of learning and acquisition are used interchangeably.

experience as a German learner, he summarised clearly the problem this method entailed: despite the fact that he mastered German grammar and learnt the dictionary by heart, when he attempted to put into practice his knowledge, he could not understand or produce orally a single utterance. Therefore, he made two proposals: (1) oral language had to be promoted and (2) modern languages had to be presented in context in order to facilitate L2 learners' understanding (Smith, 1893). In this context, Berlitz developed a new Language Teaching (LT) method: *The Direct Method* (DM; Stieglitz, 1955). This method advocated using, as much as possible, the FL in the classroom, as, in his view, it cannot be learnt by means of just translation exercises. In contrast to the G-T, in the Direct Method, oral skills were developed first. Demonstrations, pictures or objects were presented orally when teaching. Moreover, learners were not only exposed to the FL from the first moment, but they were also encouraged to use it, usually through the use of question-and-answer exchanges. However, as in the case of the G-T approach, lexis was relegated to a second placed and was only conceived of as a facilitator of communication.

The advent of *The Reading* and *Situational Language Teaching* methods brought with them the consideration of vocabulary learning as a central aspect in language learning (Moreno-Espinosa, 2003). These methods emerged, respectively, in the United States in the 1920s and in Great Britain in the 1930s (West, 1930), and were based on structuralist theories (Richards & Rodgers, 1986). Their main aim was to develop oral and reading skills, and L2 vocabulary, especially the most-frequent vocabulary identified by Thorndike's word-frequency list (1944), was considered a way to develop them. In this respect, West (1930, p. 514) noted that "the Primary thing in learning a language is the acquisition of a vocabulary, and practice in using it. The problem is what vocabulary; and none of these 'modern textbooks in common use in English schools' have attempted to solve the problem". At the same time, Palmer (1917, 1921) and Hornby (1950) in Great Britain, developed a framework of language teaching mainly focussed on practicing basic structures in meaningful situations. To choose the structures, they highlighted three main criteria: selection, gradation and representation of language structures (Richards & Rodgers, 1986). Thus, in these two approaches, as Zimmerman (1997, p. 10) states, "for the first time, vocabulary was considered one of the most important aspects of a second language learning and a priority was placed on developing a scientific and rational basis for selecting the vocabulary content of language courses".

Towards the mid-20th century, and in the context of the second World War (WWII), a new LT method appeared: the *audiolingual method* (Fries, 1945). The USA army, in need of quickly training their troops for basic communication in various languages, developed a method based on behaviourist (Pavlov, 1897, 1927; Skinner, 1948) and structuralist (Bloomfield, 1933) theories. Soon the method became widely used, not only in the army, entering L2 classrooms in the 1950s. The *audiolingual method* implied several changes in comparison to previous LT approaches: first, it promoted "fluency *with* accuracy" (Boers & Lindstromberg, 2008, p. 2); and in order to achieve this fluent expression, the

memorization and repetition of dialogues as well as drills were proposed. Moreover, for the first time, the development of the four main language skills was taken into account and a specific order of development was proposed: the first skill would be listening, followed by speaking, then reading and finally writing. Therefore, it emphasised the teaching of listening and speaking. However, despite the large implications it had for the teaching of foreign languages, in this method, the role of vocabulary was downgraded, as it was conceived of just as a resource to help translations, mainly promoted by means of repetition and memorization tasks (Zimmerman, 1997).

The 1970s saw the birth of one of the most influential approaches of the 20th century, the *communicative language teaching (CLT) approach* (Brumfit & Johnson, 1979; Wilkins, 1972). This approach mainly arose as a reaction to grammar-based approaches and it was usually considered an "outgrowth of the functionalist view of language" (Whong, 2011, p. 129). The CLT approach understood the learning process as a creative construction and advocated the real language use as the main way to learn it. Therefore, it promoted meaningful communication in class as a key methodology. Language skills would be integrated as a means to develop communicative competence (Richards & Schmidt, 2010). Finally, in contrast to previous approaches, it emphasised fluency over accuracy, as trial and error was considered part of the learning process. In general, it brought a great revolution to language teaching as it meant a re-consideration of goals and classroom activities. Its tenets have been incorporated into subsequent approaches, such as the task-based approach or content-based instruction. Regarding vocabulary, the CLT approach was, at first, mainly concerned with communication and with how to help students communicate effectively, and lexis was conceived as a facilitator of the final aim, i.e., communication. However, the approach evolved over the years, and brought with it some re-consideration of the role of vocabulary learning within communicative competence, ending with a progressive introduction of explicit vocabulary teaching in the classroom, especially after the flourishment of a new field of research in which lexis became the core area of study in mid-1980s (Boers & Lindstromberg, 2008; Moreno Espinosa, 2003).

As can be seen from this brief overview of the different LT approaches, despite their various origins and the great differences in their proposals, there is a fact that has remained stable: with the exception of the *Reading Method* and the *Situational Language Approach,* it seems that in none of the rest of approaches, vocabulary knowledge was thought to be a key aspect that could lead to L2 proficiency. In fact, Ketabi and Shahraki after a systematic review of the treatments vocabulary knowledge has received in the different language teaching approaches, state that "vocabulary teaching has not yet reached the level of consistency and systematicity that the teaching of other language skills enjoy such as grammar, although it has recently gained much attention in second language acquisition research" (2011, p. 729). On the contrary, it seems to have been developed in parallel with the language teaching approaches proposals, which, in some cases, have partly implemented some vocabulary teaching proposals on the basis of the linguistic demands.

This way of approaching vocabulary by the different LT approaches seems to explain its marginal role throughout history. However, apart from the lack of research and the omission of vocabulary development in the teaching practice, there are some authors (e.g., Milton, 2009) that point out other causes. In his view, teachers and learners' beliefs have also played a significant role in this trivialization of the relevance of vocabulary, as both practitioners and learners seemed to be reluctant to accept its importance in communicative achievement. Moreover, he also argues that the use of explicit vocabulary teaching methods has been traditionally seen as ineffective and this has not helped the implementation of new language teaching approaches with a focus on lexical development.

All in all, unsubstantiated beliefs that seem to have been deeply rooted in researchers, teachers and students' minds. And this, together with the little relevance given to lexical learning in the different language theories, hampered the development of a solid area of research on the role of the lexical competence in language learning. This situation continued until the 1980s, when a new conceptualization of the place of vocabulary within LT practice started to materialise. The advent of new LT theories, under Cognitivist paradigms, brought a new conception of vocabulary. In Dóczi's words (2017, p. 61), "in contrast to earlier theories, using vocabulary is [was] no longer viewed as filling the slots in a sentence with the help of various transformations. Instead, it is viewed as linguistic construction units (e.g., words, lexical phrases, and formulaic expressions) in their own right, which cannot be separated from their syntactic regularities". Studies showing the benefits of teaching vocabulary became more and more common (Laufer, 1986, 1990; Meara, 1980, 1996a, 1996b; Nation, 1974, 1975, 1983, 1990; Richards, 1976; Widdowson, 1978; Wilkins, 1972; Xue & Nation, 1984) and the main findings of these studies were transferred to the teaching practice. Practitioners and researchers started to consider students' lexical development an area as important as other linguistic elements, such as grammar or phonological components. For instance, Widdowson (1978, p. 115), in an attempt to describe the interaction between vocabulary and grammar in the classroom, stated that "lexis is where we need to start from, the syntax needs to be put to the service of words and not the other way round". Similarly, Wilkins stressed the relevance of lexis in comparison to grammar in the acquisition of a foreign language when saying that "without grammar, very little can be conveyed; without vocabulary, nothing can be conveyed" (1972, p. 111).

These new ideas and findings were translated as well into new LT materials. Consequently, since the 1990s, syllabi and curriculums started to include explicit information on the kind of vocabulary that should be taught, and teaching materials considered 'scientific' information about the selection of vocabulary that should be taken into classes. Similarly, this paradigmatic change was not reduced to the creation of materials in which vocabulary development was conceived, but it corresponded to the emergence of language teaching approaches where vocabulary was considered. In 1993, Lewis published *The Lexical Approach,* a language teaching approach that gave primacy to lexical knowledge. Considering principles from different disciplines, such as Corpus Linguistics, Discourse Analysis and

Psycholinguistics, Lewis developed an approach in which the grammar-vocabulary dichotomy was no longer valid, as language was conceived as a mix of different chunks made up of grammar and lexical elements. Therefore, in Lewis' words "language does not consist of grammar and words, and that much of our mental lexicon is stored as prefabricated multi-word chunks" (1997, p. 20). This new conception of language revolutionized teaching methods and principles. He proposed the use of an inductive and deductive methodology that promoted (1) the development of receptive skills, (2) activities based on L1-L2 comparisons, (3) the use of the dictionary as a resource, (4) the use of authentic texts in class, (5) the revision and recycling of vocabulary and (6) the organization of the lexical contents following mental lexicon organization. However, this approach has in fact not been put into practice very frequently, and has mostly remained as a theoretical proposal, as other approaches, such as the CLT were preferred as mainstream LT methods. One of its salient features is the consideration of lexis and grammar as interdependent elements, following, in this sense, the tenets of the Systemic Functional Linguistics (Halliday, 1990), that considered grammar and vocabulary as a single unit, introducing the term 'lexico-grammar' to make reference to this unity. Vocabulary, thus, needs to stop being either seen as isolated linguistic items or taught in isolation. It should be considered as a central element closely related to others, and, therefore, it should be taught in combination with the development of different language skills.

The reappraisal of L2 vocabulary teaching did not only give rise to language teaching approaches such as the *Lexical Approach*, but went a step further and even arrived at the level of national and supranational language teaching policies. This is the case, for instance, of the Common European Framework of Reference for Languages in the early 2000s. In my view, the identification of lexical competence as a distinct competence marks a turning point in the culmination of the shift of thought and demonstrates the extent to which lexis has reached a central role in the field of second language acquisition. Moreover, it implied further interest in the area from a more varied perspective. Whereas in the first stage the focus was mainly on more concrete aspects related to EFL instruction, thus proving the linguistic benefits of implicit lexical explanations, identifying the kind of lexical elements that should be taught, developing materials, or exploring the elements that made up word knowledge, in recent times, the focus has been broadened and enriched, thanks to the contribution of different linguistic schools.

This is the case, for example, of Cognitive Linguistics (CL), a linguistic paradigm, that places emphasis on the interaction between language and cognition (Richards & Schmitd, 2010). CL has meant a review of the importance of the lexical component in language proficiency and has helped to better understand and elaborate on the concept of lexical competence (Boers & Lindstromberg, 2006, 2008; Piquer-Píriz & Boers, 2019). For cognitive linguists, there is not a strict distinction between grammar and lexis, so they cannot be treated as different or autonomous language aspects, but they should be treated as a continuum (Ibarretxe-Antuñano, 2004; Langacker, 1987) and, therefore, "they correspond

to very specific conceptualisation, i.e., the lexicon for specific entities or relations, the grammar for more abstract conceptualisations" (Ibarretxe-Antuñano, 2004, p. 11). Thus, lexical knowledge consists of "memorised symbolic items" (Boers & Lindstromberg, 2008, p. 7), that range in size between "bound morphemes […], words […] and phrasal expressions […] and complete phrases […]" (Boers & Lindstromberg, 2008, p. 7). In other words, vocabulary encompasses not only words, as commonly theorised by generativists or structuralists, but also phrases. Definitely, CL, with its vision of 'motivated' vocabulary not consisting solely of single words but interrelated items, has enriched and widened the lexical competence construct, as it shows that the importance of vocabulary learning lies, not only in learning countless series of items, but also on understanding how these items are related to each other (Piquer-Píriz & Boers, 2019).

In short, vocabulary learning has experienced a great boom in the last forty years, reaching the same importance as other linguistic elements, such as grammar, that traditionally have received more attention. However, from my point of view, there are, at least, two main issues that still may hinder a better L2 learners' lexical development. First, although a great deal of attention has been given to vocabulary studies in research, it has not been translated into a greater emphasis on vocabulary teaching in the classroom. For example, it is still quite frequent to find that EFL textbooks present vocabulary items in lists, rather than introducing them in a more contextualised and meaningful way. Second, and related to what I posited in section 1, there are still some inconsistencies in terminology that may hamper teachers' understanding of vocabulary teaching. In fact, Ketabi & Shahraki, after a systematic review of the different language teaching approaches, state that "vocabulary teaching has not yet reached the level of consistency and systematicity that the teaching of other language skills enjoy such as grammar, although it has recently gained much attention in second language acquisition research" (2011, p. 729). Examples of this inconsistency are, for instance, the boundaries of the lexical competence in relation to other linguistic competences. As can be seen in table 2.1, the CEFRL breaks the linguistic competences into more specific sub-competences, differentiating between lexical and semantic competences. Lexical competence is related in this document to knowledge and use of L2 vocabulary, and, therefore, requires knowledge of the meaning of the words, whereas semantic competence is related to "learner's awareness and control of the organisation of meaning" and encompasses Lexical Semantics, that deals with questions of word meaning, Grammatical Semantics and Pragmatic Semantics. Both competences are so closely related to knowledge of word meaning that it makes it difficult to establish where the limit between one and the other is. I concluded Section 1 arguing that a clearer clarification in lexical competence terminology is needed. Now, in light of what has been shown in this section, I will add a new element to this proposal. After clarifying the terminology, this consensus should be also transferred to language teaching practice. Language teaching research is expected to aim at improving teaching practice, and, in this sense, research and implementation in the classroom should go hand in hand, having a close and real interaction.

4.2. Lexical development in school-age L2 learners

Learning a foreign language is a process that differs greatly from L1 acquisition, particularly, in instructional contexts. In comparison to the first language learning experience, L2 learning results in a more challenging experience, in which learning becomes an explicit process and learners are asked to learn an L2 in a more decontextualised way. Muñoz (2010) establishes differences between natural settings, and instructed settings, in, at least, five aspects: (1) L2 instruction is limited to a number of specific sessions per week, which, in turn, (2) derives in a limited exposure to the target language. Moreover, (3) the FL is not the normal language of communication in this society, and, therefore, (4) the FL is not spoken outside the classrooms. Finally, (5) the instructor's fluency may be limited. These features of instructed settings mean that learners cannot make use of the strategies they used when acquiring their mother tongue and results in a need to develop new skills to foster language learning.

This is especially difficult for young and very young learners (YLs and VYLs, respectively). In recent years, these learners are found more and more frequently in the FL classrooms, but they present specific features that differentiate them from adult L2 learners (Lightbown & Spada, 1999) and which are not fully taken into account in SLA research. First, young children and teenagers are still developing as individuals, therefore, there are some relevant abilities for L2 learning that they do not have fully developed. For example, depending on the age, cognitively, they are still developing their metalinguistic awareness, i.e., they may not be able to understand grammatical rules and explanations about language. Secondly, they have a shorter span attention in comparison with adult learners, so the teaching practice should be adapted. Finally, especially in the case of very young learners, school-age L2 learners are still developing their L1, and this may cause interference.

These differences between adults and school-age L2 learners may affect the learning process and, thus, should be reflected in SLA research. In fact, learners' age has been identified as "one of the crucial issues in the area of second language (L2) acquisition" (Muñoz, 2010, p. 39) and the issue of age of onset has been explored in relation to a number of different language skills, such as receptive skills (Cadierno et al., 2020; Jaeskel, Schurig, Florian & Ritter, 2017), oral performance (Muñoz, 2014), pronunciation (Flege & MacKay, 2011), receptive grammar (Cadierno et al., 2020), morphosyntax and fluency (see Muñoz, 2010, for an overview), general English proficiency (Cenoz, 2003) and other language learning related factors, such as motivation (Graham, Courtney, Tonkyn & Marinis, 2016; Mihaljević Djigunović & Lopriore, 2011, or Pfenninger & Singleton, 2016) or levels of language learning anxiety (Johnstone, 2009).

Focussing exclusively on lexical knowledge, Miralpeix (2007) studied the impact of age of onset and amount of exposure on lexical production. To do so, three groups of secondary-school learners who differed in age of onset and/or amount of exposure to English were asked to write a composition and the outcomes were compared. In light of her results, she concluded that an early start in formal contexts

does not result in richer vocabulary production. Similarly, Agustín-Llach and Jiménez-Catalán (2018) compared the vocabulary production of children (aged 11) and adults EFL learners who shared the same amount of exposure to the FL. They concluded that adults obtained better and richer results than children, despite having a similar amount of exposure and language level (A2). In one of the latest studies on this issue, Cadierno et al. (2020) examined the learning rate and development of receptive vocabulary, receptive grammar, and phonetic discrimination skills of two groups of early learners who were introduced to English at different ages: the first group started to attend EFL lessons in 1st grade, whereas the second group started EFL lessons when they were 3rd graders. They came to two strong conclusions: first, the latter group outperformed the former in all the tests. Second, later starters showed a more advanced learning rate when it comes to receptive grammar, whereas in the case of phonetic discrimination, although late starters presented a better performance in the tests along the three years of study, differences narrowed given the larger rate of development of early starters.

However, despite the evidence that the age of onset affects language learning and the substantial differences in young and adult learners' processing capacity, the bulk of research in vocabulary development has been conducted based on adults' understanding and performance. There are different reasons that can justify the selection of adults as the main population of study. First of all, it is easier to access adult L2 learners. Doing research with under-18 is often seen as problematic, given that it requires parents, schools and educational authorities' approvals to access the participants, while in the case of adults, only their explicit consent is needed. Moreover, given that children and young teenagers are still in the process of developing their capacities, adults are thought to be more able to reflect on their own learning process and to better express themselves. Finally, in the case of young learners, their developing linguistic abilities may prevent them from transmitting, with accuracy, what they mean even in their L1.

For these reasons, and, despite a reality in which the number of young, school-age L2 learners is increasing, studies on very young learners and young children and teenagers' vocabulary development are still scarce (Cadierno et al., 2020; Miralpeix, 2006, 2007).

There are two important aspects that, in my view, need further exploration. Firstly, a great deal of L2 objectives are based on what L1 speakers are able to do. The 'native-likeness' —quality of using a language in a similar way to a native speaker— is often considered as the final aim of language teaching, so L1 adults' performance is compared to that of L2 learners, and L1 corpora analyses are performed to obtain information about the frequency of linguistic features. Data is used to establish how L2 learners should be able to perform at the different language learning stages. In this respect, more often than not, the learning objectives for L2 young and very young learners are exactly the same as adults regarding vocabulary knowledge. However, children and young teenagers differ greatly from adults and whether these differences in contexts may have repercussions on the frequency of occurrence of lexical items, remains an unexplored issue up to now. Therefore, a first step in setting suitable and realistic

objectives for L2 children and young teenagers would be the compilation and analysis of these learners' production (L2 learners corpora). There are various initiatives in this direction, and different corpora have been compiled making use exclusively of L2 school-age learners' productions (see Table 2.9 below). However, to the best of my knowledge, despite the potential of this tool, corpora are only used to explore linguistic features of L2 learners or to be compared to L1 corpora, rather than to understand what school-age learners are able to do in accordance with their age and maturity and set feasible objectives that help school-age learners to develop L2 in the specific context of a formal setting.

Table 2.9

Corpus compiling school-age English L2 learners' productions

Corpus	L1	Medium	Learners' age	Size in words
The TELEC Secondary Learners Corpus	Chinese	Written and spoken	Secondary-school learners	c. 2 million
CORYL (Corpus of Young Learner Language)	Mainly Finnish	Written	Secondary-school learners	191,568 tokens
FUSE (Finnish Upper Secondary School Corpus of Spoken English	Finnish	Spoken	Secondary-school learners	n/a
The Corpus of Young Learner Interlanguage (CYLIL)	Dutch, French, Greek & Italian	Spoken	Primary and secondary-school learners	c. 500,000
The English of Malaysian School Students corpus (EMAS)	Malay	Written	Secondary-school learners	c. 500,000
The EVA Corpus of Norwegian School English	Norwegian	Spoken	Secondary-school learners	c. 35,000
The Young Learner Corpus of English (YOLECORE)	Greek	Spoken	Primary-school learners	1,5 million types

Corpora selected from: https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html. Own elaboration.

Secondly, and very closely connected to the previous aspect, in the case of vocabulary measuring tools, there is a complete lack of materials adapted to VYLs and YLs needs. Most of the instruments are based on frequency lists based on L1 corpora analyses, and, therefore, they establish the vocabulary

knowledge of the test-taker in relation to an adult's performance. Moreover, they are piloted exclusively with adults, so the possible biases that may appear when using them with school-age learners are usually disregarded. In addition, vocabulary knowledge tests do not consider YLs and VYLs' maturational features, such as their L2 literacy level or their attention span, and how this may affect the results. Up to date, the only vocabulary test aimed at school-age learners is the Peabody test, but it presents two shortcomings that reveal the need to develop new vocabulary receptive and productive knowledge tests for young children and teenagers. First of all, the Peabody test is aimed at pre-literacy levels, so it may be used with VYLs, but it is not so suitable in the case of young children and teenagers. Secondly, and more relevant from a scientific point of view, it does not establish what levels of proficiency it measures or creates an equivalence to language proficiency, so it avoids making measurable comparisons or relating vocabulary knowledge and skills performance.

As the age of onset in L2 classes has been lowered, this area of research has progressively grown in interest. At first, the few researchers exploring vocabulary acquisition in school-age children have used tests developed with adult data. However, the instruments available have not been conceived for use with children or young teenagers, and to what extent the age of the test-takers has a significant effect on the results is an issue yet to be explored. It is a well-known fact that reliable measuring tools are essential in SLA research. Therefore, the development of new vocabulary tests, adapted for school-age learners, is a real need and demand.

This piece of research attempts to take a step forward in the transition from research on YLs with adult-based criteria to research on YLs with a YLs-friendly approach. As will be seen in the following chapter, this dissertation provides an instrument adapted to young teenagers to gauge their ability to reflect on their own learning process. Unlike other vocabulary learning strategies questionnaires, this adapted proposal presents a series of specific properties that make it especially suitable for young teenagers: first, a short number of items o were selected taking into account their shorter attention-span teenagers are expected to have. Second, the instrument was presented in the learners' L1 in order to facilitate their understanding and to avoid any misinterpretation. Finally, it has been validated exclusively with young teenagers, to prevent comprehension problems.

4.3. L2 vocabulary knowledge in formal settings

One of the main concerns regarding vocabulary teaching is often related to the selection of the most appropriate and useful vocabulary for L2 learners. In the definition of lexical competence in section 1, it was already stated that an adult educated native speaker of a language is expected to master between 16,000 and 20,000 words. Given the impossibility of teaching the complete lexical repertoire of a language, Corpus Linguistics has helped to identify the most common words. With such analysis, vocabulary was broken into bands according to the frequency of occurrence in the corpora examined. The creation of frequency lists was foundational to other related topics, such as the analysis of the

vocabulary coverage of texts. As mentioned in the section 1.2.1 in these analyses, researchers concluded that highly frequent words cover the majority of the running words in spoken and written texts (Schmitt, 2010; Webb & Nation, 2017). In light of this finding, it would seem that words belonging to the first bands should be taught in first place, as they are the most common items in English oral and written texts and its knowledge will result in a better command of the language.

Consequently, as the communicative competence of the learners improves, so should their mastery of less frequent words. For example, as cited in Milton (2010), Meara and Milton (2003) associated Cambridge Exams scores with the vocabulary size measured with the Xlex and found that larger vocabulary sizes were used in the most advanced levels. Similarly, Milton (2010) attempted to tie vocabulary knowledge with the CEFRL levels. In his view, in order to progress beyond an elementary level of competence, a knowledge of about 3,000 words seemed to be needed. When examining the advanced levels on the CEFRL, they were associated with scores of 4,000 words or better. To achieve such advanced levels of competence, authors such as Laufer and Ravenhorst-Kalovski (2010) insist on the importance of setting adequate goals on the basis of the level of comprehension of the learners. This vision is in line with Krashen's input hypothesis (1985), according to which learners need to be exposed to an input slightly more advanced than their comprehension level (i+1), or with the Vygotskian theory of Zone of Proximal Development (ZPD, 1934), that puts forward that for the learning process to occur successfully, contents should be slightly more difficult than learners' knowledge.

A second important concern is related to how this vocabulary should be presented to learners. Traditionally, as discussed in the first section of this Chapter, vocabulary occupied a secondary role in the language classroom and, until quite recently, it was not even considered when planning lessons. With the implementation of new language teaching approaches in which vocabulary was part of the curriculum, there was a need to specify how this vocabulary should be brought to class. Since it started to be studied, vocabulary learning has been closely related to the input learners are exposed to. Exposure is central to lexical competence development, and soon, vocabulary learning was related to incidental learning when practising reading skills, as books are one of the main sources of vocabulary available to learners (Webb, 2008). However, under the consideration that reading was not enough for learners to achieve a full command of vocabulary and that some learners could fail when learning vocabulary by reading, new methodologies started to be explored. It is in that context that the concepts of implicit and explicit vocabulary learning began to be used. Implicit vocabulary learning refers to the process of learning vocabulary in which no special focus is placed on it, for example, when watching a film or reading a text. On the contrary, explicit vocabulary learning is defined as "explicit learning through the focused study of words" (Schmitt, 2000, p. 116). With regard to this aspect, Nation emphasises that "second language learners should not rely solely on incidental vocabulary learning from context" (Nation, 2001, p. 238) but that "direct vocabulary learning and incidental learning are complementary activities" (2001, p. 238). Similarly, Schmitt (2000, p. 116) indicates that "for second language learners

at least, both explicit and incidental learning are necessary, and should be seen as complementary". Therefore, equal opportunities for both types of learning are needed. In addition, in the selection of the kind of approach that should be used to present vocabulary items to learners, there is a second factor that should be borne in mind: learners' needs. Learners at different levels may need different approaches. As Schmitt (2000, p. 121) states:

> not only is vocabulary acquisition incremental, but it is incremental in a variety of ways. First, lexical knowledge is made up of different kinds of word knowledge and not all can be learned simultaneously. […] this indicates that word learning is a complicated but gradual process.

This gradual introduction of vocabulary items may result in different learning needs. Depending on the stage of the learning process, learners and teachers should be aware of them in order to create the most optimal environments to facilitate language learning and, concretely, vocabulary learning. This links directly to the third aspect of discussion in this section: the tasks and activities to bring to the classroom. Research, not just focussed on vocabulary, (Nunan, 2004; Willis, 2004) often identifies three main language learning activities in the language classroom. These three types of activities, as well as their definition are shown in Table 2.10.

Table 2.10

Language learning activities

Activity	Definition
Exercises	They provide the learners with controlled practice of decontextualised language.
Acting activities	In these activities, language looks realistic, but the learners are not asked to create the full message, but messages are often provided, and they have to adapt them. The major focus is on practicing particular structures.
Communication tasks	"Any structural language learning endeavour which has a particular objective, appropriate content, a specified working procedure, and a range of outcomes for those who understand the task" (Breen, 1987, p. 23).

In general, tasks stand out as the most relevant communicative language activities. However, the potential benefits of the other two types of actions should not be undervalued, especially in vocabulary learning. For example, focussing now on vocabulary and taking up Nation's proposal of word knowledge again (see p. 46 for further information), mechanical exercises could be a good option to make learners familiar with the spoken and written form of the word, as well as with its form and meaning. Once this knowledge is achieved, 'acting activities' could be introduced in order to settle this knowledge and expand it to other dimensions, such as associations, collocations or grammatical functions among others. Finally, the introduction of communication tasks, which are thought to be the

most demanding learning activities, could help to enhance the knowledge of the term. In short, it is not only the use of a particular activity, but the combination and synergies between them that really facilitates language learning.

In sum, language learning activities are at the 'heart' of the language learning process and contribute to vocabulary gain. The selection of appropriate vocabulary learning activities will definitely result in a larger vocabulary growth and language development. In this respect, the CEFRL also gives great importance to the role that activities have in language teaching. This document, and especially its companion volume (2017), not only establish and describe different language proficiency levels together with indicators for each level, but also provide meaningful and detailed information about the language learning process:

> The acquisition of proficiency is in fact seen as a circular process: by performing activities,
> the user/learner develops competences and acquires strategies. This approach embraces a
> view of competence as only existing when enacted in language use, reflecting both (a) the
> broader view of competence as action from applied psychology, particularly in relation to
> the world of work and professional training and (b) the view taken nowadays in the
> sociocultural approach to learning

CEFRL Companion Volume, 2017, p. 33

This statement can be applied to any field, as there is a broad body of literature (Nunan, 1989; van den Branden, 2006) showing the positive impact that the use of meaningful tasks on learning and on the development of competences. In some way or another, the different issues discussed above are closely related to the kind of teaching approach introduced in the EFL classroom. The different LT approaches differ in aspects such as the type of activities suggested, the way vocabulary is presented (either in an explicit or implicit way) or even, in some cases, such as specific programmes, the approach followed determines the vocabulary presented to the L2 learners. This PhD dissertation focusses on the lexical development of CLIL learners. Therefore, a more thorough explanation of the implications of the CLIL approach for lexical development is in place.

4.3.1. Lexical development in CLIL

Since the setting up of the European Union in 1957, and, considering the linguistic needs of some specific multilingual regions and nations, a new teaching approach began to emerge in different European countries, based on Canadian and American bilingual and immersion programmes. Content subjects began to be taught through foreign/second and minority or regional languages. Such was the magnitude of the impact and acceptance of these innovative forms of education that, at one point, the European Commission started to consider this new trend. The term CLIL was first used in 1994 to describe good practices in different schools where the teaching-learning process took place in an

additional language, considering 'additional language' any language different to the first language (Cenoz et al., 2013, p. 3; Coyle et al., 2010, p. 3). The most extended definition of CLIL describes it as

> a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language. That is, in the teaching and learning process, there is a focus not only on content, and not only on language. Each is interwoven, even if the emphasis is greater on one or the other at a given time. CLIL is not a new form of language education. It is not a new form of subject education. It is an innovative fusion of both.

> Coyle, Hood, & Marsh, 2010, p. 1

This definition highlights the major features of CLIL. Concerning language, first, it uses the expression "additional language", avoiding, thus, referring to specific languages (Cenoz et al. 2013, p. 3). Second, it establishes that CLIL is, at the same time, both a language and an educational approach. In that sense, it introduces a novel idea in language teaching practice, as the role of language varies. Whereas in traditional language teaching, language is seen as the content and the goal, in this new approach, language is conceived as a tool and aim. In this respect, the real novelty lies in the inclusion of the L2 development as an aim *per se* as well as the vehicle for communication in content subjects. This integration of language learning as part of the learning process results in an implicit learning of that language through its active use. In that way, the L2 reaches a different status, as now it is conceived also as a transmitter of ideas and contents rather than just as a goal of learning. Thus, learners focus on meaning rather than on form when learning the foreign language.

But not only has the role of language varied, there is also a different kind of language needed. Whereas in other language approaches the objective is the use of everyday language, in CLIL, language is used in an academic context, although the former is also present (Llinares et al, 2012). Table 2.11 summarises the language present in both approaches.

Table 2.11

A comparison between CLIL and mainstream EFL subject language.

CLIL language	EFL subject language
- Everyday language	- Everyday language
- Academic language	
- Technical language	
FOCUS ON MEANING	**FOCUS ON FORM**

This dichotomy had been acknowledged in the literature long before the appearance of CLIL. Cummins (1979) showed the existence of two main types of languages, BICS and CALP. The first acronym stands

for Basic Interpersonal Communicative Skills (BICS) and refers to "conversational fluency in a language" (Cummins, 2000, p. 487), whereas the second stands for Cognitive Academic Language Proficiency (CALP), and refers to "students' ability to understand and express, in both oral and written modes, concepts and ideas that are relevant to success in school" (Cummins, 2008, p. 487). Similarly, Bernstein (1999), when exploring the language and registers found in school contexts, reached a similar conclusion and made the distinction between 'horizontal discourse', made up of everyday life knowledge, and 'vertical discourse', closely related to the language of schooling, which is not commonly understandable, and it is not part of the experience of children experience until they start school. This distinction between both types of languages is of prime importance in CLIL. Whereas mainstream EFL is devoted to the development of everyday language, in the case of CLIL, both kinds of language should appear in real practice as the aim is to transmit academic contents. Nonetheless, although reaching a proper academic language level is crucial in CLIL, it should be achieved in a progressive way: teachers should start from what is familiar to the learners, 'horizontal discourse', or BICS, and through the use of scaffolding techniques reach more challenging language skills, i.e., 'vertical discourse' or CALP.

Due to this shift in the conception of language, and in order to facilitate teachers' understanding of the new role of languages in a CLIL context, Coyle (2007) presented an analytical framework, known as the language triptych, which included the language *of* learning, *for* learning and *through* learning. The 'language *of* learning' refers to the language of the content subject, the key terminology of the discipline but also the language functions that are necessary to express the required concepts in a specific discipline. The 'language *for* learning' encompasses that kind of language that helps learners understand and communicate content, including the language required to discuss, analyse and synthesise or to apply concepts. As can be seen, it is not applicable to one concrete subject, as happens with the 'language *of* learning', but is common to all kinds of academic subjects. Finally, the 'language *through* learning' is the new language that emerges from the learning situation and by which learners express their understanding and create new meaning. It is the integration of those three components of language that makes CLIL successful.
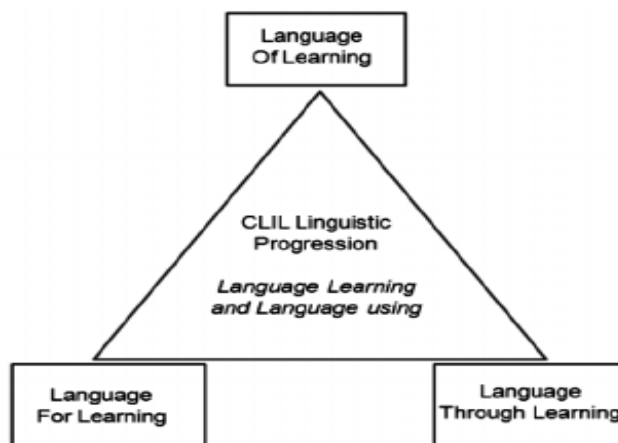
Figure 2.12. The language triptych. Source: Coyle, 2007.

This shift in the conception of language in CLIL seems to have had a positive impact on language learning. In the case of L2 development, a considerable amount of research has been carried out. Studies on specific language components (Agustín-Llach, 2009; Agustín-Llach & Canga Alonso, 2016; Canga Alonso, 2013a, 2013b, 2015; Canga Alonso & Arribas García, 2014; Jiménez Catalán & Ruiz de Zarobe, 2009; Ruiz de Zarobe, 2008) have reported better results in CLIL vs traditional EFL in both primary and secondary educational levels. Aspects such as productive and receptive vocabulary size, pronunciation, grammar, language transfer, or fluency are some of the ones analysed. For example, Agustín Llach (2009) explored the differences between primary-school CLIL and EFL learners as regards L1 language transfer in writing tasks, concluding that EFL learners had more transfer episodes than CLIL learners. Ruiz de Zarobe (2008) studied the differences in speech production, focussing on different items such as pronunciation, vocabulary, grammar, fluency and content, and CLIL students outstrip non-CLIL students in all the categories. Machado Osado (2015) explored CLIL learners' profile regarding three aspects: productive vocabulary size, grammar and aptitude, concluding that CLIL learners' level of aptitude was high, whereas the grammar and productive vocabulary size results were considered average.

The implications of CLIL for vocabulary learning, have been, by far, the most fruitful area of analysis (Agustín-Llach & Canga Alonso, 2016; Canga Alonso, 2013a, 2013b, 2015a; Jiménez Catalán & Agustín-Llach, 2017; Jiménez Catalán & Ruiz de Zarobe, 2009). Vocabulary research has become widespread, and has targeted different vocabulary dimensions, such as receptive and productive vocabulary size or academic vocabulary attainment (Lorenzo & Rodríguez, 2014). Moreover, vocabulary size of learners in Primary, Secondary and Vocational schools have been explored through cross-sectional and longitudinal studies.

Most studies have found positive evidence that CLIL learners outstrip EFL learners in both, primary and secondary levels. The studies carried out in Spain have been numerous and diverse. In the case of receptive general vocabulary size, studies have been carried out with two different aims in mind:

(1) to know the general vocabulary size of CLIL and mainstream EFL learners separately and (2) to examine the vocabulary learning benefits resulting from the use of different language approaches, mainly, CLIL and mainstream EFL approaches. Starting with the former, Jiménez Catalán and Ruiz de Zarobe (2009), as well as Canga Alonso (2013a), situate receptive vocabulary size of CLIL primary school learners within the first band of vocabulary after 1,000 hours of instruction. Fernández Fontecha and Terrazas Gallego (2009) point to a larger receptive vocabulary size after a similar exposure to the L2, concluding that CLIL learners knew, on average, 1,215 words out of the 2K most frequent ones. As for the latter, comparing CLIL and mainstream EFL learners has been a more common procedure, with, mainly, two main types of studies. On the one hand, some research has compared learners with the same age, but with different exposure to English input. These studies have found that CLIL learners outperformed mainstream EFL learners. For example, Agustín Llach (2012) explored 4th grade CLIL and mainstream EFL learners' vocabulary size and found a larger receptive vocabulary size in CLIL learners in comparison to their EFL counterparts. In turn, Arribas (2016) explored CLIL and EFL 10th grade learners regarding a series of aspects in which receptive vocabulary size was included finding that: first, CLIL learners presented a knowledge of 1,300 words out of the 2K most frequent ones, being in line with Fernández Fontecha and Terrazas Gallego (2012); and, secondly, he found differences between CLIL and EFL groups in favour of the CLIL learners. However, both studies could not strictly demonstrate that such difference was related to the approach followed, as it could also have been produced by a larger exposure to English CLIL learners received.

Other studies have opted for exploring differences between CLIL and regular EFL learners by controlling the amount of exposure, even if this meant comparing students of different ages. For example, Canga Alonso (2015) examined the receptive vocabulary knowledge of two groups of learners with the same amount of exposure in different grades (6th grade CLIL learners vs 10th grade mainstream EFL learners) without finding significant differences between both groups. Similarly, Fernández Fontecha (2014) compared the receptive vocabulary size of CLIL and EFL learners exposed to 734 hours of English at school. As in the study carried out by Canga Alonso, she compared learners with different ages: EFL learners were in the 1st grade of Secondary Education, whereas CLIL learners were in 4th grade of Primary Education. She found that mainstream EFL learners presented a larger receptive vocabulary size in comparison to their CLIL counterparts. In one of the latest studies on this issue, Agustín-Llach and Jiménez-Catalán (2018) compared the vocabulary production of CLIL children (aged 11) and adult EFL learners who shared the same amount of exposure to the FL. They concluded that adults obtained better and richer results than children, despite having a similar amount of exposure and language level (A2).

This kind of comparison has been carried out to explore not only receptive (Agustín-Llach & Canga Alonso, 2016) and productive (Alejo & Piquer-Píriz, 2016a; Jiménez Catalán & Agustín-Llach, 2017) vocabulary knowledge, but other language dimensions, such as oral comprehension and

production (Pérez-Cañado & Lancaster, 2017). With respect to productive knowledge, although there are some studies exploring production of CLIL learners (Pérez Cañado, 2018; Pérez-Cañado & Lancaster, 2017), studies comparing CLIL and EFL' productive knowledge of vocabulary are scarce. There are three studies (Canga Alonso & Arribas García, 2014; Merikivi & Pietilä, 2014; Moreno Espinosa, 2010) that have explored productive vocabulary knowledge of CLIL learners with similar features, finding conflicting results: a productive knowledge of 645 words in the case of Moreno Espinosa (2010) and Merikivi and Pietilä (2014) vs a recalling of 840 words in the case of Canga Alonso and Arribas García (2014). Therefore, these results support the hypothesis previously explained that receptive vocabulary size seems to be larger than its productive version. Jiménez Catalán and Agustín-Llach (2017) examined the lexical availability of 70 CLIL and regular EFL learners in their 8th and 10th grade respectively, concluding that the CLIL group were able to retrieve a higher number of words.

All in all, in those studies in which CLIL and mainstream EFL participants' performance was compared, authors have indicated that these comparisons may be problematic, as the practical implementation of different teaching approaches implies the use of different pedagogical techniques and different time of exposure to the foreign language. In fact, both of the approaches used to explore CLIL and regular EFL learners' differences discussed above present methodological drawbacks: in the case of learners with the same age, the problem is related to the isolation of the effect of different amounts of exposure to the L2, whereas the greatest concern in comparing CLIL and EFL learners with the same exposure to English is the maturational constraints derived from comparing learners with different ages. To avoid the methodological problem of comparing learners with different ages or different amount of input to English, there is a current trend towards exploring vocabulary growth rather than vocabulary size itself (Alejo & Piquer-Píriz 2016a; Pérez-Cañado & Lancaster, 2017). The use of longitudinal studies may help to clarify some doubts about the impact of other variables, such as the age and maturational level of the participants or the number of hours of exposure to the L2, on the outcomes of cross-sectional studies.

Table 2.12 below summarises the main findings in studies about CLIL and EFL learners' general vocabulary knowledge. It presents the kind of approach the learners followed, an estimation of the amount of exposure (AoE) to English test-takers had by the time they placed the tests, the grade in which they were, the tests used and the estimation of the number of words.

Table 2.12

A summary of studies conducted in Spain estimating L2 learners' vocabulary knowledge

Study	Tuition	IAoE (in hours)	Year	Vocabulary measured	Test used	Estimation of no. of words
López-Mezquita (2005)	EFL	1,049		Receptive		941
Jiménez Catalán & Ruiz de Zarobe (2009)	CLIL	960	6th	Receptive	VLT	800
Agustín Llach (2012)	CLIL	734	4th	Receptive	VLT	470
	EFL	419	4th		VLT	595
Fernández Fontecha & Terrazas Gallego (2012)	EFL	944	9th	Receptive	VLT	1215
Canga Alonso (2013a)	CLIL	839	5th	Receptive	VLT	696
Canga Alonso (2013b)	EFL	524	5th	Receptive	VLT	499
Fernández Fontecha (2014)	CLIL	734	4th	Receptive	VLT	471.26
	EFL	734	7th			779.54
Canga Alonso (2015a)	EFL	1,049	10th	Receptive	VLT	936
	CLIL	949	6th		VLT	903
Canga Alonso (2015b)	EFL	499	5th	Receptive	VLT	524
	CLIL	696	5th		VLT	839
Arribas (2016)	CLIL	Not provided	10th	Receptive	VLT	1330
	EFL				VLT	1200
Castellano Risco (2018)	CLIL	2,010	10th	Receptive	Yes/no	1,663
	EFL	1,200				1,301
Moreno Espinosa (2010)	-	-		Productive		645
Canga Alonso & Arribas García (2014)	CLIL	1,109		Productive	PVLT	813
	EFL	1,049			PVLT	640

At this point, it is important to remember the potential benefit of language learning in CLIL: the development of, not only general, but also academic language. As shown in this section, there are several studies examining CLIL learners' receptive knowledge of vocabulary and those examining productive knowledge are becoming more common as well. However, little research has been carried out on a central aspect of study in CLIL: academic language. CLIL is an approach in which L2 language learning occurs while learning content subjects. Therefore, it would be expected that these learners develop a series of structures that help them to transmit the subject contents, the so-called language *for* learning (Coyle, 2007). In this sense, there have been some attempts to examine the development of academic vocabulary and structures in CLIL. For instance, Lorenzo & Rodríguez (2014) compiled a corpus of 244 historical narratives written by CLIL secondary-school learners in four different years and examined the evolution of lexical and syntactic complexity and cohesion of the texts. Focussing on the lexical analysis they concluded that "changes are continuous but unstable, with higher peaks reaching significant levels in the uppermost course" (p. 70). However, given the importance of developing an appropriate level of academic language in CLIL, more studies on academic language should be carried out. In this sense, this PhD dissertation seeks to explore lexical knowledge development in CLIL settings from a broader perspective: not only does it focus on general receptive vocabulary knowledge as most of the previous studies, but it also includes an analysis of productive knowledge and expands the scope of analysis to academic vocabulary.

5. Conclusion

In this chapter, an introduction to vocabulary acquisition research has been presented and some of the main aspects surrounding vocabulary acquisition, such as lexical competence, the dimensions of word knowledge, the measurement of vocabulary knowledge or the acquisition of vocabulary in instructed settings have been discussed.

The chapter starts with a definition and conceptualization of the theoretical construct of *lexical competence*, providing a review of the main studies on this issue. *Lexical competence* has been deeply explored throughout the previous decades. Prominent scholars have examined the term using two kinds of approaches: linguistic and psycholinguistic. Despite the large number of proposals, there are some concepts that are common to most of the studies in the first group, such as the existence of two main dimensions of vocabulary knowledge: receptive and productive. As for the second group, another dichotomy —size and breadth of vocabulary— was found in most studies. Then, in section 2, an overview of the main models concerning L2 and lexical processing and storing has been presented. After this conceptualization of lexical competence, the last two sections of the Chapter have focussed exclusively on vocabulary knowledge, the first of the two main elements of the lexical competence, whereas learners' decisions and actions will be taken up again in the following Chapter. In section 3

the main instruments for measuring vocabulary knowledge have been presented. Given the complexity of *vocabulary knowledge*, and, under the evidence that the different dimensions correlate, the scope of analysis has been delimited to the analysis of vocabulary size and some of the main instruments for it have been presented, making use of the distinction between productive and receptive knowledge to scaffold the contents. The Chapter concludes with a review of lexical development in instructional settings, with a focus on issues such as the attention of vocabulary teaching in the main LT approaches, the implications of the age of the learners or the influence of the language teaching approach followed in vocabulary learning have been discussed.

Some of these aspects are central to this PhD dissertation, which, after all, focusses on lexical competence of CLIL secondary-school learners; in other words, it explores the relationship between the selection of vocabulary learning strategies and vocabulary knowledge. For this reason, it is relevant to consider the vocabulary acquisition process and the dimensions that word knowledge entails. From this starting point, the focus was placed exclusively on specific dimensions of word knowledge, specifically, the recognition and production of the written part of the word, or, in other words, the receptive and productive vocabulary size (for further information about the rest of dimensions, see Nation, 2001 and section 2.1). Similarly, in order to collect the data, it is relevant to explore the different kinds of tools available and whether the format affects the results. This information was provided in section 3 and will be taken up again in Chapter Five in order to justify the instruments used in this study. Finally, it is also central to the interests of this PhD dissertation to understand what CLIL implies for vocabulary learning. Results in this area have pointed to a better performance in receptive and productive vocabulary tests in CLIL learners. Therefore, previous research in the area may aid the understanding of the results.

CHAPTER THREE:
LANGUAGE LEARNING STRATEGIES IN THE ACQUISITION OF VOCABULARY IN AN L2

1. Introduction

This chapter deals with the concept of vocabulary learning strategies (VLSs), understood as observable manifestations of strategic competence to acquire new lexical items. As already indicated, the present thesis includes an essential section on the VLSs used by students in CLIL and regular EFL classes, and it seemed important to establish the theoretical background that underpins the analysis being carried out. Thus, in the present chapter, the connection between the two main aspects of my research — vocabulary knowledge and learning strategies — is established by referring to key concepts from Chapter 2: communicative and lexical competence. As discussed above, strategic competence is a component of communicative competence (Celce-Murcia, 2008). A comprehensive understanding of L2 lexical competence development has to take this important aspect on board.

This chapter will first deal with the Language Learning Strategies (LLSs) construct and connect it to strategic competence. Then, it will focus on the strategies that the literature has identified as specific to vocabulary (VLSs).

2. The Language Learning Strategies construct (LLS)

In a broad sense, LLSs are defined as actions taken by learners to develop the target language. They are thought to reflect the learning processes that take place in the mind and are usually considered to be a manifestation of learners' strategic competence (Phakiti, 2008).

In this section, I will start with the notion of strategic competence, and then define LLSs and identify the main features that distinguish a strategic action. Finally, I will discuss some of the key factors that may affect L2 learners' selection of LLSs.

2.1. Strategic competence

In most recent models of communicative competence, strategic competence is defined as an "inventory of communicative, cognitive, and metacognitive strategies that allow a skilled interlocutor to negotiate meanings, resolve ambiguities, and to compensate for deficiencies in any of the other competencies" (Celce-Murcia, 2008, p. 44). This definition has significantly evolved from Canale and Swain's (1980) approach, where its principal role was considered to be helping meaning transmission. Bachman (1990) extended the concept of strategic competence to include a second function, i.e., assessing one's performance. A third step in the evolution of the concept came when Celce-Murcia (2008) proposed including L2 learning practice itself. This third step involved a significant change in the notion of strategic competence. Whereas previous proposals conceived strategic competence as a tool to aid communication, Celce-Murcia's model involves an analysis of the whole communicative act to learn new language from it. This view is also shared by Agustín Llach and Canga Alonso in one of the latest publications on LLSs. For these authors, "a well-developed strategic competence will aid the process of second language acquisition and will therefore contribute to improving learners' communicative competence" (Agustín Llach & Canga Alonso, 2020, p. 13).

This theoretical development of the notion of strategic competence contributes to placing LLSs in their appropriate context. According to Cohen (2014), strategic competence materialises out of more general language strategies (LSs), which should not be confused with LLSs. He divides LSs into two main groups, according to their different functions. On the one hand, language use strategies (LUSs) aim to facilitate communication and evaluate one's performance, i.e., they match the first two objectives of strategic competence. On the other hand, LLSs, intended to help language development, are related to the third function of strategic competence. Thus, we get a more encompassing view of LLSs and perceive the importance of a theoretical understanding of strategic competence.

Among these various strategies, LLSs have been less explored in the literature (Cohen, 2014). However, from an educational perspective, their relevance is paramount. The use of these strategies has been proved to have a direct impact on language learning (Cohen, 2014; Oxford, 2017), and its exploration is thought to shed some light on the mental processes employed by a learner when developing his or her L2 interlanguage. For this reason, from here on, I will concentrate exclusively on LLS.

2.2. Definition of the LLSs construct

From the beginning, defining LLSs has not been straightforward. For this reason, given the various approaches that have been taken to LLSs, it is important to start with a review of the evolution of the field to have a full understanding of the approach to LLSs that is going to be used in this study.

2.2.1. Contentious issues in the development of the definition of LLS

The origins of the analysis of LLSs are found in researchers' interest in identifying the actions taken by good language learners. From the beginning, researchers attempted to pinpoint the learning behaviours and actions that lead to language proficiency. At that time, most studies were mainly taxonomic; they created numerous LLSs classifications but lacked a solid theoretical basis. When scholars started to develop the concept of LLSs, they quickly found two main points of contention: the identification of the key features that characterise strategically used actions, and terminological inconsistencies.

The first contentious issue in constructing a solid theoretical framework was the distinction between the strategic use of a learning action and the standard learning action itself. In an attempt to solve this problem, scholars decided to complement the definition with some determining features that could distinguish a normal use of a learning action from the strategic use of the same action. The first lists of characteristics (such as Oxford, 1990) were so long that subsequent proposals (Cohen, 1998; Van Patten & Benati, 2010; Weinstein, Husman & Dierking, 2000) focussed exclusively on the constituent that, in their view, was key to recognise strategically used actions: the element of choice, which will be discussed in the following section.

The second problem was related to discrepancies in terminology. Depending on the definition, LLSs were labelled as 'actions' (Cohen, 1996; Scarcella & Oxford, 1992; Oxford, 1990, 2011, 2017), 'behaviours or thoughts' (O'Malley & Chamot, 1990; Weinstein & Mayer, 1986), 'operations' (Oxford, 1990; Rigney, 1978), or even 'steps' (Cohen, 1996; Xhaferi & Xhaferi, 2008). Each of these terms has specific connotations that make them vary to a large extent.

In my view, the inconsistencies in terminology are not contradictions *per se,* but a reflection of the evolution of the LLSs field. As summarised in Table 3.2, the development of the LLS construct has gone through four main stages, each of which is linked to a language learning paradigm. In the earliest proposals, LLSs were defined using behaviourist terms, such as 'behaviours' or 'thoughts'. This first approach to LLSs was soon deemed insufficient, and scholars started to include cognitivist theories, prevailing at that time (the late 1970s, 1980s). Some cognitivist-oriented models likened language learning to computer processing. LLS scholars used this view to conceive LLSs as part of a broader process (the learning one) and defined them as 'operations' that help store and process linguistic information. A third step came with the inclusion of some sociocultural-oriented notions, such as the

importance of learning context and the communicative purposes of L2 learning. In this stage, LLSs are seen as 'learning actions' (O'Malley & Chamot, 1990; Oxford, 1990), and most taxonomical proposals emphasised the social and cognitive dimensions of the LLSs by adopting the terms 'metacognitive', 'cognitive' and 'social' strategies. Finally, the irruption ID factors led to the fourth stage in the development of the LLSs construct. Issues such as 'learning styles', 'ID factors', or 'self-regulation' dominated LLSs scholarly debates. Discussion of the first two elements was soon abandoned, and self-regulation principles were applied to the main definitions. Table 3.1 summarises the four stages in the development of LLSs, including information about the main features of each stage and the different taxonomies developed within them.

Table 3.1

A summary of the evolution of the language learning strategies construct

Stage	Paradigms/ Theories	Features	Representative studies & taxonomies
1	**Behaviour oriented observation**	Mainly taxonomic and focussed on ' good' language learners' actions.	Tax.: Naiman et al. (1975), Rubin (1975)
2	**Cognitivist oriented research**	- Incorporation of SLA. - Models (i.e., processing based perspectives or communicative competence models) are considered in the elaboration of taxonomies.	Tax.: Bialystok (1978), Rubin (1981)
3	**Sociocultural oriented research**	- Importance of cognition in (1) the conscious selection of LLSs and (2) underlying processes that control this selection. - Emphasis on the relevance of social or communicative learning strategies as the final aim of language learning is to enable learners to communicate in society.	Tax.: Oxford (1990), O' Malley & Chamot (1990), Stern (1992)
4	**Constructivist oriented research**	Appearance of individual factors and self-regulation leads up to: 1) Abandoning the LLS construct and move to self-regulation: In this view, LLS are indicators of self-regulation capacity. One of the main criticisms received in this respect is that the self-regulation construct is as broad and vague as the LLSs one (Boekaerts & Corno, 2005). 2) Considering self-regulation as an LLS feature: In this view, studies on LLSs have yielded, at least, a set of defining characteristics that determined learning strategies use, being self-regulation one of them.	Studies: Banisaeid & Huang (2014), Dörnyei (2005) Dörnyei & Ryan (2015) Tax.: Tseng et al. (2006) Studies: Griffiths (2008, 2013), Gu (2012), Macaro (2006) Tax.: Oxford (2011, 2017)

2.2.2. Current definition of LLSs

The LLSs field has benefited from incorporating different language learning notions, which have helped scholars reach an agreement about its most problematic issues. In fact, one of the most widely recognised proposals nowadays, Oxford's (2017), fully integrates some constructivist, cognitivist, and sociocultural notions within her definition of LLSs. As the author puts it, LLSs are

> complex, dynamic thoughts and actions, selected and used by learners with some degree of consciousness in specific contexts to regulate multiple aspects of themselves (such as cognitive, emotional, and social) for the purpose of (a) accomplishing language tasks; (b) improving language performance or use; and/ or (c) enhancing long-term proficiency. Strategies are mentally guided but may also have physical and therefore observable manifestations. Learners often use strategies flexibly and creatively; combine them in various ways, such as strategy clusters or strategy chains; and orchestrate them to meet learning needs. Strategies are teachable. Learners in their contexts decide which strategies to use. Appropriateness of strategies depends on multiple personal and contextual factors.

Oxford, 2017, p. 48

This definition is the one that will be used in this study. Its main advantage lies in the fact that it summarises the main agreements in LLS research. First, it stresses the self-regulatory character of LLSs: learners can select the strategies they consider the most appropriate for any given linguistic task. Second, it shows the multi-dimensional component of language learning, by integrating some elements related to the inner self (cognitive and emotional) with others more related to the conception of language learning as a tool to connect with other people (social). Third, it emphasises the teachable character of the strategies, leading to the conclusion that the teaching practice can have an impact on the LLS selection. In this way, it emphasises that LLSs are not ID factors, but just features of individual learning styles, as they are not fixed and may vary over time. Finally, it also includes the elements of choice and consciousness, which, as will be explained later on, is central to determine the strategic component of learning actions.

In short, Oxford's definition highlights three key features that, in my view, characterise LLSs: self-regulation, consciousness, and learning potential. These three features will be discussed more thoroughly, following the structure of the most widely accepted conceptualisations of LLSs, that is, a definition complemented with some determining features (see, e.g., Cohen, 1996; Oxford, 1990, 2011, 2017).

2.2.2.1. Self-regulation

Some of the more recent definitions and theoretical accounts of LLSs include, in one way or another, the idea of *self-regulation*. Self-regulation is seen as "a property of the person-in-situation and

attend[s] to domain-specific self-regulatory skills that develop through experience within and across situations" (Boekaerts & Corno, 2005, p. 200).

The most renowned theories about self-regulation capacity emerged in the 1990s as an educational research response to the lack of empirical evidence regarding how students became aware of and controlled their learning process (Zimmerman, 2000). Their main premises remain relevant to today's development of educational system policies. For instance, in the case of the latest Spanish Educational laws[2], under the umbrella of the European legislation, learning is considered to comprise, at least, the development of seven key competencies, including 'learning to learn', or, in other words, developing students' ability to regulate their learning process, which corresponds to the development of self-regulation capacities.

Towards the end of the 1990s, self-regulation theories began to be applied to SLA. Studies on LLSs had already pointed to learners' capacity to select strategies, which linked to self-regulation theories. However, until the IPOLLS —International Project on Language Learner Strategies— project and Dörnyei's (2005) claims that LLSs should be considered as individual difference (ID) factors, no efforts were made to clarify the differences between these notions.

In 2004, an international research project —known as IPOLLS— was developed to clarify terminological issues. As part of the project, a questionnaire evaluating their conceptions about LLSs was administered to 23 experts in this field. Part of the questionnaire was devoted to the relationship between LLSs and self-regulation theories. Most participants expressed the conviction that LLSs and self-regulation theories cannot be separated, as some LLSs foundations can only be explained by specific self-regulation notions. However, they disagreed on the weight this construct should have within the LLS construct (Cohen & Macaro, 2007; Cohen, 2014).

The publications derived from IPOLLS coincided in time with Dörnyei's chapter about LLSs in his seminal book on individual differences (2005). In this chapter, the author argued that LLSs and self-regulation capacity were actually two sides of the same coin, as learners' LLSs selection resulted from their self-regulation capacity. Therefore, he stated that there was no longer a need for studying both constructs —learning strategies and self-regulation capacity— and advocated for the use of the self-regulation capacity construct due to its greater acceptance in the academic world and its more comprehensive application.

The contributions of IPOLLS and Dörnyei brought self-regulation into the LLSs field and resulted in an increasing number of researchers studying self-regulation theories to find solutions to the contentious points of LLS theory (Oxford, 2011, 2017; Tseng et al., 2006). In this respect, Oxford's

---

[2] *Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa* (2013) and *Ley Orgánica 2/2006, de 3 de mayo, de Educación* (2006).

effort (2011) to integrate LLSs and self-regulation resulted in one of the most widely accepted proposals. Without abandoning the LLS perspective, Oxford designed a model that unified different positions by integrating self-regulation theories into LLS proposals: the Strategic Self-Regulation (S^{2}R) model.

S^{2}R combines traditional and new tendencies, opening an original path with an inclusive proposal in which different views concur. Oxford integrated some basic self-regulation principles into LLSs by making two significant modifications to previous taxonomies: renaming of the groups and including metastrategies. She changed the traditional names given to the groups of strategies (such as cognition, metacognitive or social), to coincide with self-regulation layers (cognition, affect, and interaction). She also introduced a new element, metastrategies, which refers to learners' mechanisms to regulate their own use of strategies. The author identified a total of eight metastrategies, which she subdivided into three groups (meta-cognitive, meta-affective, and meta-social interaction [meta-SI] strategies) based on the dimensions they regulate. Traditionally, most taxonomies had included metacognitive strategies. With the addition of the other two groups of metastrategies, Oxford's model presents a better balance of dimensions and adds weight to the emotional and social dimensions, which were scarcely present in previous taxonomies.

Self-regulation theories have enriched the LLS field, and Oxford's proposal is a good example of how scholars have incorporated self-regulation principles into the LLS area. Their application to LLSs has gone beyond the mere development of taxonomies and has tackled the core of the field: the conceptualisation of the learning process itself. Early LLS researchers found it difficult to explain the mental processes by which LLSs were chosen. With self-regulation notions, most LLS theorists now emphasise the critical role the learner has in his or her own learning process. It is the learner who is expected to regulate his or her learning and select the most appropriate tools or actions to make the most of any language experience. As Oxford puts it, strategically self-regulated learners are considered to have an active role in their own learning process, by controlling their "cognitive and affective states, […] their observable performance […] and the environmental conditions for learning" (2011, p. 15). This notion of self-regulated learners is central to this PhD dissertation, as one of its objectives is to explore and analyse how the implementation of a CLIL approach may affect the vocabulary learning process, which may be observed in the VLSs selection a learner makes.

2.2.2.2. Consciousness and explicitness of the action

From the beginning, one of the most complex issues in the understanding of LLSs was the distinction between learning strategies and ordinary actions carried out in the learning process, i.e., learning in itself. In this respect, one of the first proposals came from Cohen (1998), a pioneer in identifying the element of choice as a key definitional feature. In his view, there exists a conscious

process underlying the selection of learning strategies, and learners' election is what distinguishes LLSs from any other process taking place in language learning.

Consciousness and learners' choice were also recognised as central elements in LLSs by most respondents of the IPOLLS survey (Cohen, 2007; 2014). For this reason, in his most recent proposals Cohen (2007, 2014) emphasised the idea of *choice*, stating that "if the behaviour were so unconscious that the learners are not able to identify any strategies associated with it, then the moves or functions associated with this behaviour should probably be referred to simply as processes, not as strategies" (2014, p. 11). However, he agreed with Dörnyei in describing *consciousness* as a "notorious vague term" (2009, p. 132) and proposed instead the term *attention,* which encompasses "a variety of mechanisms or subsystems, including alertness, orientation, detection, facilitation and inhibition" that control access to consciousness (Dörnyei, 2009, p. 132). In his view, attention can be viewed as a continuum, ranging from "the learner being fully focused on the strategy at one end, to the learner giving the strategy only minimal attention to the other" end (Cohen, 2014, p. 11). He further explains that, on this continuum, the attention given to a strategy may shift during the process, because (as some respondents of the survey highlighted) "the strategy might be at the centre of attention, but as the plan is carried out, the strategy is then reduced to peripheral attention, then to a stand-by mode, and perhaps ultimately to a "no attention" mode" (p. 12). However, this reconceptualization does not clarify the main question, i.e., the degree of attention needed for an action to be considered strategic.

In this sense, one of the latest contributions to this discussion has been made by Oxford (2017). In her view, the notion of consciousness and attention in LLSs can be explained using Schmidt's model of consciousness in language learning (1995). Schmidt identifies four elements of consciousness for language learning: attention, awareness, intention, and cognitive effort. Oxford argues that while the first three elements are constant and central when using LLSs, and cognitive effort is involved in many cases, although it is not always necessary. Figure 3.1 below summarises Oxford's adaptation of Schmidt's model of consciousness.
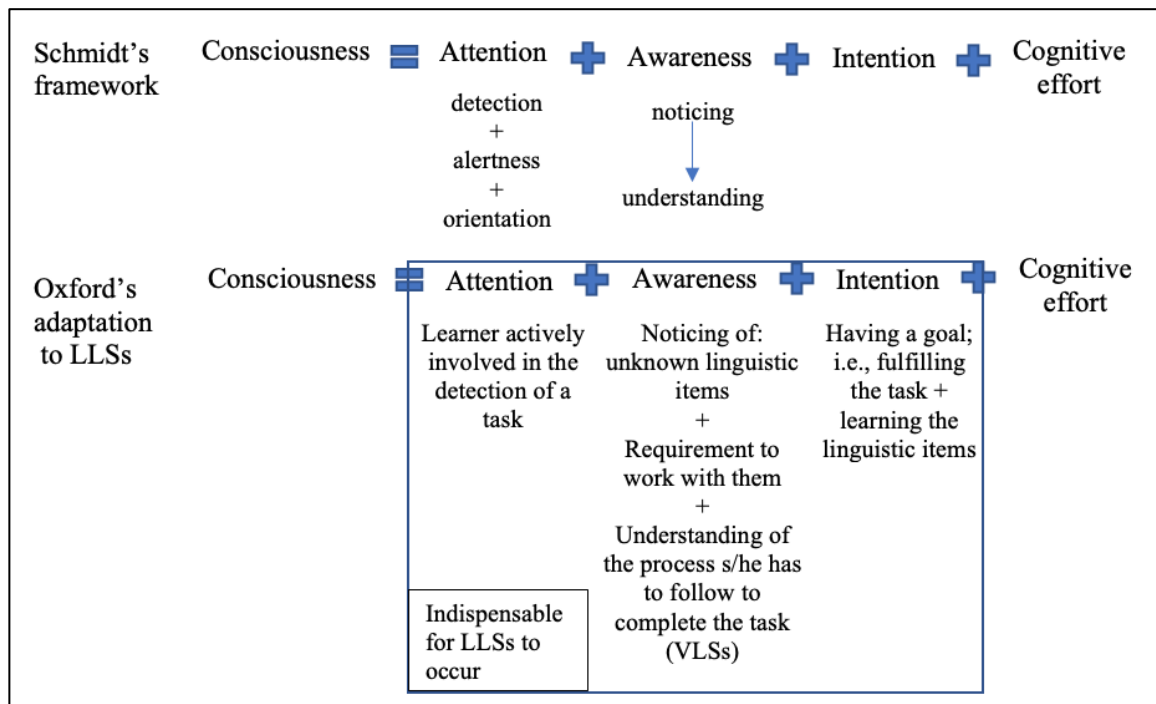
Figure 3.1. Visual representation of Oxford's adaptation to LLSs of Schmidt's model of consciousness for language learning.

2.2.2.3. Potential for learning

Oxford (2017) highlights that 97% of the most widely accepted definitions of LLSs include, either explicitly or indirectly, the sense of learning. Thus, a vast body of research has attempted to identify the most useful strategies for language learning (see, for example, Dreyer & Oxford, 1996; Green & Oxford, 1995; Wharton, 2000; Ying-Chun, 2009), with little consensus among them (Pawlak, 2011).

While researchers disagree about the most useful strategies, they are in agreement that learning is more likely to happen if strategies occur in clusters. Research on this issue highlights that the use of a strategy *per se* does not ensure language development; success depends upon how the learner uses and combines strategies to face linguistic tasks (Chamot, 2004; Cohen, 2007, 2014; Gu, 2003; Cohen, 2007; 2014; Ehrman, Leaver & Oxford, 2004; Hsiao & Oxford, 2002; Oxford, 2002, 2003). In this respect, some authors (Oxford, 2002, 2003; Ehrman, Leaver & Oxford, 2003) argued that the usefulness of any strategy is determined by three main conditions: how well it relates to the task, the extent to which it fits with a particular student's learning preferences, and how it is incorporated with other relevant strategies. For instance, strategies such as 'use of word lists' or 'skimming the text' may suffice for accomplishing simple tasks. Still, they may not be enough when dealing with more complex communicative situations or tasks, in which a combination of strategies is needed.

This section has identified three definitional features of the LLSs construct: their self-regulatory character, learners' consciousness, and potential for learning. These three features constitute the basis of the analysis of VLSs in this dissertation.

Concerning the first feature (learners' self-regulatory capacity), L2 learners are thought to select the most appropriate LLSs according to, among other factors, a preliminary analysis of the task and previous experience. Given that participants in this study were exposed to two different teaching approaches, they are expected to differ in their language learning experiences and to have other linguistic demands. The aim is to examine whether these differences affect learners' VLSs selection and how this is affected.

The notion of choice and consciousness, for its part, will be used to discriminate learning actions from VLSs. In this dissertation, any learning action is considered strategic if a learner can recognise and report on his or her own use. The ability to report, in my view, indicates that the learner is conscious of the use of the strategy, as he or she (1) is aware of the performance of the action, (2) identifies it as strategic, and (3) can identify and pay attention to the process.

Finally, as for learning potential, research shows that LLSs are likely to foster L2 development when various LLS are grouped and used together. This finding will be used to build the framework of analysis of VLS use.

Given the importance of strategy combination for language development, strategy choice will be explored in two ways. First, the use of each strategy will be studied in isolation to identify the most and least preferred strategies. Second, the use of strategies in groups will be examined. Given that learners can combine LLSs in different ways, strategies will not be clustered according to their function (i.e., cognitive, metacognitive, social, or memory) but will be grouped according to the ways the participants in this study cluster them. This discussion will be taken up again in Chapter Four when explaining the methodological details of the analysis.

2.3. Factors Influencing the LLSs Selection

As shown in previous sections, there is a component of personal choice in the selection of LLSs. This selection has been proven to be affected by a series of factors, which are going to be explored in this section. Following Oxford's (1989) and subsequent studies, factors are grouped into three main areas: learners' linguistic background, learners' ID, and teaching approaches.

2.3.1. Personal factors: learners' linguistic background and ID

As for the role of learners' linguistic background, two factors have been examined: L1-L2 resemblance and L2 proficiency. Regarding the former, some studies suggest that the degree of similarity between the L1 and target languages affects the selection of strategies (Chamot, O'Malley, Küpper, & Impink-Hernandez, 1987; Politzer, 1983; Wharton, 2000). However, given the little research available on this issue, this claim should be taken with caution. In contrast, studies into the connection between strategy-selection patterns and L2 proficiency are more numerous and conclusive. They show

that LLS selection seems to be a dynamic process, clearly influenced by the level of L2 proficiency (Chamot et al., 1987; Nyikos, 1987; Politzer, 1983; Oxford, 1989, 2011; Oxford & Nyikos, 1989; Park, 1997; Psaltou-Joycey & Kantaridou, 2009a; Vrettou, 2011; Wharton, 2000). Besides, the research analysed suggests that (1) the higher the L2 level, the greater inclination for strategy use, and (2) highly advanced L2 learners tend to reduce their use of affective strategies, and over time increasingly adopt metacognitive strategies. However, most of these studies explore the selection of LLSs by advanced and intermediate learners, with little research examining beginners. In my view, the field would benefit from an analysis of the relationship between CERFL levels and LLSs. Considering CERFL would result in more inclusive research, in which use of LLSs by a more varied sample of L2 learners (not only intermediate and advanced L2 learners) would be explored. Moreover, it would allow determining whether the selection of LLSs varies, and, if so, in which ways, as language proficiency increases.

Learners' ID factors, for their part, this has been an extensively studied phenomenon, with efforts mainly concentrated on factors such as learning styles, gender, and age. By far, the most productive area of study has been the influence of learning style on strategy choice. As pointed out above, learning styles and learning strategies were commonly confused in early research. This need to differentiate the terms gave rise to numerous studies, in which the influence of learning styles on learning strategies was acknowledged. Ehrman and Oxford (1990) and Griffiths (2004) concluded that learning styles play a crucial role in learners' choices. However, research also suggests a change in the predetermined LLS selection when strategies are explicitly taught; in other words, given the teachable nature of LLSs, learners' selection of LLS will vary as they are exposed to other actions and methods of learning. In this respect, Ehram and Oxford (1990) found that, for example, although extroverts felt more comfortable using social strategies than did introverts or thinkers, introverts resorted to social strategies more often when explicitly trained.

Regarding gender and age, findings are somewhat inconclusive. In the case of gender, early research seemed to show that females made greater use of strategies (Ehram & Oxford, 1989; Kaylani, 1996; Lan & Oxford, 2003; Oxford & Nyikos, 1989; Oxford, Nyikos, & Ehram, 1988; Peacock & Ho, 2003; Politzer, 1983) due to their "greater social orientation, stronger verbal skills, and greater conformity to norms" (Oxford 1989, p. 238). More recent studies have reached opposite conclusions (Tercanlioglu, 2004; Wharton, 2000), and some studies have even found that there is no evidence of gender differences (Ehram & Oxford, 1990; Griffiths, 2003; Psaltou-Joycey, 2008). Research on this issue has become to a dead-end, because, as males and females now usually take classes together, the issue is no longer a priority from an instructional perspective (Chamot, 2004).

Age, for its part, has repeatedly been explored due to the importance of maturational constraints in SLA. Research indicates that students of different ages use different strategies, with older o students using more sophisticated LLSs (Bialystok, 1981; O'Malley & Chamot, 1990; Politzer, 1983; Psaltou-

Joycey & Sougari, 2010; Tyacke & Mendelsohn, 1986). However, studies do not show consistent results regarding the strategies fostered over time. Peacock and Ho (2003), for instance, explored the use of LLSs by two groups of adult English for Academic Purposes (EAP) learners and concluded that older learners (aged 24-29) made more use of strategies in general, and memory, metacognitive and affective strategies in particular, in comparison to younger learners (aged 18-22). For her part, Griffiths (2003) analysed a sample of secondary-school learners (aged 14-16) studying English as a foreign language and concluded that age did not affect the use of strategies. However, Griffiths (2003) suggested that her findings could be due to the short age difference of the groups. In this respect, Psaltou-Joycey (2010) pointed out that discrepancies among studies may be an effect of the interrelation of age with other factors, such as L2 proficiency level, or culture. Therefore, there is still some dispute about the influence of age on LLS selection. In this respect, this dissertation aims to make a contribution to this field by exploring LLSs use by teenagers. As has been observed, most research into LLSs explore adult LLSs. Foreign languages are increasingly taught to young and very young learners in many parts of the world (Castellano-Risco et al., 2020; García-Mayo & Gutiérrez-Mangado, 2020). The exploration of LLSs at these early ages would greatly contribute to understanding whether this variable influences how language is processed.

In sum, there are some learners' features that seem to affect the selection of LLSs. The three main features explored in the literature are learning styles, gender and age. In the case of the first feature, some relationship between the selection of particular learning strategies and the learner's learning style has been shown to exist. As for the second feature, it seems to be generally accepted that females make more use of strategies than males, although this difference is not significant. Finally, there is a lack of clear and conclusive findings regarding age.

2.3.2. Language approaches: LLS selection and CLIL

A third branch of research has focussed on the influence of the teaching practice itself. The relationship between the selection of strategies and the type of linguistic tasks to which learners are confronted has been extensively explored in the literature. Bialystok (1981), a pioneer in the area, reported that learners used different strategies depending on the objective of the task, finding that some strategies were only considered useful for certain activities. Following Bialystok, there have been a growing number of studies analysing the relation of strategies and tasks devoted to the development of the four language skills: listening (Bacon, 1992; Vandergrift, 1997), reading (Barnett, 1989; Hayati, 2005; Konishi, 2003), writing (Manchón, 2001; Trenchs, 1996) and speaking (Cohen, Weaver, & Li, 1998).

The kind of activity is not the only issue concerning the influence of the teaching practice; the teaching methods and approaches employed have also been related to the area of LLSs. Researchers have found that LLSs selection depends, to some extent, on the type of teaching approaches used. In

this respect, the intrinsic nature of CLIL, the teaching approach explored in this dissertation, could be taken as an indication that this educational approach may affect the L2 learning process. Content learning in an L2 is more cognitively challenging than in the L1, as content is both conceptually and linguistically new to learners. Besides, CLIL promotes the use of certain kinds of tasks that place greater cognitive efforts on learners. Based on constructivist principles, CLIL is learner-centred, and learners are often prompted to develop their own knowledge by adopting a problem-solving approach.

Mehisto et al. (2008) made use of Bloom's taxonomy of educational objectives to explain the cognitive demands involved in CLIL. As shown in Figure 3.2, Bloom (1984) classified thinking skills into six groups ranging from Lower Order Thinking Skills (known as LOTS) to Higher Order Thinking Skills (HOTS) and arranged them into a pyramid. Mehisto et al. (2008) consider that CLIL promotes HOTS tasks in class, as learners are exposed to progressively higher levels of cognitive challenges.
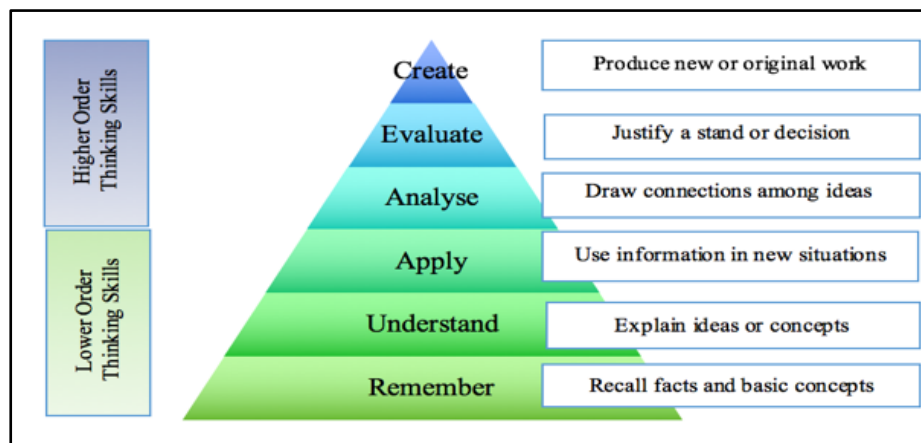
Figure 3.2. Adaptation of Bloom's taxonomy.

Thus, it could be said that CLIL involves great cognitive demands derived from (1) the use of a foreign language to learn new content, and (2) the range of activities and methodologies employed. However, despite the importance of cognition in CLIL, research into the cognitive implications of CLIL is still scarce, with most efforts concentrating on providing a clear and strong theoretical basis for real-world practices.

To the best of my knowledge, only two studies have investigated LLS selection in CLIL: Psaltou-Joycey, Mattheoudakis & Alexiou (2012) and Milla & Gutierrez-Mangado (2019). In the first study, CLIL and EFL primary school learners' LLSs were compared (grades 4-6; age 9-12), concluding that CLIL learners made greater use of (1) strategies in general, (2) strategies that help them to solve very challenging situations, and (3) strategies more focussed on communication. The second study also focussed on primary-school learners (grades 5 and 6; age 10-12), but the profile of participants and the aim of the study were different. In this case, the study dealt with Basque/Spanish bilingual learners of English as a third language in a CLIL context, and the objective was to explore the selection of LLSs by these CLIL learners and how they varied according to age, language proficiency and gender. To

achieve this aim, an adaptation of the SILL questionnaire (Oxford, 1990) and the Cambridge English FLYERS test were administered to 131 learners. Results showed that upper-primary school children preferred social, memory, and metacognitive strategies. However, no significant differences considering language proficiency and gender were reported.

This section has discussed several issues related to LLSs, two of which (teaching approach and age) are central to the subject of this dissertation as they seem to have some impact on the selection of strategies. As for teaching approach, the few studies on CLIL have concluded that these learners make great use of strategies aimed at fostering communication in the classroom. They are called 'social strategies' by Millá and Gutierrez-Mangado (2019), or 'strategies focused on communication' by Psaltou-Joycey, Mattheoudakis and Alexiou (2012). Moreover, the only study comparing CLIL and regular EFL learners (Psaltou-Joycey et al., 2012) finds that CLIL learners make greater use of strategies in general, and metacognitive and social strategies in particular. Given the considerable success of CLIL in Europe, greater efforts should be carried out to explore its impact on language processing. LLSs could well serve this purpose. As for the second factor (age), its importance in this dissertation is paramount, given that participants are secondary-school learners studying English in a formal context: the school. As shown, most research focusses on strategy choice made by adult learners of English. However, the little research available suggests that learners' LLS selection changes with age, with an increasing number of strategies being used as learners grow. Table 3.2 summarises the main findings in relation to how learners' language background, learners' characteristics and the teaching context affect the selection of LLSs.

Table 3.2

Main factors influencing language learning strategies selection

	Factor	Findings	Studies
Learners' language background	Language proficiency	Inconclusive results as regards the relationship between language proficiency and the number of strategies used.	- Oxford & Nyikos, 1989; Park, 1997; Psaltou-Joycey & Kantaridou, 2009a, 2009b; Vretou, 2011; Wharton, 2000: positive correlation. - Mullin, 1992; Hong-Nam & Leaven, 2006: negative correlation - Curvilinear relationship: Philips (1991).
	Duration of the L2 learning process	More and less experienced learners' strategy use differs.	- Griffiths; 2003; Oxford & Nyikos; 1989.
Learners' ID factors	Learning style	Predetermined used of LLSs influenced by the learning style, although LLSs use may vary when their use is explicitly trained.	- Ehrman & Oxford 1990; or Griffiths, 2004.
	Gender	Greater use by females	- Oxford & Nyikos, 1989; Ehram & Oxford, 1989; Jiménez Catalán, 2003: females' higher use of strategies. - Ehram & Oxford, 1990; Griffiths, 2003; Psaltou-Joycey, 2008: no differences.
	Age	Inconclusive results	- Peacock & Ho, 2003: older students make more use of memory, metacognitive and affective strategies. - Gavriilidou, 2004: increase of cognitive and metacognitive strategies. Lower use of socio-affective strategies with age.
L2 teaching	Type of approach	Strategies vary depending on the teaching context	- CLIL fosters the overall use of LLS (Psaltou-Joycey et al., 2012). Increment mainly observed in social, memory and metacognitive strategies (Milla & Gutierrez-Mangado, 2019).
	Type of task	Strategies vary depending on the task	- Bacon, 1992; Cohen et al., 1998; Hayati, 2005; Konishi, 2003; Trenchs, 1996.

3. Vocabulary Learning Strategies (VLS)

So far, I have referred to LLSs rather than to VLSs. LLSs are usually classified according to the language skill (e.g., reading or writing) or area (e.g., grammar or vocabulary) to which they are applied. It was not until the 1990s that LLSs started to be applied to lexical development. Many lexical studies emphasised the need to understand how the L2 lexis was developed. Together with the increasing research into LLSs, this concern resulted in applying the findings of this research to strategies for vocabulary acquisition. In this section, the main VLSs definitions, taxonomies and studies on VLS use will be reviewed.

3.1. Definition

Finding a concise definition for VLSs is just as difficult as in the case of LLSs, because there are as many definitions as approaches to the construct. Some authors directly refer to VLSs as a sub-group of LLSs. For instance, Nation (2001, p. 217) defines VLSs as a "part of language learning strategies, which in turn are a part of general learning strategies". Although it is undoubtedly a suitable definition, Nation does not explain what features or characteristics define VLSs; he simply frames VLSs as a part of LLSs.

Other authors (e.g., Cameron, 2001; Jiménez Catalán, 2003; Intaraprasert, 2004, and, more recently, Oxford, 2017) have attempted to narrow down the definition. Most of these researchers emphasise that L2 learners make use of VLSs with three main aims: (1) understanding new words, (2) consolidating word meanings, and (3) expanding lexical knowledge. However, they differ in the term serving as a frame for the definition. Thus, VLSs are defined as 'actions' (Cameron, 2001; Jiménez Catalán, 2003), 'techniques' (Intaraprasert, 2004), 'behaviours' (Intaraprasert, 2004; Oxford, 2017), 'thoughts' (Oxford, 2017), or even 'knowledge about the mechanisms' (Jiménez Catalán, 2003). As with LLSs, it is clear that we are referring to different things when we refer to them as 'actions' and 'knowledge'.

However, except for Oxford (2017), most VLSs definitions were developed in the early 2000s, and do not reflect notions that nowadays are considered essential to understanding the construct. Given the vast amount of field research conducted in the last two decades, to consider VLSs to be part of the LLSs construct, one needs to include the new elements discussed in the previous section, such as choice, potential for learning, and self-regulation theories.

The real question is to determine whether a definition of VLSs is even needed in the first place. We have to consider that LLSs taxonomies, given their general outlook, cannot include every single strategy used to deal with linguistic skills or areas of any given language. In his or her own opinion, each author merely collects the most relevant or representative strategies to develop the different

language areas and skills. Following the same principle, the taxonomies of VLSs could be simply a sort of 'repository' of all the LLSs related in some way or another to vocabulary development. Under this approach, the definition of VLSs would be included in that of LLSs.

The definition of VLSs I adopt in this doctoral dissertation is an adaptation of Oxford's latest LLS definition (2017) in which I have replaced the primary purposes of LLSs with the ones identified by Jiménez Catalán (2003) and Intaraprasert (2004) for VLSs. Thus, in this dissertation, VLSs are regarded as complex, dynamic thoughts and actions, selected and used by learners with some degree of consciousness in specific contexts to regulate multiple aspects of themselves for the purpose of (a) understanding and retaining in long-term memory the meaning of previously unknown words, (b) recalling this knowledge when in need, and (c) expanding vocabulary knowledge. Strategies are mentally guided but may also have physical and therefore, observable manifestations. Learners often use VLSs flexibly and creatively; combine them in various ways, such as in strategy clusters or strategy chains; and orchestrate them to meet learning needs. Strategies are teachable. Learners in their contexts decide which strategies to use. Appropriateness of strategies depends on multiple personal and contextual factors.

In my view, this adapted definition provides a clear idea of what VLSs are, their primary purposes, and how they work in practice. It also includes some basics of lexical knowledge research, such as the dichotomy of receptive and productive knowledge, together with other more psychological-based notions.

3.2. Main taxonomies

The lack of agreement on what VLSs are has not hindered the development of numerous taxonomies. However, inventories differ significantly, making both the extrapolation of findings from and comparison among studies difficult.

Researchers have tackled the development of VLS taxonomies from three different perspectives. Some authors (Cook, 2001; Decarrico, 2001; Nation, 2001; Webb & Nation, 2017) have presented theoretical proposals, not supported with empirical data, based on specific vocabulary learning theories and notions, such as the incidental component of learning (Decarrico, 2001), and autonomous learning (Webb & Nation, 2017). Other authors (e.g., Lawson & Hogben, 1996) have followed a qualitative approach to identify and classify VLSs. In these studies, data is gathered through interviews and think-aloud procedures, which results in data sets rich in details. However, samples are usually small, and findings often lack the significance level needed to extrapolate those findings. The third, and predominant, approach is illustrated by those studies that have adopted a quantitative approach through which VLSs are identified by making use of large samples, and classifications are grounded on robust

statistical analysis (Stöffer, 1995; Gu & Johnson, 1996; Schmitt, 1997; Hedge, 2000; Intaraprasert, 2004; and Tseng et al., 2017).

As a consequence of this diversity, taxonomies differ greatly. By taking a closer look at Table 3.3 below, discrepancies can be perceived. First, even though taxonomies are expected to be limited to strategies used to learn vocabulary, the strategies included in some VLSs taxonomies, such as the self-regulating capacity in vocabulary learning scale (Tseng et al., 2006), could well be applied to language learning in general or even to learning in general, rather than to vocabulary learning specifically. Moreover, proposals differ significantly in the number of VLSs identified and in the number and kind of participants, if any, used to pinpoint the strategies. In this regard, most scholars have examined adults' selection of VLSs in an EFL setting. However, there are some exceptions, such as Lawson & Hogben (1996), who explored Australian learners of Italian, or Schmitt (1997), who examined VLSs selection by adolescents. Finally, these classifications are usually based on the authors' own perceptions rather than on data and statistical analysis. In fact, some inconsistencies are found in the grouping itself: it is common to see together included in categories aspects more related to psychological issues, such as 'cognitive strategies' or 'metacognitive strategies' (Schmitt, 1997), and others more related to methods or even materials, such as 'dictionary strategies' (Stöffer, 1995).

Table 3.3 summarises the main VLS taxonomies with details about the number of items included, the kind of participants used to identify the VLS, if any, and the main categories identified. A more detail account of the taxonomies, with a list of the different strategies included in each proposal is found in appendix A.

Table 3.3

A summary of the main VLS taxonomies

Taxonomy	No. of items	Sample	Categories	
Stöffer (1995)	53	707 university students	1. Authentic language use 3. Overcome anxiety 5. Create mental linkages 7. Creating activities 9. Memory strategies	2. Organise words 4. Physical action 6. Self-motivation 8. Visual/auditory strategies
Gu and Johnson (1996)	91	850 advanced adult Chinese learners	1. Beliefs about vocabulary learning 3. Dictionary strategies	2. Metacognitive regulation 4. Note-taking strategies 5. Memory strategies
Lawson & Hogben (1996)	15	15 adult EFL learners	1. Repetition 3. Word feature analysis	2. Simple elaboration 4. Complex elaboration
Schmitt (1997)	58	600 teenagers and adults EFL learners	1. Discovery: determination and social strategies	2. Consolidation: cognitive, metacognitive memory & social strategies
Nation (2001)	13	-	1. Planning 2. Source 3. Processing	
Cook (2001)	7		1. Strategies for getting meaning	2. Strategies for acquiring words
Decarrico (2001)	4	-	1. Guessing meaning from context 3. Use of vocabulary notebooks	2. A mnemonic device or the keyword method 4. Other learner strategies
Intaraprasert (2004)	33	133 EST adult learners	1. Strategies to discover the meaning of new items. 3. Strategies to expand the knowledge of vocab. Items	2. Strategies to retain the knowledge of newly-learned items
Tseng et al. (2006)	20	193 adult learners	1. Commitment Control 3. Satiation Control 5. Environment Control	2. Metacognitive Control 4. Emotion Control
Webb & Nation (2017)	6	-	1. Finding ways to encounter the L2 outside the classroom 3. Use of flashcards 5. Learning word parts	2. Use of dictionaries effectively 4. Finding ways to use the L2 outside the classroom 6. Guessing from context

Among the different proposals, Schmitt's taxonomy (1997) was selected for this study. Three reasons support this decision. First, although this taxonomy was designed over twenty years ago and does not include the most recent notions that characterise a VLS, its development was grounded on a solid theoretical background (such as Oxford, 1990; Purpura, 1994) and I considered that its theoretical conceptualisation could be updated to meet the purpose of this study by including principles such as self-regulation, choice and learning potential. Second, as will be seen later, it is the only taxonomy in which L2 secondary-school learners were included as part of the sample. This fact is fundamental, as Schmitt's sample is similar in age to the one studied in this dissertation. Secondary-school learners are adolescents who have not reached adulthood's cognitive maturity and their selection of strategies may differ from that of adults, so it is important to use a taxonomy adapted to this age group. Finally, the large number of strategies that Schmitt's taxonomy includes and, more importantly, its careful selection, make it one of the most inclusive proposals. In this respect, given the need to adapt the taxonomy to the participants of my own study, the existence of many items is an advantage. Below, the compilation procedure of the VLSs included in Schmitt's taxonomy, and my resulting taxonomy, are described.

3.2.1. Schmitt's taxonomy

In 1997, Schmitt presented one of the best-known VLS taxonomies up to date. Taking as a starting point Rubin's definition of learning, i.e., "the process by which information is obtained, stored, retrieved, and used" (1987, p. 29), Schmitt attempted to identify the main LLSs that served vocabulary learning. The compilation process took information from various sources. First, he examined a vast number of manuals and reference books, from which he selected the majority of initial VLSs to be included in the final taxonomy. Second, Japanese intermediate level students were asked to write a report about how they studied vocabulary items. From the analysis of these reports, some additional strategies were added. Finally, some teachers were requested to give their opinion about the list and to add other strategies they considered relevant. The initial list consisted of 40 strategies. However, to check whether there were omissions, a survey was administered to a Japanese intermediate learners pilot group. As a result, the final taxonomy list included a total of fifty-eight strategies.

Once the strategies were compiled, the second step was to categorise them, for which Schmitt required a theoretical framework to support the classification. After a thoughtful analysis of several taxonomies for LLSs (such as Bialystok, 1981; Chamot, 1987; O'Malley & Chamot, 1990; and Oxford, 1990), Schmitt decided to adapt Oxford's, due to its detailed analysis and grounded theoretical basis.

Oxford organised the LLSs into six main groups, which, in turn, were grouped in two broader ones: direct and indirect strategies. Schmitt decided to maintain the two layers of grouping but changed some of the categories. First, he claimed that in Oxford's proposal there was no category representing the way learners discovered the meaning of a new word without invoking someone else's help. Therefore, he decided to create a new category: determination strategies. Second, he considered that some of Oxford's

strategies were inadequately classified. Based on the idea that LLSs are multi-faceted and can serve to approach vocabulary learning from more than one perspective at a time, he decided to rearrange some of the strategies so that they fitted better into his framework.

Schmitt identified an imprecision in the categorisation of the most common clusters. In his view, memory and cognitive items shared the same aim, i.e., "to assist recall of words through some form of language manipulation" (1997, p. 6), which he saw as a flaw. To clarify the issue, he resorted to Purpura's (1994) classification of memory strategies. According to Purpura, memory strategies can be divided into six methods: repeating, using mechanical means, associating, linking with prior knowledge, using imagery, and summarising. Schmitt suggested that cognitive strategies were more related to the first two items, as they were less obviously linked to mental manipulations. In contrast, memory strategies involved the other four methods, because, in his view, they "are somewhat closer to traditional mnemonic techniques which either organise mental information together or transform it in a way which makes it more memorable" (2001, p. 16).

Finally, Schmitt classified the VLSs he had previously identified into five groups (determination, social, memory, cognitive, and metacognitive), which, in turn, were collapsed into two more general clusters: those strategies involving actions to understand the meaning of a new word (determination and social) were included in the 'discovery' group, whereas those strategies that aided the retention of the new meanings (social, memory, cognitive, and metacognitive), were considered 'consolidation' strategies. Each group is defined in Table 3.4 below, which shows Schmitt's classification of VLSs.

Table 3.4

Schmitt's taxonomy (1997)

Groups	Sub-groups	Definition	No. of items
Discovery	Determination	L2 learning actions taken by learners to individually discover the meaning of an unknown word.	8
	Social	L2 learning actions in which learners interact with other people (teachers, classmates, L2 speakers…) to understand new meanings of unknown words.	5
Consolidation	Social	L2 learning actions learners carry out to retain the meaning of L2 words by interacting with other people.	3
	Memory	L2 learning actions used for "relating the word to be retained with some previously learned knowledge, using some form of imagery or grouping" (Schmitt, 1997, p. 15)	28
	Cognitive	L2 learning actions that involve learners' manipulation or transformation of the target language.	9
	Metacognitive	L2 learning actions that involve "a conscious overview of the learning process and making decisions about planning, monitoring or evaluating the best ways to study" (Schmitt, 1997, p. 17).	5

3.3. Studies into VLSs selection by young and adult learners of English

Since the emergence of research into VLSs, its main focus has been the identification and classification of strategies and the understanding of their selection by learners. As soon as the first classifications emerged, studies on the use of VLSs flourished. These studies had two main foci: some studies attempted to identify the preferred and least used strategies by L2 learners, while other studies aimed to relate the selection of VLSs to other variables. Given that the studies differ, among other things, on the age of the participants, and considering that age and duration of the L2 learning process may affect the selection of VLSs (see section 2.2) and that this dissertation focusses on secondary-school learners, the studies reviewed below are clustered according to the sample (secondary-school learners or tertiary learners) they analyse.

3.3.1. University students' VLSs selection

Studies exploring university students' choice of VLSs are quite varied in relation to the aim of the study, the type of learners examined, the sample size, and the instruments used.

Studies exploring the selection of strategies in isolation, i.e., without considering other factors, are rarely found in the literature and are usually part of broader studies aiming to compile VLSs taxonomies.

Such studies have attempted to identify the most and least used VLSs, but results are somewhat inconclusive. On the one hand, authors such as Lawson & Hogben (1996) and Shabazian (2004) have found that their subjects' preferred strategies are repetition of new words, memorisation, and mnemonics. For these authors, adults prefer mechanical strategies, which place lower cognitive demands on the learners. On the other hand, other authors (Gu & Johnson, 1996; Schmitt, 1997; Xhaferi & Xhaferi, 2008) have found that adult EFL learners prefer more 'meaning-oriented strategies' (e.g., guessing from context, using a dictionary, paying attention to the word form, and using newly-learned words in sentences) than rote (i.e., repetition) strategies.

Thus, these studies have opposite outcomes. Differences in methodological aspects may partly cause these discrepancies; as can be seen in Table 3.5, the quantitative features of the samples vary greatly: while some studies are based on the analysis of quite large samples (500+ participants in Gu & Johnson, 1996; or Schmitt, 1997), it is also possible to find other studies with much smaller samples (35 participants in Lawson & Hogben, 1996). Moreover, the studies provide no detailed information about the participants' L2 learning backgrounds and levels, and they also differ in the methods used to gather data: some authors preferred self-report protocols (Schmitt, 1997; Sahbazian, 2004; Xhaferi & Xhaferi, 2008); others (Lawson & Hogben, 1996) chose to use interviews, think-aloud protocols and direct observation. Table 3.5 summarises the methodologies and results of the studies reviewed in this section.

Table 3.5

A summary of some studies on university L2 learners' VLSs selection

Study	Methodological aspects	Preferred or most widely used strategies
Gu & Johnson (1996)	Instrument: *VLSQ* Sample: 850 Chinese undergraduate learners of English.	- Guessing from context, using dictionary, paying attention to a word form and using new learned words in sentences.
Lawson & Hogben (1996)	Instruments: interviews and think-aloud protocols Sample: 15 adult learners of Italian	- Repetition of the new words and their meanings.
Schmitt (1997)	Instrument: VLSQ based on Schmitt's taxonomy Sample: Japanese intermediate EFL learners.	- Bilingual dictionary and guessing from textual context.
Sahbazian (2004)	Instrument: 35-item questionnaire	- Memory and mnemonic strategies.
Xhaferi & Xhaferi (2008)	Instrument: *VLSQ* Sample: Albanian EFL undergraduate students	- Asking teachers for meaning, making guesses and making list of new words.

Studies on how the use of VLSs is related to other variables are, by far, more frequent in the literature. Those based on gender differences stand out notably, but those that focus on other variables also deserve a closer look.

Results on gender differences seem to be inconclusive. For example, while Intaraprasert (2000) found no significant gender differences in the overall strategy use, he concluded that, at the individual level, female learners made greater use of some strategies ('attending class regularly', 'asking a classmate to solve the problem encountered in classroom lessons', and 'practising translating from Thai into English'). In line with this latter finding, Jiménez Catalán (2003) found that both genders shared eight out of the ten most frequent strategies, although females used a larger number of strategies than males. However, Xhaferi & Xhaferi (2008) found no significant gender differences in Albanian EFL learners' use of VLS. Given the limited number of studies and the lack of consistency in the findings, further research in this area is needed.

Other studies (Gu & Johnson, 1996; Hamzah et al., 2013; Wadvogel, 2013) have explored the use of VLSs in relation to the language level with diverse findings. Gu & Johnson found that while L2 learning correlated positively with 'contextual guessing', 'use of dictionaries for learning purposes', 'note-taking', 'paying attention to word formation', 'contextual encoding', 'intentional activation of new words' and 'verbal repetition', the use of 'visual repetition' was a predictor of smaller vocabulary size. Likewise, Hamzah et al. (2009) explored the relationship between VLSs and vocabulary size on

EFL undergraduates, finding that larger vocabulary size related positively to 'taking notes', 'studying new words many times', 'using English media', 'talking with native speakers', 'studying the word with classmates', and 'using physical action when learning a word'. Similarly, Wadvogel (2013) explored students' VLSs use and its relationship to L2 proficiency, reaching a very interesting conclusion: beginners make more use of determination (such as 'use of dictionaries', and 'paying attention to word-formation') and memory strategies (such as 'note-taking' or 'studying the words many times'), and increase their use of cognitive skills as their L2 improves.

Finally, some studies (Gu, 2002; Intaraprasert, 2000) have related VLSs selection to other variables. Intaraprasert explored students' choices in relation to the factors mentioned above and the type of institution (state vs private universities) where participants studied, concluding that this factor affected VLSs selection. Another example worth mentioning is Gu (2002), who examined EFL undergraduates' selection of VLSs in relation to their academic major (science vs arts majors) and gender. Significant differences were found, with learners taking science major courses selecting more often strategies related to 'memorisation of words', 'word structure' and 'word-formation rules', and arts students making more 'use of vocabulary notes'. Similarly, when examining the relationship between VLSs use and gender, females were found to make a larger use of VLSs. Table 3.6 provides a summary of the main methodological features and findings of the studies presented.

Table 3.6

A summary of the main features of some studies on university L2 learners' selection of VLSs in relation to other variables

Study	Methodological aspects	Findings
Gu & Johnson (1996)	Instrument: *VLSQ* Sample: 850 Chinese undergraduate learners of English. Variables: lexical and proficiency level.	Positive correlation of L2 level and 'contextual guessing', 'use of dictionaries for learning purposes', 'note-taking', 'paying attention to word formation', 'contextual encoding', 'intentional activation of new words' and 'verbal repetition'.
Intaraprasert (2000)	Instruments: Oral interview and written questionnaire. Sample: Thai EFL undergraduate students Variables: gender, proficiency level, and kind of institution	No significant differences in males and females' choice of strategies. L2 proficiency related to greater use of strategies.
Gu (2002)	Instrument: Large-scale survey Sample: Chinese EFL undergraduate learners Variables: gender and academic major	Female made greater use of strategies. Differences between arts and science majors were found.
Jiménez Catalán (2003)	Instrument: *VLSQ* based on Schmitt (1997) Sample: EFL undergraduate students Variables: gender	Males and females differ in the number of strategies used.
Xhaferi & Xhaferi (2008)	Instrument: *VLSQ* Sample: Albanian EFL undergraduate students Variables: gender	Preferred strategies: Asking teachers for meaning, making guesses and making list of new words. No differences between genders.
Hamzah et al. (2009)	Instrument: *VLSQ* (Benet, 2006) Sample: EFL undergraduate students. Variables: vocabulary size	Vocabulary size related to taking notes, studying new words many times, using English media, talking with native speakers, studying the word with classmates, and using physical action when learning a word.
Waldvogel (2011)	Instrument: *VLSQ* based on Schmitt (1997) and Yes/No test Sample: 475 Spanish FL students Variables: vocabulary size	Significant positive relationship between vocabulary size and management of VLSs.

3.3.2. Secondary-school learners' VLSs selection

Although the number of young and very young EFL learners is growing, studies exploring secondary-school learners' selection of strategies are scarce. Different reasons may explain this. First, secondary-school learners are under eighteen, and researchers must request explicit permission from parents or tutors to administer questionnaires or collect any type of data. Second, secondary-school learners are adolescents and are growing cognitively. This means that they are still developing their metalinguistic awareness, i.e., they may not be fully able to reflect on language properties or give explanations about language. Age affects the learning process, and it may also have repercussions on VLS selection, which may explain why researchers usually opt to explore adults' selection of learning strategies.

The little research available on adolescents (Castellano-Risco, 2018; García López, 2000; Schmitt, 1997) shows a clear preference for the use of repetition ('oral repetition' and 'written repetition') and semantic strategies (such a 'creation of word lists' or 'use of a dictionary'). While the usefulness of semantic strategies is not in question, repetition strategies raise some concern. As s put forward by the Depth of Processing Hypothesis (Craik & Lockhart, 1972) and Bloom's taxonomy (1984), the low-level information processing required may not make them useful.

There is a lack of agreement on the use of mnemonic strategies, such as the 'keyword' method, by secondary-school learners. Schmitt (1997) found a considerable use of these strategies, but they were the least used strategies in García López (2000). Given the lack of further research and contextual information, such as learners' L2 level, and considering the differences among the samples, e.g., the participants' L1, providing an explanation for this fact is difficult, as different factors, such as the teachers' roles, learners' L2 mastery, and the diverse L1 writing systems could contribute to this difference.

Finally, to the best of my knowledge, other than my own study (Castellano-Risco, 2018). little research into the exploration of VLS in relation to other variables has been carried out. My study examined secondary-school (age 14-15) EFL learners' selection of VLSs in relation to vocabulary knowledge and learning context (CLIL vs regular EFL). The main findings of this study were that (1) learners who had a larger receptive vocabulary knowledge made greater use of consolidation strategies, (2) the strategies that related most with vocabulary learning implied a larger cognitive effort ('grouping words together to study them', 'analysis of affixes and roots', 'using English media', and 'using new words in a sentence'), and (3) CLIL learners made greater use of the more beneficial strategies. Table 3.7 summarises the main features of the studies here presented.

Table 3.7

A summary of the main features of some studies on secondary-school learners' selection of VLSs

	Study	Methodological aspects	Findings
Identification of strategies	Schmitt (1997)	Instrument: *VLSQ* based on Schmitt's taxonomy	Preferred strategies: 'bilingual dictionary' and 'asking classmates', 'verbal repetition', 'written repetition', 'study of spelling', 'studying the sound of a word' and 'wordlist'.
	García López (2000)	Sample: 139 EFL learners Instrument: Self-observation questionnaire (Levin & Pressley, 1985)	- Preferred VLSs: Repetition strategies and semantic strategies. - Least used strategies: Mnemonics.
VLSs and other variables	Castellano-Risco (2018)	Instrument: *VLSQ* based on Schmitt (1997) and Yes/no test Sample: 44 secondary-school EFL students Variables: vocabulary size and learning context	- Significant impact of the teaching approach. - Greater use of consolidation strategies by more advanced learners. - Greater use of cognitive strategies by CLIL learners.

Source: Own elaboration.

This section has aimed to provide a review of some studies exploring adult and teenage L2 learners' VLSs selection and its relation to other factors. Findings relating the most-widely VLSs used are somewhat inconclusive. While most research has shown a preference on L2 learners, regardless of the age factor, for repetition strategies, other studies (Schmitt, 1997; Waldvogel, 2011) suggest that repetition strategies result in less recall than other VLSs that involve greater cognitive effort, such as 'paying attention to word form' or 'using new words in sentences'. In my view, in the absence of further research, the discrepancies reported here may be related to differences in factors such as the participants' L2 level, exposure to the L2, or L1. In this respect, some research (Schmitt, 1997; Waldvogel, 2011) suggests that L2 learners' VLSs selection evolves hand in hand with their L2 level, with repetition strategies decreasing as learners improve their L2 command.

As for the relationship between VLS use and other factors, most findings can likewise be considered inconclusive, given the short number of studies exploring the different variables and the lack of consistency in the results. Most studies have examined strategy choice in relation to language proficiency and gender, finding that while linguistic proficiency seems to affect their selection, no clear conclusions can be drawn for gender. Nevertheless, other variables relevant to the development of this dissertation, such as the learning context, have also been explored. This variable has been tackled from different perspectives: type of institution (private vs state universities; Intaraprasert, 2000), academic

disciplines studied (arts, social sciences or sciences; Gu, 2002), and learning approach (CLIL vs EFL; Castellano-Risco, 2018). These three studies agree that learning context is a variable that affects the selection of VLSs significantly.

4. Conclusion

The construct of Language Learning Strategies was developed in the 1970s to fulfil teachers' and practitioners' need to understand how L2 learners face the language learning process. This chapter has focussed on the theory underpinning the construct and has been organised around the definition of LLSs and its application to vocabulary learning with the VLSs.

The first part of the chapter dealt with the definition of LLSs. Following the most influential publications in the field, I have identified the three definitional features (self-regulation, consciousness, and contribution to language learning) that, in my view, constitute the core of LLSs. These three features assist in the construction of the framework of analysis of this dissertation. Concerning self-regulation, given that learners freely choose the strategies to use, i.e., they self-regulate their use of LLSs and VLSs, and that this selection seems to be conditioned by various factors, this dissertation will explore whether the language teaching approach affects VLS use. The second element —consciousness— will be used to determine whether a learning process is considered strategic. Only if the learner can report the use of an action as strategic will this action be viewed as a strategy. Finally, the third element —learning potential— will be used to design the VLS use analysis. Strategies are more likely to help L2 learning when taking place in groups. Since one of the objectives of this dissertation is to determine the usefulness of the strategies and their impact on lexical knowledge, the use of VLSs will be analysed not only as individual items but also in groups.

The second part of this chapter has focussed on the concept of VLSs. After a brief discussion on the VLSs definition, a review of the main VLSs taxonomies has been presented, along with a description of the compilation process of the taxonomy to be used in this study (Schmitt, 1997). The last part of this section discussed some studies on VLSs selection and its relationship with some contextual and ID factors. Although research on VLSs is less abundant than that on LLSs, this review allows us to reach two clear conclusions that justify the present study. First, most research in the area is based on adults' VLSs selection, with little attention paid to younger learners. This dissertation can thus contribute to filling the existing gap about how L2 teenage learners process and learn vocabulary. Second, the little research that is available indicates that the learning context has significant influence on VLS use, particularly with respect to CLIL learning. Considering the acceptance of this new teaching approach, the present study aims to shed some light on how CLIL affects cognition and language processing, by looking into CLIL learners' VLSs use.

Thus, this chapter is of great importance for this dissertation, as it establishes the basis for accurate and complete analysis and interpretation of VLSs use. As anticipated throughout this chapter, part of the Chapter Four discussion on VLSs will be devoted to justifying the selection of Schmitt's taxonomy (1997). Moreover, I will return to this Chapter Three in Chapter Six to discuss the results of this study.

PART TWO

THE STUDY

CHAPTER FOUR:
METHODOLOGY

1. Context of Study

1.1. CLIL as a Result of Multilingualism Policies in Europe and Extremadura

In the last 50 years, the world has become significantly more interconnected than ever before. This sense of global citizenship and interconnection has led to the fostering of new kinds of relationships between countries and offered greater ease of movement. As a result, mastering more than one language becomes a necessary skill.

This need was soon identified in the European Union. With 28 countries, 24 official languages and 60 regional and minority languages, this institution has, since its inception, aimed to overcome its characteristic linguistic division to achieve political and economic integration. The first step on this path was setting the official languages of the European Economic Community in 1958. However, it was not until the 1970s that the EU started developing a common language teaching framework with the final aim of promoting multilingualism (for further information, see Marsh, 2002; 2013). As expressed in the European Commission's white paper *Teaching and Learning towards the Learning Society* (1995), the main reasons for promoting multilingualism were primarily related to economics:

> Proficiency in several Community languages has become a precondition if citizens of the European Union are to benefit from occupational and personal opportunities open to them in the border-free Single Market. This language proficiency must be backed up by the ability to adapt to working and living environments characterised by different cultures. Languages are also the key to knowing other people. Proficiency in languages helps to build up the feeling of being European with all its cultural wealth and diversity and of understanding between the citizens of Europe.

> Multilingualism is part and parcel of both European identity/citizenship and the learning
> society.

<div align="right">European Commission's White Paper, 1995, p. 47</div>

This vision of multilingualism was reinforced and expanded in 2002 with the '1+2 objective', intended to achieve that "every European citizen should have meaningful communicative competence in at least two other languages in addition to his or her mother tongue" (European Commission, 2002, p. 4). To achieve this aim, various actions were carried out, being especially relevant the Action Plan 2004-2006 and its subsequent updates: *A New Framework Strategy for Multilingualism* (European Commission, 2006) and *Multilingualism: an Asset for Europe and a Shared Commitment* (2008). While the first action plan set some broad objectives regarding language learning in three main areas —life-long language learning, improving language teaching and creating a language-friendly environment—, the subsequent plans addressed the promotion of multilingualism by recognising the importance of fostering an intercultural dialogue between the different European languages and cultures and focussing on four interrelated elements —education, economy, society, and research.

| **1958:** Official languages in the EEC are set | **1994:** Emergence of the term CLIL | **2002:** Setting of the MT+2 policy | **2006:** *New plan to promote multilingualism* |
| **1975:** Establishment of the Threshold Level | **1995:** European Commission's white paper: *Teaching and Learning towards the Learning Society* | **2004:** Action Plan 2004-2006: *Promoting Language Learning and Linguistic Diversity* | **2008:** *Multilingualism: an Asset for Europe and a Shared Commitment* |

Figure 4.1. European Union's language policy evolution. Source: own elaboration.

The practical implementation of these plans was diverse. Each member state carried out different initiatives, ranging from lowering the onset of foreign language learning to introducing new language teaching approaches. It is within this latter language policy that the development of the CLIL approach is framed. CLIL is an educational approach that promotes the development of content subjects through the use of a foreign language (Coyle et al., 2010; Mehisto et al., 2008). This educational approach was conceived in an effort to (1) reinforce Europe's level of multilingualism (reactive reasons), but, at the same time, (2) improve the foreign language competence in those places where there was a deficit (proactive reasons; Pérez Cañado, 2012). Due to its various advantages, it was even explicitly included in the Commission's Action plan 2004-2006 as one of the most beneficial actions to foster language learning:

> Content and language integrated learning (CLIL), in which pupils learn a subject through
> the medium of a foreign language, has a major contribution to make to the Union's

language learning goals. It can provide effective opportunities for pupils to use their new language skills now, rather than learn them now for use later. It opens doors on languages for a broader range of learners, nurturing self-confidence in young learners and those who have not responded well to formal language instruction in general education. It provides exposure to the language without requiring extra time in the curriculum, which can be of particular interest in vocational settings. The introduction of CLIL approaches into an institution can be facilitated by the presence of trained teachers who are native speakers of the vehicular language.

European Commission, 2003, p. 19

In general, CLIL was well received along with the member states and soon, it was extensively implemented throughout most of them (Lasagabaster & Doiz, 2016). This quick expansion is readily observed when comparing the maps showing the CLIL provision in Europe in the academic years 2004-05 and 2010-11 (Figure 4.2). While in the academic year 2004-05, most CLIL projects were in pilot stages, in the 2010-11, most member states that offered CLIL had systematised this educational practice.
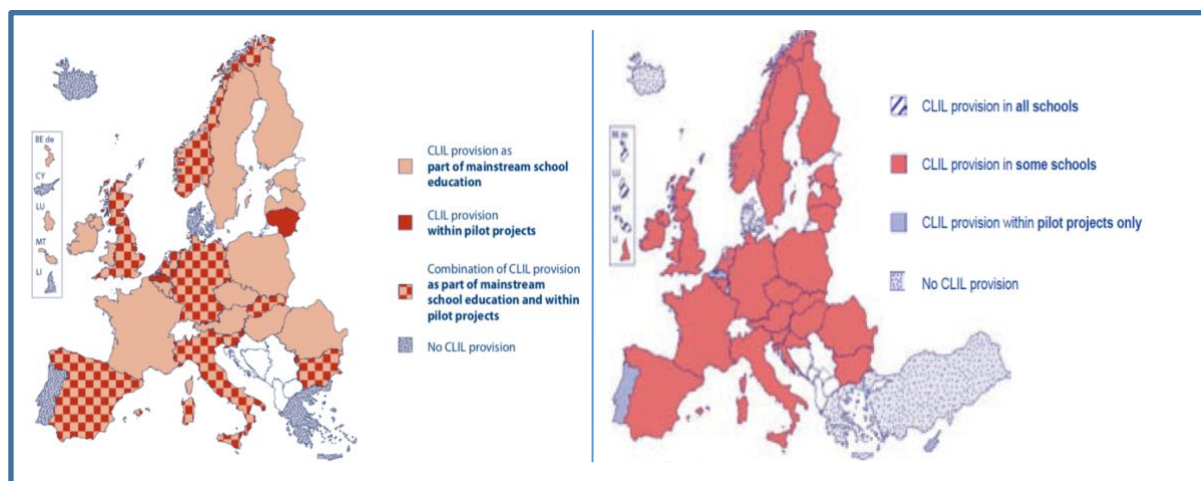
Figure 4.2. Existence of CLIL provision in primary and secondary education in the academic years 2004-05 (on the left) and 2010-11 (on the right). Source: Eurydice report, 2006, 2017.

However, its implementation has not been as straightforward as it seems at first sight. Traditionally, there have been three main areas of concern regarding the implementation of CLIL at schools: the vagueness with which the interaction between language and content approaches is defined, the lack of egalitarianism and the lack of theoretical clarity of the construct.

After the first boom, some sceptical voices arose questioning those presumed benefits as well as the validity of the research conducted on this issue (Bruton, 2011a, 2011b, 2013; Paran, 2013). In this sense, a broad body of research attempted to provide clear evidence of the benefits of implementing such programmes in different learning aspects, such as motivation (Lasagabaster, 2011; Lasagabaster & López Beloqui, 2015), affective factors (Lasagabaster & Doiz, 2017), L2 learning (Agustín-Llach &

Canga Alonso, 2016; Canga Alonso, 2013b; Heras & Lasagabaster, 2015; Lasagabaster, 2008; Merino & Lasagabaster, 2017), or even content knowledge (Fernández-Sanjurjo, Fernández-Costales, & Arias Blanco, 2017; Ouazizi, 2016; Pérez Cañado, 2018; Xanthou, 2010, 2011). The positive findings of these studies also contributed to its warm welcome and implementation.

Furthermore, some other authors (Bruton, 2011b, 2013; Lorenzo, Casal, & Moore, 2009; Ruiz Gómez & Nieto García, 2009) pointed out that the implementation of CLIL attacks equality of opportunities in schools. In their view, more often than not, students are asked to decide whether they want to join CLIL programmes and those with higher socio-economic statuses (SES) are found frequently opting for CLIL (Bruton, 2011b, 2011a, 2013; Lorenzo, 2007). However, the latest research on this issue (Lorenzo et al. 2011; Pérez-Cañado, 2017) suggests that CLIL actually favours egalitarian access to education.

Finally, a lack of theoretical clarity (Bruton, 2011b, 2013) has also been attributed to CLIL. This approach covers a wide range of methodologies, to the extent that it is sometimes seen as an "umbrella term" (Mehisto, Marsh, & Frigols, 2008, p. 12). This flexibility has benefitted the implementation of CLIL programmes adapted to the different languages and realities present in the European landscape (Pérez-Cañado, 2017) but has led to a lack of conceptual clarity (Cenoz et al., 2013, p. 5) which may hamper the CLIL construct. Therefore, great efforts have been made to identify those variables determining the type of CLIL programme to be implemented (see, for example, Rimmer, 2009; Smitt, 2007; Wolff, 2005).

In spite of the criticism, nowadays, CLIL succeeded in Europe and its implementation in the context of this study, the region of Extremadura, is a reaction to the European Commission's strong commitment to multilingualism. Extremadura, a monolingual region located in south-western Spain, on the Portuguese border, can be considered a sparsely populated area, with approximately a million inhabitants disseminated in a large extension of about 41,634 km^2.

The region was proclaimed an Autonomous Community in 1983, and since then, the Regional Authority has gradually been assuming the management of different competences, such as the regional legal system, environmental or health care competences (2002) transferred by the Spanish National Government. In 2000, the Spanish government transferred the Educational competences as well to the regional authorities.

From the beginning, the Extremaduran Educational Authority has displayed a real commitment to the promotion of second languages. As can be seen in Table 4.1, the Extremaduran language policy encompasses six different kinds of actions for promoting L2 learning: (1) the lowering of the age at which learners are introduced to the first foreign language; (2) the introduction of a second foreign language; (3) the implementation of CLIL approaches; (4) the implementation of the European Language Portfolio (ELP); (5) the development of a plan for promoting multilingualism, the *Linguaex*

Plan (which, eventually, was not fully implemented), and (6) the Portuguese Language and Culture programme. In general, a transition from the conception of L2 learning as a synonym for English to a new and broader picture of L2 education as a multilingual opportunity, can be observed. Furthermore, the target population of these initiatives has become broader: while the initial policies were devoted to promoting L2 learning in formal settings with young and very young learners, later initiatives seek to promote L2 education in all sectors of society.

Table 4.1

A summary of the main actions taken to raise foreign languages learning and multilingualism in Extremadura.

Action	Act	Aims
English as a Foreign Language subject since Pre-Primary Education	*Orden de 30 de agosto de 2000* *Orden de 10 de agosto de 2001*	- Lowering the age of onset in the English as a Foreign Language subject (3-5 years of age).
Introduction of a second foreign language in Primary Education	*Orden de 27 de mayo de 2004*	- The incorporation of a second foreign language (French or Portuguese) in the third cycle of Primary Education (10-12 years of age). - Promotion of multilingualism.
The "Bilingual Sections" programme	First regulation: ORDEN de 19 de mayo de 2005	- Promotion of multilingualism. - Improving and supporting L2 learning. - The implementation of CLIL programmes in the region in a balanced way between rural and urban schools.
The European Languages Portfolio (ELP)	*Orden 1 de septiembre de 2008*	- To incorporate to the schools a tool in which L2 learners can register all their languages experiences.
Plan Linguaex (2008-2013)	Plan Linguaex (2008)	- Promotion of multilingualism. - Supporting L2 learning outside the Compulsory Education. - Improving the quality of language teaching. - Promotion of other languages rather than English, with particular emphasis on the Portuguese language.
The Portuguese Language and Culture programme (1988-2013)	*Instrucción 24/2013 de 5 de septiembre de 2013*	- The promotion of the Portuguese language and culture. - The integration of Portuguese and the Portuguese-speaking learners. - Promotion of the appreciation for cultural differences.

Although the different proposals can converge and various initiatives can be carried out simultaneously, below, I will focus on the core issue concerning the interest of this dissertation: the incorporation of CLIL to the Extremaduran educational system through the 'bilingual sections' programme.

1.2. CLIL in Extremadura

In the academic year 2004–2005, the 'Bilingual Sections' programme, based on CLIL principles, was officially regulated. The programme aimed at the experimental implementation of content-based teaching approaches in the region. However, this was not the first attempt to introduce bilingual learning experiences as, since the academic year 1997–98, two primary schools in the region —Luis de Morales in Badajoz and Alba Plata in Cáceres— had joined the British Council–MEC agreement. This agreement sought to integrate both the Spanish and British curricula so that children could learn English while they were supposed to obtain both diplomas. In practice, this meant that the British Council explicitly trained teachers and that 40% of the teaching hours were in English.

The 'Bilingual Sections' programme was a step forward in the implementation of content-based teaching approaches in the region. This programme was intended to implement CLIL practices in the teaching/learning of some disciplinary subjects, which were partially taught through a foreign language (mostly English). One of its key differences, compared to other CLIL programmes, was that CLIL was not implemented throughout school but only in some specific groups (known as 'sections') of each grade. The programme commenced with the setting up of six 'bilingual sections' in both the primary and secondary levels. Since the fifth grade of primary education had been established at the beginning of the year for the programme, the implementation proceeded as follows: first, the programme was implemented in the third cycle of primary education (grades 5 and 6, 10–12 years of age), then the experience was progressively introduced in the second (grades 3 and 4, 8–10 years of age) and first (grades 1 and 2, 6–8 year of age) cycles. The procedure was different in secondary education: the grade of onset was the first grade of secondary education.

In general, the Educational Authorities regulated some key issues to ensure the homogeneity of the different experiences. First, regarding the languages used in the programme, English and French initially, and Portuguese later on, were the only three options given to students. Second, given the sparse population of Extremadura, a balance between the urban and rural areas was aimed at for ensuring that the bilingual sections were equally promoted in both contexts, avoiding population bias. Third, the partnered schools were also encouraged to help students to continue with the project throughout their compulsory educational life. Fourth, the role of the teachers was also coordinated: both content- and language-specialist teachers played a specific role in the programme and all the teachers, regardless of their speciality, were required a B2 level in the foreign language used (Alejo & Piquer-Píriz, 2010).

Finally, the amount of exposure to the FL was also regulated: CLIL learners had to attend subjects in English, either disciplinary or EFL subjects daily.

With this policy as a cornerstone, the CLIL picture in the region has significantly evolved in the last decade, becoming one of the hallmarks of this regional language policies. With a total of 295 bilingual sections in the academic year 2019–2020, the picture of the region with respect to L2 teaching has completely changed. When the programme kicked off it only included initiatives in primary and secondary schools; now, it has expanded to encompass practices in primary, secondary and vocational levels. Furthermore, some features with respect to its implementation have also changed: in the case of Primary Education, now, new projects start in the first grade and are progressively developed in the remaining grades, and the different models of implementation in secondary education (varying in the number of languages used in the bilingual section and subjects taught in the foreign language) were unified in 2015 (Junta de Extremadura, 2015). Finally, a new actor has also come into play: CLIL schools. In the Linguaex plan, it was established that all new state schools had to be 'CLIL schools', i.e., they had to provide CLIL programmes in all the school groups. As a result, currently, six CLIL primary schools are working in the region. Table 4.2 shows the current picture of the CLIL panorama in the region.

Table 4.2

Number of bilingual sections in Extremadura classified into levels, and languages (academic year 2019-2020)

	Pre-Primary	Primary Education			Secondary Education				Vocational Education	A-levels	TOTAL
	English	English	French	Portuguese	English	French	Port.	Mixed	English	English	
By L2	9	147	3	2	110	3	2	2	8	9	295
By level	9	152			117				8	9	

Source: translated from the Extremaduran Educational authorities' webpage.

2. Research questions and variables under study

The purpose of this doctoral dissertation is to explore the impact of language teaching approach (CLIL vs mainstream EFL approach) and Instructed Amount of Exposure (IAoE) on two elements of lexical competence: selection of VLSs and vocabulary knowledge. In line with this objective, three research questions have been posed:

1. Does the implementation of a CLIL approach enhance learners' vocabulary knowledge?
   a. Does it result in significant improvements in the recognition of high-frequency vocabulary?
   b. Does it result in significant improvements in the production of high-frequency vocabulary?
   c. Does it result in significant improvements in the recognition of academic vocabulary?
   d. Does it result in significant improvements in the production of academic vocabulary?
2. Does the implementation of a CLIL approach result in significant changes in the selection of VLSs? If so, are CLIL learners making significantly greater use of those most beneficial strategies?
3. Is IAoE (Instructed Amount of Exposure), among the different changes a CLIL approach implements, a variable that explains results both in VLSs and vocabulary knowledge?
   a. Is the knowledge of the 2K and academic bands determined by IAoE?
   b. Does IAoE affect the selection of VLSs?

Therefore, this study presents a broad objective organised around two independent variables (language teaching approach and IAoE) and their implications for the two elements of lexical competence: vocabulary knowledge and VLSs. Below, the different variables are specified and explained in detail:

- <u>LT approach</u>: this variable refers to the type of language instruction to which learners are exposed in the classroom. In this study, two main approaches are considered: a mainstream EFL approach, which has been the main and almost the only approach for the last three decades in the region (Extremadura, Spain) in which the study was carried out, and the CLIL approach, which started to be officially implemented in some schools in the academic year 2004/05. As stated in previous chapters, these two approaches differ in several aspects, such as objectives, the methodology followed by the teachers, IAoE, contents, or kind of input L2 learners received. Therefore, it is assumed that all these differences would result in different learning experiences and outcomes.

- <u>Instructed Amount of Exposure in English (IAoE)</u>: in SLA, a greater exposure to the L2 is related to more extensive linguistic and vocabulary gains.

- <u>Selection of VLSs</u>: in Chapter Three, VLSs were defined as the conscious actions the students take to learn vocabulary. Considering that learning actions reflect the way the brain processes new items, an analysis of their selection can be used as an indicator of how part of the learning process works. Therefore, the use of VLSs is going to be explored in order to get a better understanding of how secondary-school learners process and learn new L2 items.

- <u>Vocabulary knowledge</u>: every learning experience has as a key objective the acquisition of some kind of content. This study aims to explore the impact of variables related to L2 input on

vocabulary learning, understood, in this context, as learners' receptive and productive knowledge of high-frequency and academic terms.

In short, as shown in Table 4.3, in this doctoral dissertation, the relationship between the four variables is going to be explored. Two of these variables —LT teaching approach and IAoE— are taken as independent variables, and their effects on the dependent variables (selection of VLSs and vocabulary knowledge) are examined.

Table 4.3

Independent and dependent variables

Independent variables	Dependent variables
LT approach	Selection of vocabulary learning strategies
IAoE	Vocabulary knowledge

3. Participants

The sample was composed of 138 secondary-education students in their third grade (aged 14–15). Participants were from four different secondary state schools in Badajoz, the largest town in Extremadura with a population of around 150,000 inhabitants. Some criteria pertaining to the schools and learners' features were established to select the school and type of participants taking part in the project. In Badajoz, there are a total of 24 state high schools. Faced with the impossibility of collecting data from all of them, the following selection criteria were established. Schools had to be urban state schools with at least one English bilingual group per year. Furthermore, the CLIL experience had to be a consolidated practice in the school, i.e., the number of years of CLIL practice should be longer than five. Finally, to avoid any risk of bias, schools were chosen considering the SES of the learners, preventing the selection of schools with different SES levels. After this pre-selection, four schools were invited to participate —three of which were in the same area; the fourth was in the town centre, but it showed similarities regarding the SES of the students and the kind of linguistic programmes and initiatives developed at the school.

With respect to the age of the participants, students from the first and second grades were not chosen mainly for two reasons. First, certain instruments of this study required some level of self-knowledge, and it could not be ensured that young learners were able to reflect on their own learning process in such a profound manner. Second, some students had joined CLIL programmes in secondary school and —considering some of the objectives of this study, i.e., to analyse the effects of CLIL in relation to IAoE— the most suitable participants were those learners who had been enrolled in CLIL programmes for at least two academic years.

Table 4.4 summarises the participants' features according to school, gender, and whether they study a third language or the approach they use to learn English.

Table 4.4

Distribution of the sample by schools

	N	Tuition	N	Gender		Second FL		
				Male	Female	French	Portuguese	Others
School 1	47	CLIL	36	17	19	3	8	-
		EFL	11	6	5	-	4	-
School 2	31	CLIL	15	7	8	34	6	-
		EFL	16	8	8	4	6	-
School 3	30	CLIL	10	4	6	2	4	-
		EFL	20	9	11	20	-	-
School 4	30	CLIL	21	10	11	6	8	-
		EFL	9	5	4	1	5	Galician

Considering the language teaching approach participants followed, two groups of learners could be drawn. On the one hand, 56 participants learnt English in a regular EFL programme. They took English lessons mainly devoted to the development of general English skills four times a week. On the other hand, 82 participants took part in CLIL experiences. They attended EFL lessons five times a week and some content subjects which were delivered in English. As stated in the previous chapters, both language teaching approaches (CLIL vs regular EFL) are intrinsically different, and these differences may have an impact on learners. The following sections undertake the profiles of learners according to the programme they are enrolled in.

3.1. CLIL learners

The CLIL group was made up of 82 learners (average age: 14.2 years old at T1), among which 38 were males and 44 females. It was a heterogeneous group with evident differences with respect to their language learning experience.

First, participants had had different language learning experiences. Thus, due to the regional language policies in place at the time, all participants had started attending EFL lessons since their pre-school education, but they had joined CLIL programmes at different ages. Twenty-three learners had started their experience with a content-based language learning programme in the first grade of primary education, under the coordination of the British Council-MEC agreement. In contrast, 25 had commenced the regional CLIL experience in the fourth, fifth or sixth grade of primary education and 34 participants had joined CLIL programmes in secondary education.

Moreover, depending on their school and time of joining the CLIL programme, participants had attended different subjects in which English was used as the medium of instruction. As shown in Figure 4.3, most participants who had started CLIL at the primary education level had been taught Science and Arts and Crafts partly in English. However, the responses become more diverse when CLIL was introduced at the secondary level in terms of the number of subjects taught, with most participants attending subjects such as Natural Science, Music, Technology, and Geography and History partially in English (see Figure 4.3). In the end, despite this freedom of choice, it seems that most schools offered the same CLIL subjects; hence, most participants had been enrolled in similar subjects when taught through English.
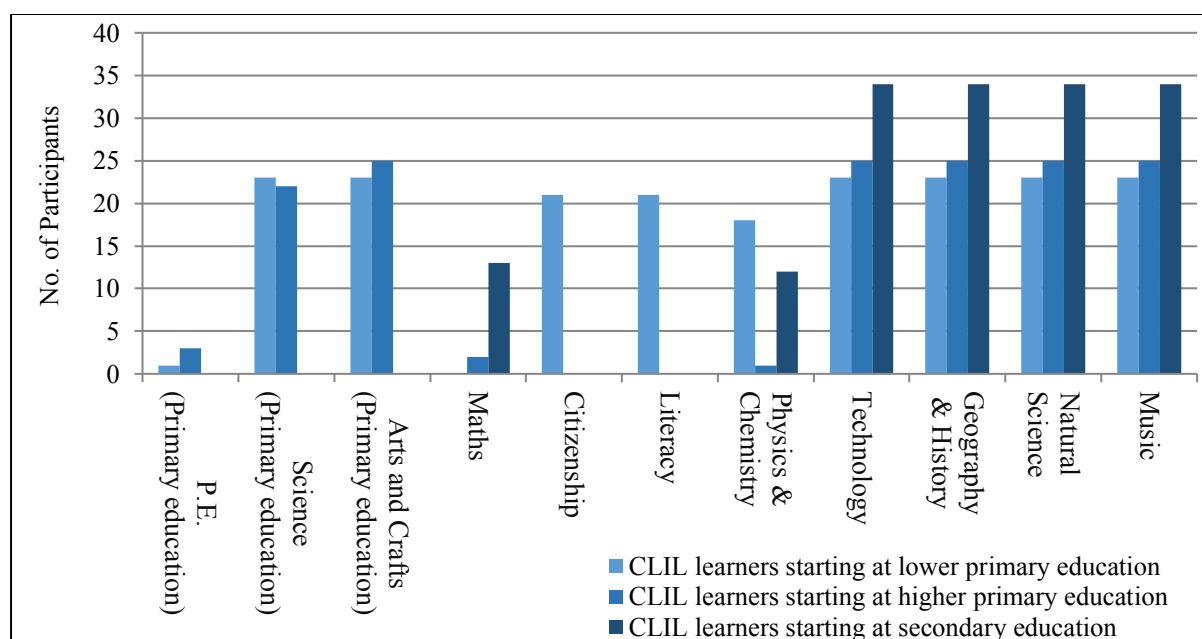
Figure 4.3. Subjects learnt through English.

Derived from their different ages of onset, participants have had different instructed amounts of exposure (IAoEs). That is, those learners starting the CLIL programme in 1st grade of primary education have had a more extensive exposure that those who joined the CLIL programmes in secondary education. In order to estimate the IAoE CLIL learners have received, participants are classified into three groups, considering the age at which they joined the programme. To calculate the IAoE, the regional curricula, officially regulated (Junta de Extremadura, 2007a, 2007b, 2015b) were checked, resulting in the following three groups:

- CLIL 1 ('early CLIL learners'): this group consisted of 23 learners who had participated in CLIL experiences from the first grade of primary education. These learners were enrolled in the specific programme developed by the British Council in cooperation with the Spanish Ministry of Education and Science, in which both the Spanish and British curricula were integrated. In primary education, they learnt Social and Natural Science, Arts, and Literacy through English. In secondary education, they joined standard CLIL secondary schools, in

which between 2 and 3 disciplinary subjects were taught in English. The subjects ranged from Geography and History to Biology, Technology, Music, or Physical Education (PE). Additionally, learners had an extra hour of EFL per week. In total, these learners had been approximately exposed to 3,000 hours of English at T1 —1,300 of EFL tuition and 1,700 hours of CLIL content subjects. At T2, the amount of exposure accounted for 3,315 hours.

- CLIL 2 ('standard CLIL learners'): this group comprised 25 learners who joined a CLIL programme in the fourth, fifth or sixth grade of Primary Education. The subjects they learnt through English in primary school were typically Natural and Social Sciences, and Arts and Crafts, although some of the participants reported having attended PE lessons in English. In secondary education, the subjects varied depending on the school and level. These learners had been exposed to approximately 2,400 hours of English at T1, a figure that increases up to 2,715 hours at T2.

- CLIL 3 ('late CLIL learners'): this was the group with the highest number of participants, 34. The learners had started CLIL at the beginning of secondary education, which means that their only input in primary education consisted of EFL classes. The disciplinary subjects learnt through English varied every year as in the previous groups. They had received at T1 an approximate amount of input of 2,000 hours in total —1,300 of EFL tuition and 700 of CLIL content subjects. At T2, this figure increases up to 2,315 hours.

For this dissertation, it is also relevant to specify the number of learners from each school that makes up each group, as it may affect the interpretation of some of the results. This information is summarised in Figure 4.4.
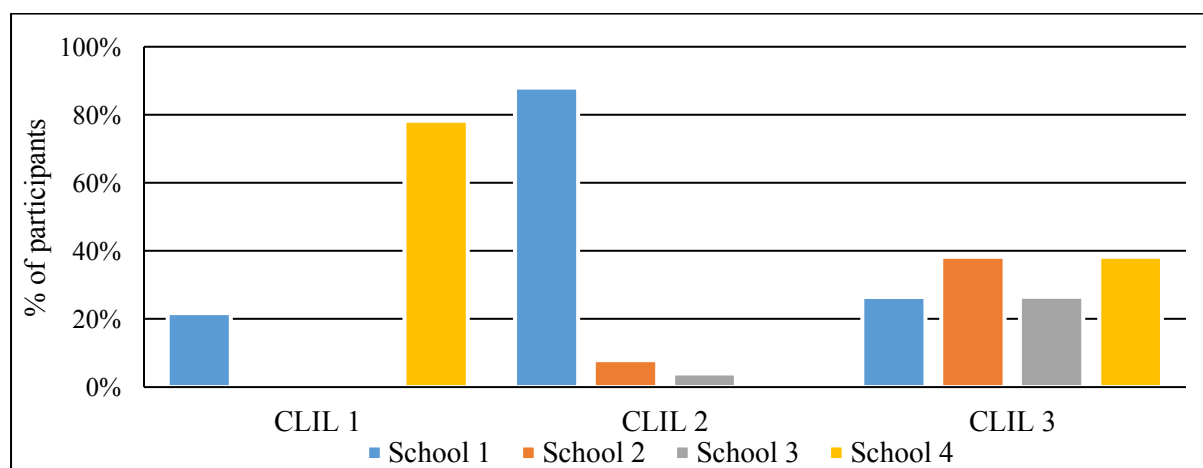
Figure 4.4. Percentage of participants from the different schools making up each of the CLIL subgroups.

Finally, concerning their attitude towards FL learning, in general, the respondents seem to have a positive attitude towards language learning based on their interest in studying some subjects in CLIL. However, there are more signs of this interest in foreign languages. Forty-one students had also learnt a second foreign language at school: 15 learners had attended FFL (French as a Foreign Language) and

26 studied PFL (Portuguese as a Foreign Language) at school. Moreover, 55% of the CLIL learners participated in extra-curricular activities in which English was either the learning aim or the language of communication. These activities included courses in the *Escuela Oficial de Idiomas* (Official School of Languages), summer camps, and private English lessons.

In short, the prototypical CLIL participant was a 14-year-old secondary-school learner who had been receiving EFL lessons since they were three years old and had joined the CLIL programme in high school. They were taught several academic content subjects in English, such as Science, Technology, or Music with EFL lessons five times a week. Moreover, they attended extra-curricular English language activities weekly. In sum, these participants were exposed to at least nine hours of English weekly.

3.2. Mainstream EFL learners

This group was made up of 56 learners (average age: 14.4 y.o. at T1) who had only attended EFL lessons since pre-primary level. They belonged to the four schools explored, with most informants (36%) coming from school 3, followed by 28% from school 2 and 20% who attended lessons at school 1. Finally, a minority of these learners (16%) were from school 4.

These participants had had EFL classes since pre-school; therefore, with the information derived from the Official Syllabus of Extremadura (Junta de Extremadura, 2007b, 2007a, 2016), they had been exposed to, approximately, 1,200 hours of English at T1, while, at T2, this figure reached 1,332 hours.

Their exposure to foreign languages were not restricted to English lessons. In fact, 71.42% of them took an additional language as an optional subject, possibly an indication of a positive attitude towards FL learning. Specifically, 25 learners had FFL as a subject, whereas 15 attended PFL lessons at school. Additionally, 37.5% of them were exposed to English outside the school, attending extra-curricular activities such as courses at the Official School of Languages, language academies, or private lessons.

In conclusion, the prototypical mainstream EFL learner was a 14-year-old Spanish native speaker who, apart from English, learns other foreign languages at school (Portuguese or French) showing their understanding of the importance of learning an L2.

In this section, participants have been profiled through a description of their mother tongue, age, gender and relationship with foreign languages. According to the level of exposure to English, four main groups can be drawn. Table 4.5 summarizes the main features of each group. It includes information about the number of students, gender, whether they studied a second foreign language and an estimation of the IAoE to English at T1 and T2.

Table 4.5

A summary of the groups' main features

	N	Gender		Second FL			IAoE at T1	IAoE at T2
		Male	Female	French	Portuguese	Other		
EFL	56	26	30	25	15	3	1,200	1,332
Early CLIL	23	14	9	6	8	-	3,000	3,315
Standard CLIL	25	14	11	3	6	-	2,400	2,715
Late CLIL	34	16	18	6	12	-	2,000	2,315

4. Instruments

A total of four different instruments were used. They include: a language history questionnaire (adapted from Li, Sepanski & Zhao, 2006) to record participants' demographic data, the Vocabulary Levels Tests (VLT; Schmitt, Schmitt & Clapham, 2001) and the Productive Vocabulary Levels Test (PVLT; Laufer & Nation, 1999) for measuring receptive and productive vocabulary knowledge respectively, and a VLSs questionnaire adapted from Schmitt (1997). Detailed information about each of the instruments is provided in the following sub-sections.

4.1. The language history questionnaire adapted from Li, Sepanski & Zhao (2006)

In order to obtain general information about learners, the language history questionnaire (Li, Sepanski, & Zhao, 2006) was adapted (see Appendix B). It included questions related to (1) participants' initials, age, and gender; (2) parent's mother tongue; (3) languages learnt at school and additional information about when they started to learn them, the type of language approach followed (CLIL or mainstream EFL subject), and, if relevant, the subjects learnt in the CLIL programme and (4) languages they learnt in extra-curricular activities. The questionnaire was written in Spanish in order to facilitate understanding and participants had to answer it before taking the rest of the tests.

4.2. The Vocabulary Levels Test (Schmitt et al. 2001)

In the literature, several tests measuring receptive vocabulary knowledge are available, differing in terms of the target audience, test format, level measured, or the procedure of application (see Table 4.6). Therefore, considering the objective of this PhD dissertation —exploring Spanish secondary school L2 learners' receptive and productive knowledge of high-frequency (2K) and academic vocabulary— it was important to choose an instrument suitable for the features of the sample. Participants of this study were in their teens and had been learning English for ten academic years. Considering the characteristics of my sample, the Peabody test was not an efficient option, as it is not

suitable for teenagers. Within the form-recognition tests, the checklist test would not serve the purposes of measuring specific vocabulary bands either, as the test does not specify them. Among the two remaining options (VLT and VST), VLT (Nation, 1983; Schmitt, Schmitt & Clapham, 2001; Webb, Sasao & Balance, 2017) seemed a better choice since it includes an academic band and allows the administration of isolated bands, which facilitates its use with teenagers. Moreover, it has the advantage of having been used worldwide with different kinds of samples, among which, teenagers are included. In particular, this set of tests is the most widely used instrument for measuring the CLIL lexical knowledge in the Spanish context. Therefore, the use of this instrument would allow comparisons to other samples with similar characteristics.

Developed by Nation (1983) and refined by Schmitt, Schmitt, and Clapham (2001), the VLT is the most widely used vocabulary knowledge test (Schmitt, 2010). It is a form-recognition matching test that focusses on vocabulary at five levels. Four of them are based on frequency analyses and correspond to the number of word families considered sufficient to engage in daily conversation (2,000); to enable initial access to authentic reading (3,000); to enable independent reading (5,000); and to enable advanced usage in most cases (10,000). The fifth level focussed exclusively on academic vocabulary that measures the recognition of the words contained in the AWL (Coxhead, 2000).

It was conceived as a diagnostic test in which test-takers are asked to match the definitions with their corresponding words. In the first version (Nation, 1983), 18 words in six clusters were presented per level, each of which contained three definitions and six options. In the second version (Schmitt et al., 2001), the number of words increased to 30 per level.

The design of the test minimizes guessing and ensures the reliability of the results. All the words —those tested, and the ones presented in the definitions—, belong to the level tested or to lower levels. Moreover, in each cluster, the different words belong to the same part of speech. Finally, in each level, the same distribution of part of speech is presented, following a 3 (noun): 2 (verb): 1 (adjective) ratio (Schmitt, Schmitt & Clapham, 2001).

```
1   business
2   clock          _____  part of a house
3   horse          _____  animal with four legs
4   pencil         _____  something used for writing
5   shoe
6   wall
```

Figure 4.5. Example provided in the VLT test. Source: Schmitt et al. 2001.

As for the main drawbacks of this instrument, most criticism has honed in on the (1) frequency list on which the tests are based on and in (2) the lexical dimension the test actually measures. As for the first concern, frequencies were established based on Thorndike and Lorge's frequency lists (1944), Kučera and Francis' list (1967), and the General Service list (West, 1953), whereas the academic words were taken from the AWL (Coxhead, 2000). Some researchers (see Schmitt, 2010 for an overview)

argue that these lists may be slightly outdated. Moreover, some authors (Webb & Sasao, 2013) state that due to the importance of the 1K band in lexical recognition, a specific level measuring this band should have been included. Regarding the discussion on the actual measurement unit of the test, although it is used as a tool for measuring vocabulary size, it is not actually a vocabulary size test in the narrowest sense of the concept, as it does not provide a global estimation of L2 learners' size, but merely offers an estimation of the knowledge of particular bands (Kremmel & Schmitt, 2018; Webb et al., 2017).

Table 4.6

A summary of the main advantages and drawbacks of the tests explored

Test	Description	type[3]	Bands	Objective	Advantages	Disadvantages
Peabody test (Dunn, 1959)	Pictures in groups of four are displayed and test-takers have to listen to different words and point to the picture that represents it best.	F-R	Not specified	Diagnosis	- Easy to apply with very young learners.	- Not suitable for teenagers/adults
Checklist test (Meara, 1992)	A set of five tests that measures knowledge from the 2K band to the 10K band. Test-takers have to read a list of lexical items and indicate the item they recognise	M-R	Not specified	Not specified	- Easy and quick to take in class - Many items can be measured. - Straightforward and automatic scoring.	- No direct demonstration of knowledge. - Possibility of overestimation. - Bands not specified.
VLT (Schmitt et al., 2001)	Form-recognition matching test that focusses on vocabulary at five levels (2K-10K + academic).	M-R	2K, 3K, 5K, 10K and academic bands	Diagnosis	- Short definitions. - Designed to tap into the initial stages of form-meaning links. - Clusters designed to minimise aids. - Academic vocabulary band.	- Not designed for providing an estimate person' s overall vocabulary size.
VST (Nation & Beglar, 2007)	Multiple-choice meaning-recognition format test that is aimed at producing an overall vocabulary knowledge profile.	M R	14 first bands// 20 first bands	Measuring the overall vocabulary size	- It intends to measure the overall vocabulary size.	- Too long.

[3] M-R= meaning recognition; F-R= form-recognition

4.3. The Productive Vocabulary Levels Test (PVLT; Nation & Laufer, 1999)

In contrast to what happens in receptive vocabulary knowledge, the productive versions of the tests are less frequent in the literature. The primary reason is that the productive ability of learners can be measured using open essays, in which learners have to write about a given topic. This way of measuring vocabulary offers the advantage of collecting a large amount of data through a particular tool: the production of a text. However, it also presents some drawbacks: the topic may bias the findings, learners may be more used to specific kinds of topics, and it can be really time consuming for both test-takers and researchers. For this reason, researchers developed a number of tests for measuring productive vocabulary knowledge.

For the purposes of this study, different options were considered (see Table 4.6 below). For instance, the first proposal consisted of assigning participants a written task on a given topic; however, this option was discarded, primarily due to time constraints. Similarly, the possibility of using the Lex30 was explored, but this instrument did not provide a source for measuring academic knowledge and focussed on associations rather than on form-production knowledge. Finally, the PVLT was chosen for the following three main reasons: (1) it measured the same bands as the VLT; so, it allows measuring the academic band; (2) it had already been used with secondary-school learners (Canga Alonso & Arribas García, 2014; Moreno Espinosa, 2010) and (3) it is not very time-consuming, favouring the administration of the tests with secondary-school learners.

Table 4.7

A summary of the main productive vocabulary size measurement methods

	Bands	Advantages	Disadvantages
PVLT (Laufer & Nation, 1999)	2K, 5K, 10K and Academic Band	It provides information about specific bands. It is easy to implement and correct. It includes an academic band.	It does not measure real productive vocabulary but prompted one. There are varying degrees of difficulty in the prompts.
Lex30 (Meara & Fitzpatrick, 2000)	Not specified	It measures real production.	It is not highly reliable with CLIL learners (Alejo & Piquer-Píriz, 2016).
P_Lex (Meara & Bell, 2001)	Not specified	It measures real production.	It does not include an academic vocabulary.

The PVLT was developed by Laufer and Nation (1999) and is a reliable instrument to measure productive vocabulary knowledge. This battery of tests was designed by making use of the receptive

VLT version. For this reason, this test presents a similar structure to the first version of the VLT (Nation, 1983). It contains the same five levels, i.e., four frequency-based levels (2K, 3K, 5K, and 10K) and an academic band level, which, in this case, is based on the UWL (Xue & Nation, 1984). It is considered a completion test, in which there are five levels with 18 sentences each. Each sentence contains an incomplete word, and test-takers are asked to complete the word, considering the context provided with the sentence. In the following figure, an example of the 2K band is provided.

1. Every working person must pay income t_____.
2. The differences were so sl_____ that they went unnoticed.
3. There are a doz_____ eggs in the basket.
4. The telegram was deli_____ two hours after it had been sent.

Figure 4.6. Productive Vocabulary Level Tests. Source: Laufer & Nation (1999).

Two main problems are usually indicated when analysing the reliability of this test. First, the test provides a variable number of initial letters, which may result in different difficulties depending on the number of letters of the lexical item presented. To address these arguments, Laufer and Nation (1999) carried out a validation study in which they showed that higher-grade test-takers obtained better scores than examinees at lower levels.

The second concern pertains to the issue of what the test really measures. Authors such as Read (2000) and Schmitt (2010) suggest that these tests present some features that may result in the reconsideration of these instruments as form-recall tests: learners are asked to recall some words, but they are not actually producing any lexical item. From their point of view, the PVLT does not measure the vocabulary freely produced by the test-takers, but the knowledge of specific vocabulary.

4.4. Vocabulary Learning Strategies Questionnaire (VLSQ)

Finally, given the lack of instruments available for examining the selection of VLSs adapted to adolescents, a questionnaire was developed. To this aim, the first step was to carry out a literature review to select the most appropriate VLSs taxonomy to be adapted.

4.4.1. Questionnaire development

After a thorough theoretical analysis of the different proposals, Schmitt's taxonomy (1997) was chosen for two reasons. First, it was one of the most widespread used taxonomies, and it was compiled using secondary-school learners as part of the sample. This homogeneity of the sample was regarded as an advantage, as learners of the same age usually have similar cognitive maturity and metalinguistic awareness. In my view, this taxonomy reflects the strategies that teenagers use to learn a foreign language in a more accurate way than taxonomies based on university students' performance. Second, each strategy was clearly defined, and the way VLSs were clustered was justified. This kind of information was crucial for a good understanding of the taxonomy.

The use of this taxonomy also presented some shortcomings. Schmitt's taxonomy was made up of 58 strategies. A questionnaire with a high number of items was difficult to implement with secondary-school learners, considering their attention span. Moreover, there were some strategies to which learners may not be familiar, especially the ones related to specific teaching methods such as the LOCI or the PEG methods. Thus, an adaptation of the taxonomy was needed, and some criteria were established to select the items that would be finally included in the questionnaire: first, it was important for the proportion of items in each category to remain unchanged. In order maintain it, the strategies in each category were counted and the intended total number of items was established. After that, the new number of items per category was calculated from the following formula:

$$\frac{No.\,of\ items\ of\ this\ category}{58} = \frac{X}{21}$$

The total number of items per group in Schmitt's taxonomy and in the new questionnaire can be seen in following Table 4.8:

Table 4.8

Items per category in Schmitt's taxonomy and in the new questionnaire.

	Schmitt's taxonomy	VLSQ
Determination Strategies	9	4
Social for discovering meanings	5	3
Social for consolidating meanings	3	1
Memory strategies	27	7
Cognitive strategies	9	4
Metacognitive strategies	5	3

Second, some VLSs were directly discarded. These strategies were based on specific methods such as the PEG, the LOCI or the KEY word method that were completely unknown to students. Finally, based on the analysis of previous studies, those strategies that had demonstrated greater use were included in this study (García López, 2000; Gu & Johnson, 1996; Lawson & Hogben, 1996; Schmitt, 1997).

The resulting questionnaire was piloted to ensure its suitability for students of this age. Some of the questions had to be reformulated. As a result, the final questionnaire (see Appendix E) was made up of twenty-one strategies. Test-takers had to mark their use of each strategy on a Likert scale from 1 (never) to 4 (always). Table 4.9 shows the selected VLSs.

Table 4.9

Items included in the questionnaire

Group	Sub-group	Strategies
Discovery strategies	Determination strategies	Analysis of the part of speech
		Analysis of affixes and roots
		Check for L1 cognates
		Analysis of any available picture or gesture
		Using a bilingual dictionary
	Social strategies	Asking teacher for an L1 translation
		Asking teacher for paraphrase or a synonym of a new word
		Asking students for meaning
Consolidation strategies	Social strategies	Studying and practice meaning in group
	Memory strategies	Studying word with a pictorial representation of its meaning
		Connecting word to a personal experience
		Connecting the word to its synonyms and antonyms
		Using a new word in a sentence
		Grouping words together to study them
		Using physical action when learning a word
	Cognitive strategies	Verbal repetition
		Written repetition
		Word lists
	Metacognitive strategies	Using English-language media
		Skipping or passing on a new word
		Continuing to study a word over time

Source: Castellano-Risco, 2018.

4.4.2. Construct validity

The grouping and the constituents of each group were a key issue in this study. Ehrman, Leaver and Oxford (2003) and, more recently, Cohen (2014) and Oxford (2017) claim that the success of VLSs is not related to frequency of use, but to how each strategy is combined with other strategies. It is in this way that the analysis of the subgroups of strategies was needed. At the time this taxonomy was developed (nearly twenty years ago), the inclusion of inferential and statistical analysis for the development of taxonomies was not a common procedure. In fact, when Schmitt (1997) proposed the six groups (namely, determination, social strategies for discovering meanings, social strategies for consolidating meanings, cognitive, memory and metacognitive groups) his suggestion was based on a theoretical examination of the strategies and on his own perceptions and intuitions. As Tseng et al.

suggest, the "items are written in a behavioural fashion" (2006, p. 84). Besides, in this dissertation, the taxonomy was adapted to reduce the number of items.

For these reasons, an exploration of the internal coherence of each group was needed, and a Cronbach's coefficient alpha analysis was carried out. This coefficient is a measure of the internal reliability that ranges from 0 to 1.0. The greater the value, the greater internal coherence. It is generally accepted that values from 0.7 onwards show internal consistency. The analysis was carried out to explore the groups of strategies in general, and then, to explore the particular groups proposed by Schmitt (1997). As for the whole group of strategies, results show a general Cronbach's coefficient of 0.69. This was quite near to the accepted value, and therefore, it could be accepted that the data presents internal coherence. Regarding the different sub-groups presented by Schmitt, results did not show any internal coherence within each group. Table 4.10 shows the results of Cronbach's alpha coefficient.

Table 4.10

Cronbach's alpha coefficients

	Group strategy	Cronbach's alpha coefficient
Discovery	Determination strategies	0.27
	Social strategies (Discovery)	0.88
Consolidation	Cognitive	0.50
	Memory	0.65
	Metacognitive	0.07
	Social (Consolidation)	-

Results seemed to point to a lack of internal coherence in all the groups explored. In light of the results, the present classification might not work with the current sample. For this reason, I attempted to classify the strategies looking for a statistical relationship between them.

A factor analysis was carried out in order to seek the underlying structure of the whole set of VLSs examined. The factor analysis is a mathematically complex procedure that reduces a correlation matrix containing many variables to much smaller number of factors (Howitt & Cramer, 2000; Velicer & Jackson, 1990). Authors such as Cohen and Manion (1994) or Howitt and Crammer (2000) suggest its use in exploratory research where the researcher aims to impose an orderly simplification on a number of interrelated measures.

There are three main decisions to make when carrying out a factor analysis: selection of the factor analysis extraction method, factor loading value of discrimination in each factor, and rotation method. As regards the factor analysis extraction methods, there are a number of methods to choose from, such

as generalised least squares, maximum likelihood, principal axis factoring or image factoring. Following the suggestions made by Fabrigar, Wegener MacCallum and Strahan (1999), the maximum likelihood method was chosen, because, in their own words, the maximum likelihood method "[…] permits statistical significance testing of factor loadings and correlations among factors and the computation of confidence intervals" (p. 277). The factor loadings indicate the degree of relationship between the factors and the variables explored (Bachman, 1990; Seliger & Shohamy, 1989). They range between -1.00 and +1.00, following the same rules as the correlation coefficients (Howitt & Cramer, 2000). Thus, the greater the value, the more importance this factor has in the component. For this reason, factor loadings with absolute values below 0.30 were not considered to be part of the corresponding factor because they were too weak to be considered relevant, and the VLS with highest loading was used to define the factor. Finally, as regards the factor rotation method, it aims to simplify the data structure in order to facilitate the understanding. Yaremko, Harari, Harrison, and Lynn (1986) define it as follows: "In factor or principal-components analysis, rotation of the factor axes (dimensions) identified in the initial extraction of factors, in order to obtain simple and interpretable factors" (p. 78). There are a number of methods, such as Varimax, the Direct Oblimin Method, the Quartimax Method, the Equamax Method or the Promax Rotation. In this study, the Varimax method is used, as this orthogonal rotation method minimises the number of variables that have high loadings on each factor, simplifying the interpretation of the factors. Table 4.11 presents the factor analysis results, specifying the factor loading of each strategy.

Table 4.11

Factor analysis results

Group	Explanation	Strategies	Loading
Lexical analysis strategies	Renamed as 'lexical analysis' group, it encompasses all those strategies were closely related to the lexical acquisition.	Analysis of the part of speech	0.639
		Analysis of affixes and roots	0.693
		Using new words in a sentence	0.375
		Grouping words together to study them	0.421
		Connecting the word to its synonyms	0.443
		Using English-Language media	0.334
Mental imagery	These strategies concerned vocabulary learning through linking of meaning to concrete things such as pictures or personal experiences	Studying the word with pictorial representation	0.989
		Analysis of pictures and gestures	0.561
		Connecting word to a personal experience	0.303
Repetition	Strategies implied repetition actions in any form: written or spoken.	Saying a new word aloud when studying	0.573
		Written repetition	0.847
Linking	Strategies involved the creation of links with other words.	Word lists	0.865
		Using a bilingual dictionary	0.317
Kinaesthetic	This group included strategies that has a kinaesthetic component.	Using physical action when learning a word	0.992
Guessing from context	This group encompasses strategies that require context information to understand the meaning	Skipping or passing on new words	0.350
		Checking for L1 cognates	0.787
Social strategies involving interaction with teachers	This group presents a close link to the social strategies, but it only focussed on the teachers' role.	Asking teachers for an L1 translation	0.643
		Asking teachers for paraphrasing or for a synonym	0.440
Social strategies involving interaction with students	This group is related to the understanding of new words aided by other learners.	Asking other students for meaning	0.727

At first, two strategies did not seem to fit properly within their corresponding groups: 'checking in the bilingual dictionary' and 'skipping or passing new words'. As for the former, its inclusion was explained by the fact that students were asked to write down the vocabulary they looked up in the dictionary. Concerning the inclusion of the latter strategy within the 'guessing from context' category, it is the way it was expressed in the questionnaire which allowed to consider it as part of this group, as it is emphasised that the word is skipped when learners understand the gist of the text.

Besides, in the factor analysis, other two problems arose. First, some strategies could be included in more than one category. This may be explained due to the multi-faceted nature of VLSs. Schmitt suggests that "some strategies could easily fit into two or more groups, making their classification difficult" (1997, p. 8). If a strategy seemed to fit in two groups, it was included in the group in which it presented a greater factor loading.

Secondly, only nineteen out of the twenty-one strategies were included. The maximum likelihood method implies some previous analysis of the data, such as the analysis of the communalities. Such analysis indicates the variance shared by factors with given variables. If the values are small, it means that the variable do not fit well with the factor solution, so the program drops it from the analysis. This was the case of two strategies: 'studying and practicing the meaning in group' and 'continuing to study the word over time'. Thus, from here on, when exploring the use of VLSs, only the resulting nineteen strategies will be explored.

5. Data collection

The present PhD study started in November 2015. It is based on a preliminary study carried out as a MA dissertation and which main outcomes can be found in Castellano-Risco (2018). Concerning the data collection period, the battery of tests was administered at two times: T1, February-March 2016, and T2, February 2017. For each period, data collection involved two class sessions of fifty-five minutes, which were not planned for the same day, or for the same week, but for two consecutive weeks. As shown in Table 4.12, at both times, different tests were administered.

During the first T1session, the VLSQ was administered together with the VLT 2K version and the language history questionnaire. As regards the VLSQ, fifteen minutes were considered enough for students to mark the use of each strategy in a Likert scale from 1 (never) to 4 (always). Instructions were given in Spanish to facilitate understanding and test-takers were reminded the anonymity of the questionnaires. In the second session, the academic version of the VLT was administered. At T2, new tests were introduced, so the battery of test was larger, and more time was required. Together with the VLT, the VLSs questionnaire and the language history questionnaire, the 2K and academic versions of the PVLT were administered. Therefore, whereas in the first session, the VLSQ, the VLT of the 2K

version and the language history questionnaire were administered, during the second one, the academic band of the VLT and the 2K and academic bands of the PVLT were administered.

Table 4.12

Time distribution of the tests

	Day 1	Day 2
T1	- VLSQ. - 2K version of VLT. - Language history questionnaire.	- Academic band of the VLT.
T2	- VLSQ. - 2K version of VLT - Language history questionnaire.	- Academic band of the VLT. - 2K and academic bands of PVLT.

6. Data treatment

In the scoring of the VLT and PVLT, some decisions had to be taken: in both tests, if an item was not answered, it was considered incorrect. In addition, other criteria had to be followed in the PVLT. Thus, as the aim was to explore their lexical knowledge, a lexical term was correct, even if the appropriate tense or number were not used, it was considered correct. However, when the appropriate part of speech was not used, a verb was required and a noun was used instead, it was considered incorrect. Finally, if there was any spelling mistakes, it also invalidated the answer.

The results from the VLT and the PVLT were expressed in the following way: (1) the overall mean score, expressed in raw data and as an estimation of the number of words participants seem to know, (2) the frequency distribution of the scores, and (3) the percentage of participants mastering the band explored. To estimate the number of words known by the learners, Nation's formula (1990) was applied: "N correct answers multiplied by total N words in dictionary (the relevant word list) divided by N items in test" (p. 78). Similarly, to determine when a band is mastered, Schmitt et al.'s (2001) specifications were followed. In their view, a vocabulary band has been acquired when the test-taker hits, at least, twenty-six out of the thirty words.

Intaraprasert's (2000, p. 167) scale was used to interpret the VLSQ data. According to this scale, the mean frequency score of strategy use of each item ranging from 1.00 to 1.99 is established as 'low use', from 2.00 to 2.99 as 'medium use', and from 3.00 to 4.00 as 'high use'. Figure 4.7 below shows the applied measures.
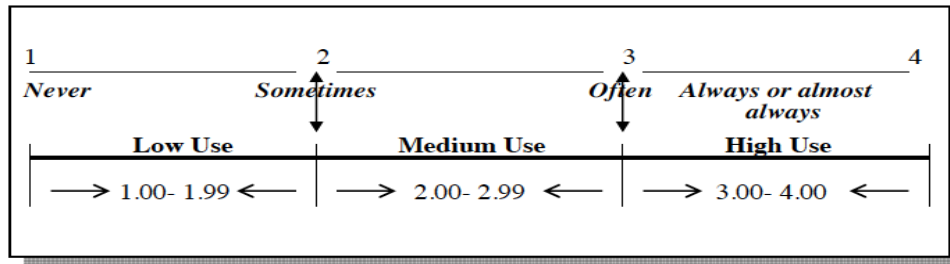
Figure 4.7. Scale used to classify strategy use. Source: Intaraprasert (2000, p. 167).

The analysis of the data was performed using three computer programmes: Microsoft Excel 2016, IBM SPSS Statistics (version 23) and R (version 3.6.1). Microsoft Excel 2016 program was mainly used in the descriptive analysis. In relation to the R program, it was only used to correct the *p* values obtained through multiple comparisons in the *t*-tests. Finally, IBM SPSS Statistics version 23 was used to run the following descriptive and inferential analyses:

- To obtain descriptors of the data, such as the mean or the standard deviation of the selection of VLSs and vocabulary receptive and productive knowledge.

- To carry out an analysis of the internal coherence of the groups and sub-groups suggested by Schmitt (1997) with the use of the Cronbach's Alpha coefficient.

- To carry out a factor analysis in order to determine the relationship between the selection of the different VLSs.

- To explore the vocabulary growth and the variations in the selection of vocabulary learning strategies in the course of one academic year.

- To analyse the differences between CLIL and EFL groups, in terms of selection of VLSs and the receptive and productive knowledge of vocabulary.

- To examine the differences within the CLIL group.

In statistics, there exist two types of analyses: parametric and nonparametric tests. The main difference between them is that parametric tests assume that the population follows a normal distribution whereas in non-parametric tests the normal distribution is not assumed. Therefore, in order to select the most appropriate statistical tests, some previous information of the data is needed. To run parametric tests, the following criteria need to be complied with (Cubo Delgado, 2011, p. 264):

- Continuous variables: only parametric tests can be used with continuous variables. Therefore, nominal or ordinal variables must be explored using non-parametric models.

- Normality: data must follow a Gaussian distribution, i.e., the distribution of the test must be bell-shaped. It is measured with the Shapiro-Wilk test.

- Homogeneity of variances (homoscedasticity): it is an assumption underlying some statistical tests, such as the *t*-tests, in which the population variances (i.e., the distribution, or 'spread' of scores around the mean) of two or more samples are considered equal (Salkind, 2010). It is checked by using the Levene test.

- Randomness of the data: a fundamental concept in statistics, "randomness exists when it is not possible to predict the outcome of an experiment or observation before it is performed" (Lewis-Beck, Bryman, & Futing Liao, 2004). It is checked with the run test of Randomness.

Every time a statistical analysis was performed, these assumptions were checked. Ass shown in Table 4.13, when one of these four assumptions were not met, a non-parametric test was used.

Table 4.13

A summary of the criteria followed to select the appropriate statistic instrument

Is it a continuous variable?	K-S	Randomness	Levene	Instrument
Yes	>0.05	>0.05	>0.05	Parametric test
No	-	-	-	Non-parametric test
Yes	<0.05	-	-	Non-parametric test
Yes	>0.05	<0.05	-	Non-parametric test
Yes	>0.05	>0.05	<0.05	Non-parametric test

Finally, in the case of parametric tests, an additional result was included in the analysis. Cohen's d values were calculated to obtain the effect sizes of the differences among the groups studied. To interpret the values, Plonsky and Oswald (2014) were followed: 1.00 and above is considered a large effect size, from 0.5 to 0.99 is considered a medium effect size and below 0.5 is considered a small effect size.

CHAPTER FIVE:
RESULTS

1. Introduction

This chapter summarises the results obtained from applying the VLSs questionnaire and the VLT and PVLT to the sample. It is organised as follows: it starts reporting on the most general results (i.e., the results from the overall sample), then it continues by exploring the differences between learners from the two different language learning background (CLIL vs EFL) examined in this dissertation, and it concludes with an analysis of the impact of the IAoE on the results.

2. Lexical competence of the overall sample

As the concept of lexical competence entails both lexical knowledge and selection of VLSs, this section is divided as follows: in 1.1, L2 lexical knowledge is described focussing first on the receptive mastery of the 2K and academic bands and moving then to the exploration of the productive knowledge of the same two bands. Then, section 1.2 details the selection of VLSs. Finally, section 1.3 relates both variables, VLSs selection and lexical knowledge, in order to determine whether the use of VLSs can be a predictor of vocabulary learning.

2.1. Vocabulary knowledge

This section presents a general overview of the sample's knowledge of the 2K and academic bands of vocabulary. It starts with an analysis of learners' recognition of the lexical items included in these two bands as well as their vocabulary growth after a year of study. Afterwards, it continues with an examination of learners' recall of the 2K and academic bands, not including, in this case, vocabulary growth, as productive data were only collected once.

2.1.1. Vocabulary recognition: high-frequency terms

2.1.1.1. Knowledge at T2

Data description: Table 5.1 shows the raw scores for the 2K band of the VLT obtained by the 138 students involved in the study. With a mean of 20.78 (SD= 6.52, max. = 30, min. = 2) at the end of Compulsory Secondary Education (CSE), participants recognise, in extrapolated values, an average of 1,378 out of the 2,000 most frequent words according to Nation's formula (1990). Although a normal frequency distribution of the scores was expected, the results show that as the percentage of hits increases, so does the percentage of population belonging to the interval (see Figure 5.1). The modal class interval is in the 27-30 interval and more than three-quarters of the population recognise more than 15 lexical items in the test.

2.1.1.2. Knowledge growth

Data description: results show a growth of 20.46% over T1, when participants achieved an overall mean score of 17.25 (SD = 6.12, max. = 29, min. = 5). In extrapolated values, this means that, at T1, participants recognised 1,009 words out of the 2K most frequent ones.

Table 5.1

2K VLT results, T1 and T2

	N	No of items	Minimum	Maximum	Mean	Std. Deviation
2K VLT (T1)	138	30	5	29	17.25	6.12
2K VLT (T2)	138	30	2	30	20.78	6.52

Distribution of the results: growth is also observed when comparing the frequency distribution and the percentage of participants mastering the 2K band at T1 and T2. As reflected in Figure 5.1, at T1, a fifth (21.74%) of participants recognised 22 to 24 words, followed by a considerable group of participants (16.67%) that recognised 13 to 15 items. This picture changes radically at T2, when more than a fifth of the participants know more than 27 terms, and nearly another fifth score 22 to 24 items. Besides, whereas 57.24% of the participants score higher than 50% on the test at T1, the figure increases substantially to 81.16% at T2.

Concerning the percentage of participants mastering the band, the data also indicates an evident change from T1 to T2: while at T1 only 7.97 % of the participants recognise more than 26 words out of the 30 examined, at T2, 27.97% of the sample demonstrate a full recognition of the band. Figure 5.1 shows the frequency distribution of the results at T1 (in yellow) and T2 (in orange).

Inferential statistics: the statistical analysis of the difference between T1 and T2 shows that the evolution in the receptive knowledge of the 2K band is statistically significant (z = -6.857, p = 0.000).
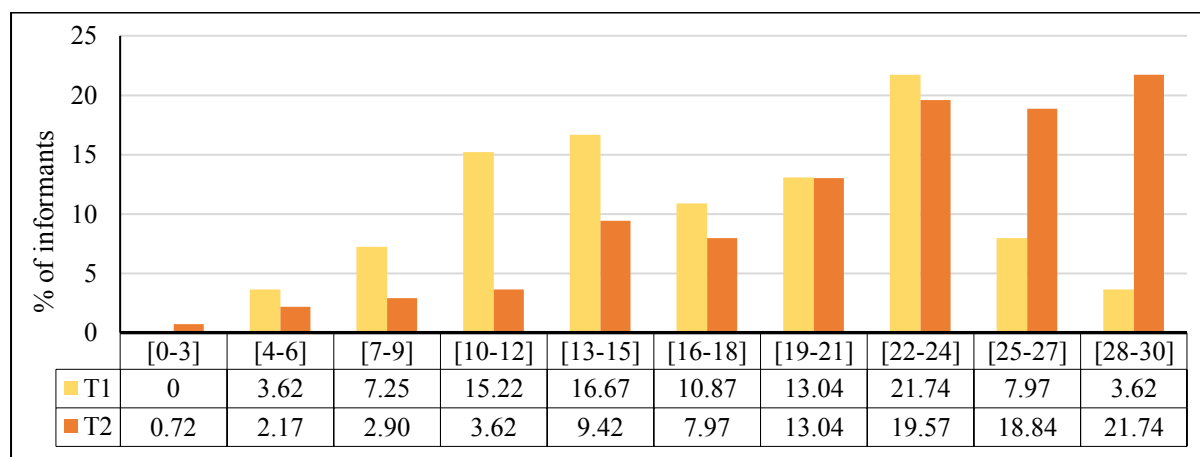
Figure 5.1. Frequency distribution of the 2K VLT scores at T1 and T2.

2.1.2. Vocabulary recognition: academic terms

2.1.2.1. Knowledge at T2

Data description: concerning the receptive knowledge of the academic band at the end of CSE, participants recognised a mean of 18.64 (SD = 7.92, max. = 30, min. = 0) words out of the 30 items included in the academic band of the VLT (see Table 5.2). This value corresponds to recognising, in extrapolated values, 354 out of the 570 word families comprising the AWL (Coxhead, 2000), according to Nation's formula (1990). Regarding the frequency distribution of the scores, the modal class interval comprises participants with scores ranging from 19 to 21, followed by the group of participants who recognise more than 27 items. Finally, the percentage of participants (11.59%) showing a full mastery of the academic band —by recognising more than 26 out of the 30 the academic terms— is also noteworthy.

2.1.2.2. Knowledge growth

Data description: participants' receptive academic knowledge increases by 23.11% over the same period of the previous year, where participants knew approximately half of the academic words ($\bar{x}$ = 15.14, SD = 7.95, max. = 30, min. = 0).

Table 5.2

Academic VLT results, T1 and T2

	N	No of items	Minimum	Maximum	Mean	Std. Deviation
Academic VLT (T1)	138	30	0	30	15.14	7.95
Academic VLT (T2)	138	30	0	30	18.64	7.92

Distribution of the results: the year variation is verified in the frequency distribution of the results shown in Figure 5.2. The main differences lie in the percentage of learners scoring 27+, which increases threefold in a year and in the group of participants recognising 25 to 27 words, which nearly doubles. Consequently, a rise in the percentage of learners mastering the band (i.e., scoring more than 26 terms)

is confirmed: whereas at T1, 11.59% of the participants show a receptive mastery of the 2K band, this percentage increases up to 21.01 at T2.

Inferential analysis: the statistical analysis shows that the evolution in the receptive knowledge of the academic vocabulary band is statistically significant (z = -5.633, p < 0.0004).

	[0-3]	[4-6]	[7-9]	[10-12]	[13-15]	[16-18]	[19-21]	[22-24]	[25-27]	[28-30]
T1	10.14	6.52	12.32	8.70	11.59	12.32	13.04	12.32	8.70	4.35
T2	6.52	2.17	3.62	6.52	13.77	10.14	19.57	9.42	15.22	13.04
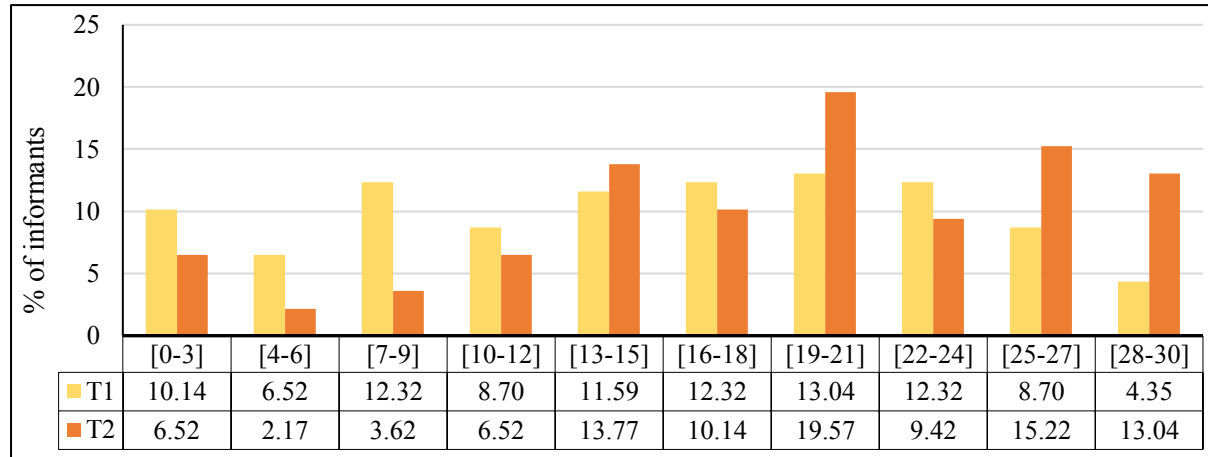
Figure 5.2. Frequency distribution of the academic VLT scores at T1 and T2.

2.1.3. Vocabulary recall: high-frequency terms

Data description: participants recall a mean of 6.70 (SD = 4.11, max. = 16, min. = 0) out of the 18 words measured in the academic PVLT. This means that participants can recall 601 words out of the 2K most frequent English terms in extrapolated values.

Distribution of the results: Figure 5.3 shows the frequency distribution of the 2K PVLT scores. An upward tendency is observed in the first frequency levels until reaching a peak in the 5-6 interval. After that point, the percentage of informants in each band sharply goes down, although a second peak is found at the 11-12 interval. Finally, it is also relevant that there are no participants recalling more than 16 words and that only two participants master the knowledge of the band productively.
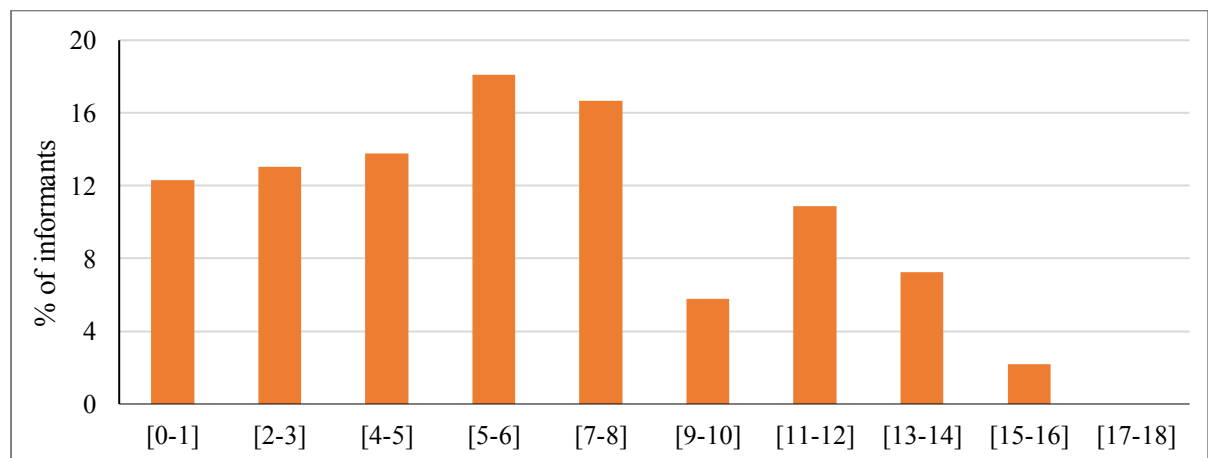
Figure 5.3. Frequency distribution of the 2K PVLT scores.

2.1.4. Vocabulary recall: academic terms

Data description: participants score 4.18 (SD = 3.11, max. = 13, min.= 0) out of the 18 words measured in the academic PVLT. In extrapolated values, this means that learners know a mean of 194 terms out of the 836 word families included in UWL (Xue & Nation, 1984)[4] in their productive dimension; in other words, they can recall, approximately, one out of five academic items.

Results distribution: Figure 5.4 shows the frequency distribution of the academic PVLT scores. As can be observed, most of the results are found in the left part of the graph, that is, the distribution is right-skewed. Only 3.61% of the sample score higher than 50%, with no participants demonstrating to master the academic band of vocabulary productively.
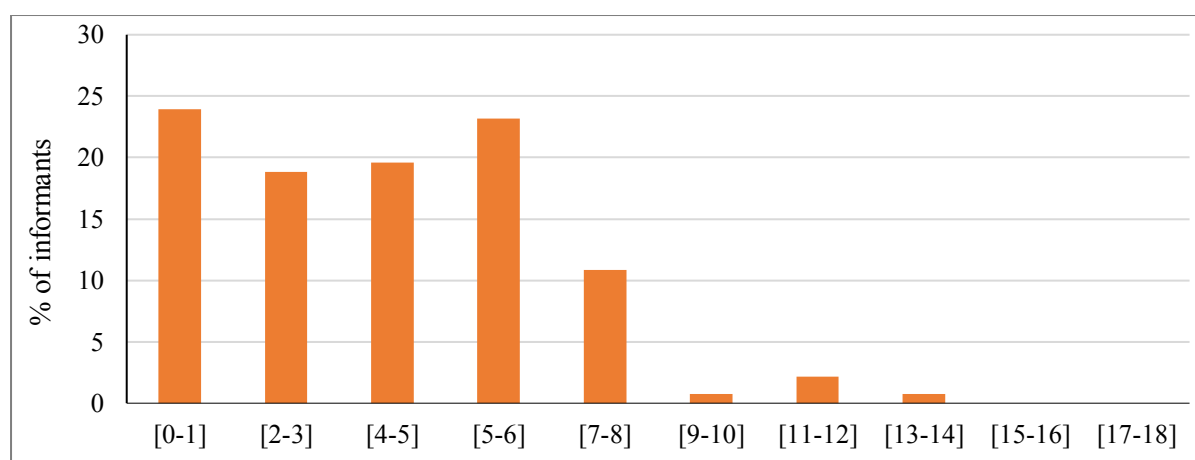
Figure 5.4. Frequency distribution of the Academic PVLT scores.

2.2. Overall use of Vocabulary Learning Strategies

This section presents an overall description of VLSs selection organised in three levels of analysis: overall use of VLSs, use of the different kinds of VLSs (according to the framework determined in section 4.4. of Chapter Four) and individual use of VLS. Intaraprasert's (2000, p. 167) scale is used to interpret the results.

2.2.1. VLSs selection at T2

Overall use of VLSs: at the end of CSE, participants make medium use of VLSs ($\bar{x} = 2.52$).

Use of the different kinds of VLSs: the preferred VLSs type is the 'linking' strategies ($\bar{x} = 2.76$), i.e., strategies involving the creation of links with other words, either in English or in their L1, to retain the meaning of new words. Following in the preference of VLSs is 'guessing from context' strategies ($\bar{x} = 2.71$), which includes VLSs used to understand what an unknown word means by connecting it to the target language or by inferring the gist of the text without knowing all the specific words. On the

---

[4] The UWL (1984) was used in the academic PVLT (1995) and the first version of the VLT (1984). However, when the VLT was updated (2001), the AWL was used to develop the academic band test. For this reason, when calculating the absolute values, the UWL is used for the academic PVLT and the AWL for the academic VLT.

other hand, the least widely used VLSs are the 'kinaesthetic' group ($\bar{x} = 1.48$). Indeed, there is extreme contrast between the use of this particular kind of VLSs and the preferred VLSs, whose use nearly doubles that of kinaesthetic strategies. Table 5.3 summarises the mean use of the types of VLSs.

Table 5.3

Descriptive analysis of the use of the different types of VLSs

Group	Mean Frequency score ($\bar{x}$)	SD	Frequency Category
Kinaesthetic strategies	1.48	0.77	Low
Mental imagery strategies	2.29	0.70	Medium
Social strategies involving interaction with teachers	2.37	0.70	Medium
Repetition strategies	2.55	0.89	Medium
Social strategies involving interaction with other students	2.61	0.86	Medium
Lexical analysis strategies	2.75	0.51	Medium
Guessing from context strategies	2.71	0.75	Medium
Linking strategies	2.76	0.75	Medium

Individual use of VLSs: this intermediate use is also observed when exploring the individual use of each VLSs. Table 5.4 displays a picture of the secondary-school learners' reported use of the nineteen items based on the mean frequency score. Only three strategies, 'checking for L1 cognates' ($\bar{x} = 3.13$), 'use of word lists' ($\bar{x} = 3.04$), and 'analysis of affixes and roots' present a high use. In contrast, two strategies are reported to be employed at the low-frequency level: 'use of physical action when learning a word' ($\bar{x} = 1.48$), and 'connection of the word to a personal experience' ($\bar{x} = 1.71$).

Table 5.4

Use of VLSs at the end of CSE (T2)

Strategies	Mean Frequency score	Standard Deviation (SD)	Frequency Category
Using physical action when learning a word	1.48	0.76	Low
Connecting a word to a personal experience	1.71	0.83	Low
Asking the teacher for paraphrasing or a synonym of the new word	2.10	0.94	Medium
Connecting the word to its synonyms and antonyms	2.21	0.89	Medium
Skipping or passing on a new word	2.34	0.95	Medium
Studying a word with a pictorial representation of its meaning	2.46	0.94	Medium
Verbal repetition	2.28	1.03	Medium
Using a bilingual dictionary	2.44	0.92	Medium
Written repetition	2.57	1.02	Medium
Asking students for meaning	2.61	0.86	Medium
Using a new word in a sentence	2.51	0.89	Medium
Grouping words together to study them	2.83	0.91	Medium
Using English-language media	2.99	0.93	Medium
Analysis of the part of speech	2.80	0.91	Medium
Analysis of any available picture or gesture	2.70	0.93	Medium
Asking the teacher for an L1 translation	2.65	0.83	Medium
Analysis of affixes and roots	3.02	0.88	High
Use of word lists	3.07	0.96	High
Checking for L1 cognates	3.09	0.94	High

2.2.2. Development of VLSs use

Overall use of VLSs

Data description: the analysis of the evolution in the VLSs selection along the academic year of study does not yield significant variations (see Figure 5.5). Overall use of VLSs slightly increases over T1 results, where participants presented a mean of use of 2.48. However, this evolution in VLSs use is almost negligible.

Use of the different kinds of VLSs

Data description: the evolution of the different types of VLSs shows scarce variations over T1 results. The preferred and least widely selected types of strategies at T1 coincide with those at T2. The most commonly used strategies are 'lexical analysis' and 'linking' strategies. In contrast, the least widely selected kind of strategy is the 'kinaesthetic' one. Figure 5.5 shows the use of strategies at T1 (in yellow) and T2 (in orange) and the increase (in grey) and decrease (in red) over one academic year. As evidenced, a generalised increase in the use of most kinds of VLSs is observed, except for 'mental imagery' and 'social strategies involving interaction with teachers', whose use decreases by 0.01 percentage points.

Inferential analysis: the analysis of the significance of these variations shows that the only significant (z = -2.577, p = 0.010) variation between T1 and T2 takes place in the use of 'kinaesthetic strategies'. That is, although, in general, slight variations are observed in the descriptive analysis, when they are analysed in-depth, only the increase in the use of the kinaesthetic category one is substantial enough to be considered significant.
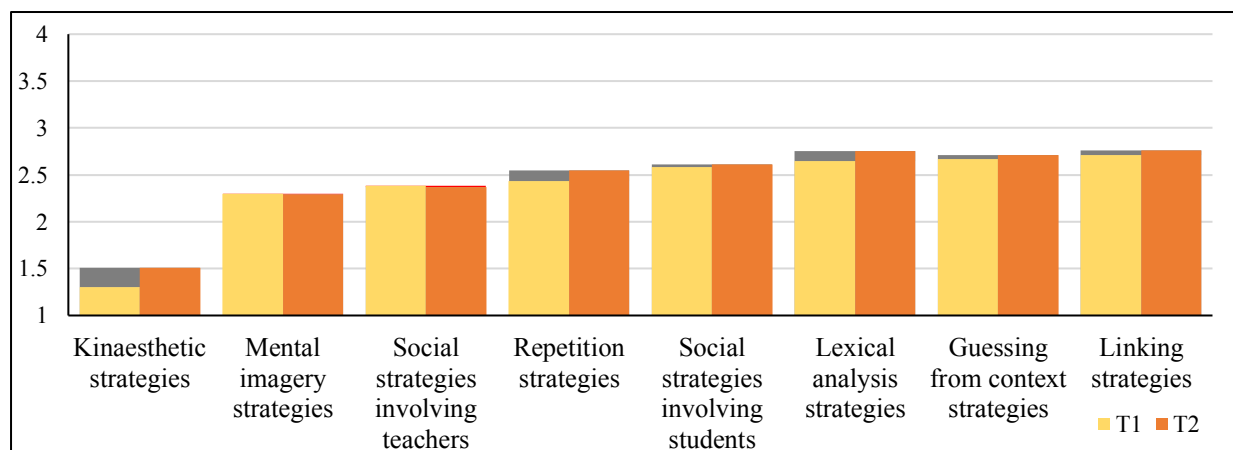
Figure 5.5. Mean selection of clusters at T1 and T2 and their variation along one year of study.

Individual use of VLSs

Data description: again, the three most widely selected strategies match at both times, 'checking for L1 cognates' strategy, 'use of word lists' and 'analysis of affixes and roots' being the most widely used. However, as shown in Table 5.5, there are differences in their use: in the case of the first strategy,

its use is reduced by 0.06%, whereas the use of the second and third strategies increases by 0.98% and 3.97% over the results at T1 respectively.

Table 5.5

A comparison of the top-three strategies in T1 and T2

	T1		T2	
	Strategies	Mean	Strategies	Mean
1	Checking for L1 cognates	3.12	Checking for L1 cognates	3.09
2	Use of word lists	3.04	Use of word lists	3.07
3	Analysis of affixes and roots	2.9	Analysis of affixes and roots	3.02

The least widely used strategies (see Table 5.6) correspond to the same three strategies already identified at T1, although the use of two of them increases in a year. The least widely used strategy is 'using physical action when learning a word', whose use increases by 0.21 percent points in one year. Concerning the second least selected VLS, it is, once again, 'connecting a word to a personal experience', whose selection remains more or less stable, with a slight reduction in its use (- 0.04). Finally, the third least widely used strategy corresponds, at both T1 and T2, to 'asking the teacher for paraphrasing or a synonym of new word', but its use increases so much way that, when at T1 it is considered to be used at a low level, at T2 it presents a medium use frequency.

Table 5.6

A comparison of the three least widely used strategies in T1 and T2

	T1		T2	
	Strategies	Mean	Strategies	Mean
1	Using physical action when learning a word	1.27	Using physical action when learning a word	1.48
2	Connecting a word to a personal experience	1.75	Connecting a word to a personal experience	1.71
3	Asking the teacher for paraphrasing or a synonym of new word	1.99	Asking the teacher for paraphrasing or a synonym of new word	2.10

Statistical analysis: Figure 5.6 shows the variations in the use of VLSs. In short, a total of 12 strategies present an increased use at T2, while in the remaining strategies, the opposite is observed. Among all these variations, the Wilcoxon test results indicate that only the use of two VLSs increases significantly: 'using English-language media' ($z = - 3.818$, $p = 0.000$) and 'using physical action when learning a word' ($z = -2.577$, $p = 0.010$).
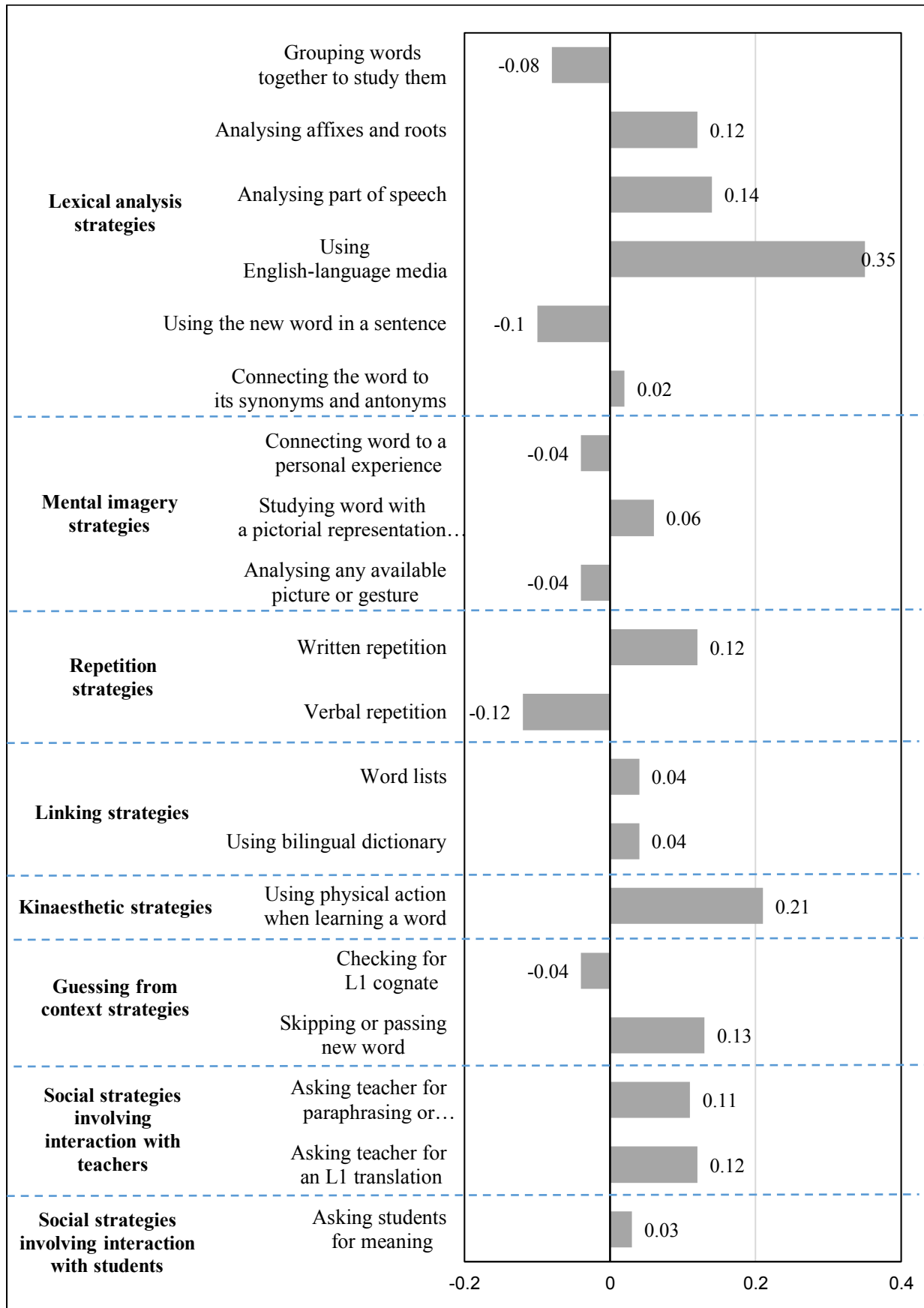
Figure 5.6. Variations in the individual use of strategies.

Given that the teaching methods may vary in the different schools and that this could affect the significance of these results, the particular evolution of these two strategies is explored in the four schools. Starting with the strategy 'using physical action when learning a word', a significant growth in its use is only found in school 4 ($z = -2.143$, $p = 0.032$), whereas in the case of school 1 ($z = -1.628$, $p = 0.104$), school 2 ($z = -0.378$, $p = 0.705$) and school 3 ($z = -1.265$, $p = 0.206$), no significant increase is found. In light of these results, variation may not be said to be a generalised phenomenon. Rather, it would seem to be related to other factors. The analysis of the strategy 'using English-language media' allows us to draw a different picture. The significance of the growth rate of this strategy is a constant in two out of the four schools (schools 2 [$z = -2.147$, $p = 0.032$] and 3 [$z = -2.303$, $p = 0.021$]), and in a school 1 is on the verge to be significant ($z = -1.865$, $p = 0.062$). Therefore, it could be argued that the evolution of the use of 'using English-language media' is not related exclusively to the teaching practices or teachers in specific schools and it could be considered a widespread phenomenon.

2.3. Correlation between the use of vocabulary learning strategies and vocabulary learning

In this section, the relationship between lexical knowledge and use of VLSs is analysed to identify which VLSs are related to successful vocabulary learning. With this aim, the selection of VLSs at T2 is correlated with T2 receptive and productive knowledge of the 2K and academic bands.

Starting with the relationship between receptive knowledge of the 2K band of vocabulary and the use of specific kinds of VLSs, a Spearman's rho only finds a positive correlation between the use of 'lexical analysis' strategies and receptive knowledge of the 2K band ($r_s = 0.262$, $p = 0.001$) and academic bands ($r_s = 0.202$, $p = 0.048$), whereas no relationship is established between the use of the different kinds of VLSs and the recalling of high-frequency and academic terms. Table 5.7 summarises the results.

Table 5.7

A summary of the groups of strategies that are predictors of a larger or smaller vocabulary at T2

	Indicators of greater knowledge	Indicators of lower knowledge
Receptive 2K	Lexical analysis strategies	
Receptive academic	Lexical analysis strategies	
Productive 2K		
Productive academic		

At the level of individual use of VLSs, some particular VLSs seem to be used as predictors of larger vocabulary knowledge. As illustrated in Table 5.8, a calculation of Spearman's rho finds that a larger recognition rate is positively correlated (N = 138, $p < 0.01$) with a greater use of the following two strategies: 'analysis of affixes and roots' ($r_s = 0.195$, $p = 0.022$), and 'connecting the word to synonyms and antonyms' ($r_s = 0.233$, $p = 0.006$).

Concerning the relationship between a larger academic receptive vocabulary and the use of individual VLSs, a Spearman's rho calculation finds a positive correlation (N = 138, $p < 0.05$) with the 'analysis of affixes and roots' strategy ($r_s = 0.306$, $p < 0.0004$), 'asking the teacher for an L1 translation' ($r_s = 0.172$, $p = 0.043$) and 'asking students for meaning' ($r_s = 0.215$, $p = 0.011$). Results are summarised in Table 5.8.

Table 5.8

A summary of the individual VLSs that correlated significantly with each of the vocabulary bands explored.

	Indicators of higher level	Indicators of lower level
Receptive 2K	Analysis of affixes and roots	
	Connecting to its synonyms and antonyms	
Receptive academic	Analysis of affixes and roots	
	Asking the teacher for an L1 translation	
	Asking students for meaning	
Productive 2K		
Productive academic		

This first section has attempted to portray the overall lexical competence of the sample that took part in the study, acknowledged as lexical knowledge and selection of VLSs. Regarding lexical knowledge of high-frequency and academic terms, participants' most outstanding proficiency is found in the receptive dimension of high-frequency items, whereas they show a lower productive mastery of academic terms. In general, participants present greater receptive knowledge than productive ability. This holds true for both general and academic vocabulary and in line with previous studies in the field (Melka, 1997; Takala, 1984; Laufer, 1998; Fan, 2000; Laufer & Paribakht, 1998), which also place the productive capabilities of a vocabulary band at a lower level than receptive knowledge of the same band.

Similarly, the 2K band of vocabulary knowledge is also greater than the academic vocabulary knowledge, both receptively and productively. These findings are not surprising. Experts usually situate the academic terms on an interval ranging from 3K to 5K, as the academic band does not follow a frequency pattern. Since academic vocabulary is less frequent, lower knowledge of this band than that of the 2K band could be expected.

As for the selection of VLSs, results show that the use of 'linking' strategies overrides the remaining groups. At the same time, participants make an almost null use of VLSs that involve any kind of movement, that is, kinaesthetic strategies. Moreover, at the individual use of VLSs, participants

show a preference towards using strategies such as 'checking for L1 cognates', 'use of word lists' and 'analysis of affixes and roots'.

The analysis of the relationship between both variables has shown that the receptive knowledge of the 2k and academic bands is positively correlated with 'lexical analysis' strategies. Besides, the receptive knowledge of the 2K band is also positively related to the individual use of two strategies: 'analysis of affixes and roots' and 'connecting to its synonyms and antonyms'. For its part, the receptive knowledge of the academic band is positively related to three VLSs: 'analysis of affixes and roots', 'asking the teacher for L1 translation' and 'asking students for meaning'. Surprisingly, no correlations have been found between the use of VLSs and productive lexical knowledge.

The following section deals with the analysis of the impact of the language teaching approach (CLIL vs EFL) on lexical competence.

3. CLIL vs EFL: a comparison of learners' lexical competence

3.1. Vocabulary knowledge

As mentioned in the theoretical framework, the implementation of CLIL in Europe has had an important impact on the way languages are taught in formal settings. This educational approach involves implementing new teaching and learning techniques and has brought into the L2 classroom methods, objectives and activities that could affect vocabulary learning. This section aims to explore the differences between CLIL and regular EFL learners' receptive and productive knowledge of high-frequency (2K band) and academic terms by comparing their performance in the 2K and academic bands of the VLT and PVLT.

3.1.1. Vocabulary recognition: high-frequency terms

3.1.1.1. Knowledge at T2

Data description: as Table 5.9 shows, there is a great difference between CLIL and mainstream EFL learners in the recognition of high-frequency terms by the end of CSE. On the one hand, CLIL learners have a mean score of 23.99 (SD = 4.37, max. = 30, min. = 11) in the 2K VLT, which is to say that they recognise 1,607 terms out of the 2K most frequent ones according to Nation's formula (1990). Besides, 40.24% of these learners recognise more than 26 terms, and, consequently, they are considered to have full knowledge of the band. On the other hand, the EFL participants recognise 16.09 of the 30 words included in the 2K VLT (SD = 6.58, max. = 27, min. = 2), which, in extrapolated values, means that they recognise a total of 1,042 words. Only 8.89% of these learners recognise 26 or more words, i.e., have full recognition of the band.

Table 5.9

CLIL and mainstream EFL learners' 2K receptive VLT results at T2

| | N | Raw data | | | | | In extrapolated values | | % of participants mastering the band |
		No. of items	Min	Max	Mean	SD	No. of items	Mean	
CLIL	82	30	11	30	23.99	4.10	2000	1607	40.24
EFL	56	30	2	27	16.09	6.58	2000	1042	8.89

Frequency distribution: this better proficiency of CLIL learners is confirmed in the distribution patterns of both groups. As shown in Figure 5.7, CLIL learners' distribution of results is left-skewed, i.e., the left tail is longer than the right one, whereas, in the case of regular EFL learners, the bulk of the population is located in the centre of the graph. Moreover, the distribution of the results and the number of participants per interval differentiate the two groups. In CLIL, the number of participants per interval grows as the percentage of success increases. Half of the CLIL participants recognise more than 24 test-items, and the modal class interval is located in the 28-30 interval, with 34.15% of the CLIL sample scoring within this interval. In contrast, in mainstream EFL learners, an upward tendency in the number of participants per interval is found until reaching the 19-21 interval, which is the modal class interval. After this point, a sharp decrease in the number of informants per interval is found.

Inferential analysis: the *t*-test shows that the differences accounted for the CLIL and regular EFL groups are statistically significant ($t(84.04) = -7.978$, $p < 0.0004$, $d = 1.44$).

	[0-3]	[4-6]	[7-9]	[10-12]	[13-15]	[16-18]	[19-21]	[22-24]	[25-27]	[28-30]
CLIL	0.00%	0.00%	0.00%	1.22%	2.44%	3.66%	8.54%	23.17%	26.83%	34.15%
EFL	1.79%	5.36%	7.14%	8.93%	14.29%	17.86%	19.64%	8.93%	7.14%	8.93%
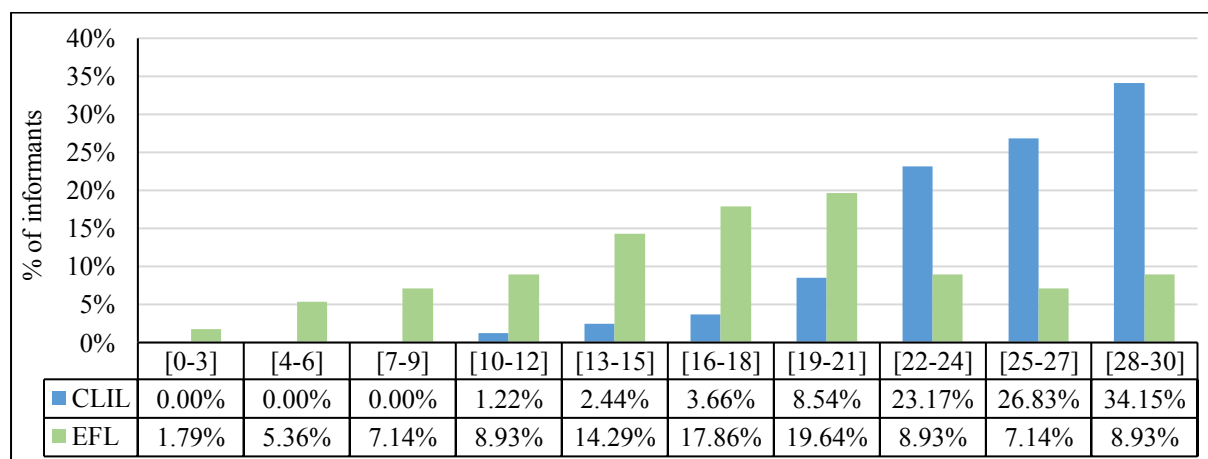
Figure 5.7. Frequency distribution of CLIL and regular EFL learners' receptive knowledge of the 2K band at T2.

3.1.1.2. Knowledge growth

To explore knowledge growth, first, T1 data is described, and the significance of the difference between CLIL and EFL learners' results at T1 is calculated. Then, the evolution of each of the two groups (CLIL and EFL) is analysed, concluding with the comparison between both groups.

T1 data

Data description: on the one hand, CLIL learners had an average score on 2K-band VLT of 20.85 (SD = 4.37, max. = 29, min. = 11), which is to say, they recognised a total of 1,352 out of the 2,000 most frequent words, with no differences among schools (F = 1.347, p = 0.265) or genders (F = 0.069, p = 0.793). Of them, 13.41% had a full mastery of the band. On the other hand, mainstream EFL students recognised an average of 11.96 out of the 30 words included in the VLT (SD = 4.18, max. = 25, min. = 5), with no significant differences among schools (F = 1.501, p = 0.225) or genders (F = 0.206, p = 0.652). That is to say that, in extrapolated values, mainstream EFL learners knew approximately 798 out of the 2K most frequent words.

Inferential analysis of the difference at T1: the difference between both groups was already significant (t(7.978) = 8.0037, p < 0.0004, d = 2.07) the first time the tests were administered.

Growth

Data description: although significant differences are found at both T1 and T2, the difference between groups decreases in one academic year: CLIL learners' recognition of high-frequency words shows an increase of about 15.06% in a year, while EFL learners' growth rate is greater, as their receptive knowledge of the 2K words increases by approximately 34.54% in a year.

Inferential analysis of each group's knowledge growth: the analysis of CLIL (t(81) = -7.191, p < 0.0004, d = -0.77) and EFL (t(55) = -4.634, p < 0.0004, d = -0.75) learners' vocabulary knowledge growth shows that the lexical development taking place in both groups is statistically significant.

Inferential comparison of both groups' growth: given the significant difference between both groups' lexical knowledge at T1, a comparison of both groups' receptive 2K knowledge growth requires transforming the extrapolated values into relative increases. To do so, the following equation is applied:

$$\frac{VLT\ result\ at\ T2 - VLT\ result\ at\ T1}{VLT\ result\ at\ T1}$$

Once the relative increases are calculated, the significance of the growth rates is determined. The analysis indicates that difference between CLIL and EFL learners' lexical knowledge growth is not statistically significant (U(N_{CLIL}=82, N_{EFL}=56)= 1938.00, z = -1.552, p = 0.121). Considering schools individually, this holds true in three schools, but a significant greater receptive vocabulary growth is found in favour of mainstream EFL learners in school 3.

3.2.1. Vocabulary recognition: academic terms

3.2.1.1. Knowledge at T2

Data description: Table 5.10 shows receptive knowledge of the academic VLT at T2 for each group. By the end of CSE, CLIL learners know 22.41 (SD = 5.37, max. = 30, min. = 0) of the 30 items

explored in the academic VLT. In extrapolated values, this means that CLIL learners recognise a total of 426 of the 570 word families included in the AWL (Coxhead, 2000). Moreover, an exploration of the internal consistency of the results finds no significant differences among genders ($F = 0.049$, $p = 0.764$) or schools ($F = 2.426$, $p = 0.072$). In comparison, mainstream EFL learners' performance is substantially lower, as these learners demonstrate knowledge of 13.11 (SD = 7.84, max. = 29, min. = 0). In extrapolated values, this means 249 of the listed words, with no significant differences among genders (F = 0.004, $p = 0.949$) or schools (F = 2.636, $p = 0.059$).

Table 5.10

CLIL and EFL learners results of the academic band of the VLT at T2

	N	Raw data					In extrapolated values		% of participants mastering the band
		No. of items	Min	Max	Mean	SD	No. of items	Mean	
CLIL	82	30	0	30	22.41	5.37	570	426	31.71
EFL	56	30	0	29	13.11	7.84	570	249	5.36

Results distribution: Figure 5.8 shows the CLIL and EFL learners' results distribution. While most EFL participants (54.82%) recognise less than 15 test-items, CLIL learners' results distribution is left-skewed, as more than 90% of the learners understanding more than 50% of the academic terms. Differences are also reflected in the percentage of participants presenting scores similar to or higher than 26 words. In the case of CLIL participants, 31.71% of the sample hits 26 out of the 30 terms or higher, so they master the band, whereas 5.36% of mainstream EFL learners reach the score needed to prove their fully receptive knowledge of the academic band.

Inferential analysis: according to the *t*-test results, CLIL learners significantly outperform ($t(89.59) = 7.731$, $p < 0.0004$, $d = -1.39$) the mainstream EFL group in the academic band of the VLT.

	[0-3]	[4-6]	[7-9]	[10-12]	[13-15]	[16-18]	[19-21]	[22-24]	[25-27]	[28-30]
CLIL	1.22%	0.00%	0.00%	1.22%	1.22%	13.41%	19.51%	15.85%	19.51%	28.05%
EFL	14.29%	3.57%	8.93%	12.50%	16.07%	14.29%	10.71%	7.14%	8.93%	3.57%
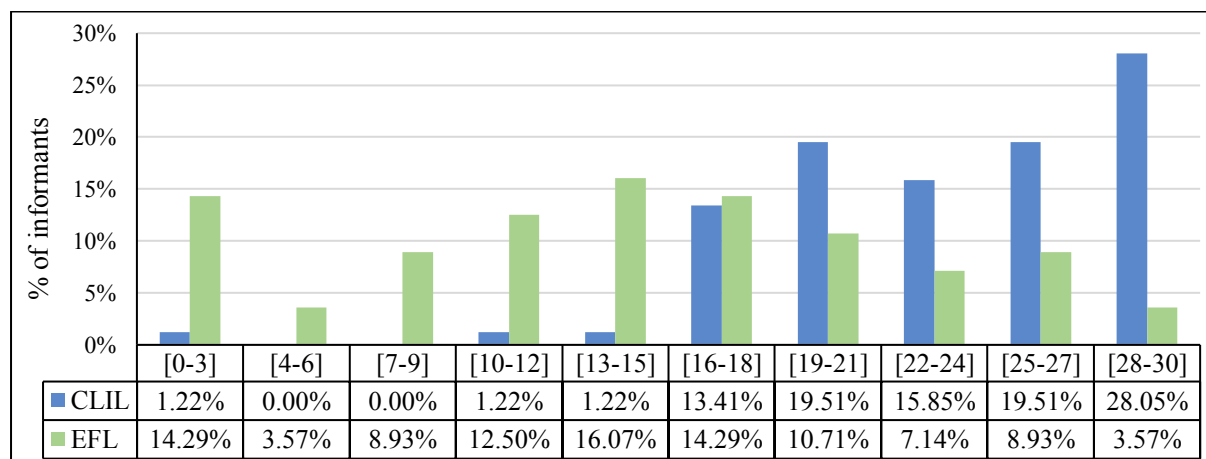
Figure 5.8. Frequency distribution of CLIL and EFL learners results of the academic VLT at T2.

3.2.1.2. Knowledge growth

This section starts with a description and inferential analysis of the differences between CLIL and EFL learners' academic knowledge at T1. Then, it moves on to explore the evolution of each of the two groups (CLIL and EFL). Finally, both groups' growth rates are compared.

T1 data

Data description: Table 5.11 summarises CLIL and EFL learners' results in the academic VLT. As can be observed, CLIL learners ($\bar{x}$= 19.54, SD =5.94, max. = 30, min. = 4) obtained better results than the EFL group ($\bar{x}$= 8.70, SD =5.87, max. = 21, min. = 0), being, in both cases, homogenous results concerning schools or gender differences. This is not the only difference between both groups: while nearly a fifth of CLIL participants master the academic band receptively, none of the EFL participants reaches 26 hits, which is the minimum number of hits required to consider that a band is mastered.

Table 5.11

CLIL and EFL learners results of the academic band of the VLT at T1

| | N | Raw data | | | | | In extrapolated values | | % of participants |
		No. of items	Min	Max	Mean	SD	No. of items	Mean	mastering the band
CLIL	82	30	4	30	19.54	5.94	570	371	19.51
EFL	56	30	0	21	8.70	5.87	570	165	-

Growth

Data description: the comparison between T1 and T2 results shows that CLIL and EFL groups improve their performance by 14.68% and 50.68 % respectively.

Inferential analysis of the evolution in both groups: the *t*-test confirms that CLIL (t(81) = -4.472, p < 0.0004, d = -0.18) and EFL (t(55) = -3.631, p < 0.001, d = -0.63) learners' receptive knowledge of the academic band grows significantly over T1 results. In the case of CLIL learners, this significant increase is found in all the schools with the exception of school 2, while in the case of regular EFL learners, this significant improvement is only found in school 2 (z = -3.159, p = 0.002).

Inferential comparison of the difference between groups: as with the receptive knowledge of the 2K band, the increase in regular EFL learners' recognition of academic terms is particularly intense. The analysis between both groups shows a statistically significant (t(135) = -3.565, p < 0.0004, d = -0.55) difference in favour of the regular EFL group. This statement remains valid for three out of the four schools, except for school 3, where no statistically significant difference is found.

3.1.3. Vocabulary recall: high-frequency terms

Data description: CLIL learners are also found to perform better than the mainstream EFL group in the 2K PVLT. Concretely, CLIL learners obtain a mean of 8.54 out of the 18 words included in the

test (SD = 3.64, max. = 16, min. = 0), which translates to a productive knowledge of 948 words, with no differences among schools or gender. On the other hand, EFL participants present a mean score of 4 words (SD = 3.17, max. = 12, min. = 0), i.e., in extrapolated values, they can recall a mean of 444 words out of the 2K most frequent terms. This group also presents homogeneous results regardless of gender or schools. Table 5.12 shows the descriptive statistics for each group.

Table 5.12

CLIL and mainstream EFL learners' 2K PVLT results

	N	Raw data					In extrapolated values		% of participants mastering the band
		No. of items	Min	Max	Mean	SD	No. of items	Mean	
CLIL	82	18	0	16	8.54	3.64	2000	948	2.44
EFL	56	18	0	12	4	3.17	2000	444	-

Results distribution: Figure 5.9 shows the frequency distribution of CLIL and regular EFL participants. The CLIL sample seems to follow a normal distribution. The number of participants per interval goes up until reaching a peak at the 7-8 band, which is the modal class interval. After that, a marked decreasing trend is observed, although another peak is observed at the 11-12 interval. In contrast, EFL learners' results are right skewed, with 94.64% of the population scoring lower than 50% of the words. This difference also links to the percentage of participants mastering the whole band: only two CLIL learners attain the minimum score to attain a complete mastery of the 2k band.

Inferential analysis: productive knowledge of the 2K band is significantly different ($t(136)$ = -7.556, $p < 0.0004$, $d = 1.40$) in favour of CLIL learners. This difference is significant in three of the schools explored, but a non-significant better CLIL learners' performance is found in school 3 (z = -1.867, $p = 0.062$).

	[0-1]	[2-3]	[4-5]	[5-6]	[7-8]	[9-10]	[11-12]	[13-14]	[15-16]	[17-18]
CLIL	3.66%	4.88%	9.76%	19.51%	21.95%	9.76%	14.63%	12.20%	3.66%	0.00%
EFL	25.00%	25.00%	19.64%	16.07%	8.93%	0.00%	5.36%	0.00%	0.00%	0.00%
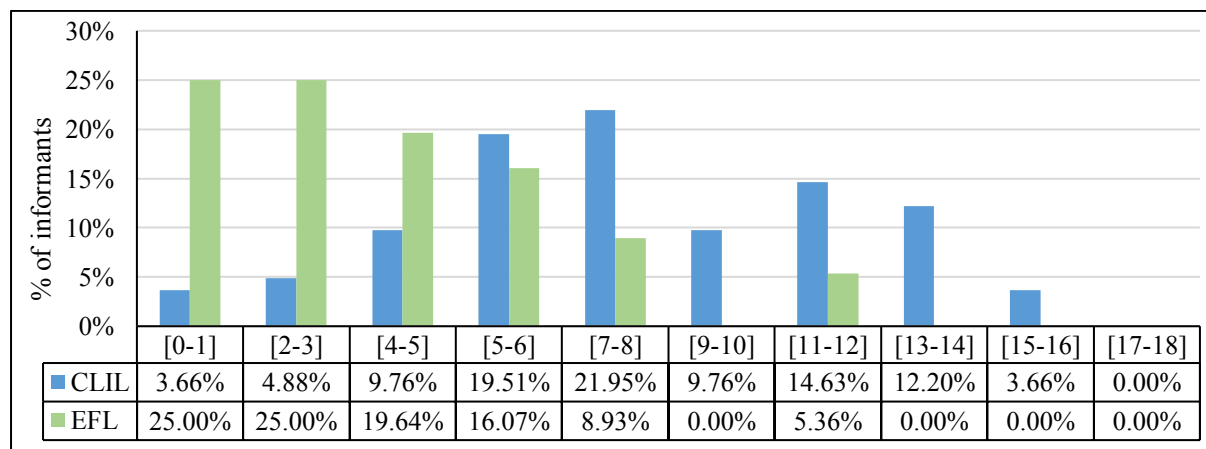
Figure 5.9. CLIL and mainstream EFL learners' 2K PVLT results frequency distribution.

3.1.4. Vocabulary recall: academic terms

Data description: as with the rest of dimensions, CLIL learners present better results in their recall of academic terms ($\bar{x}$= 5.63 out of 18, SD = 2.72, max. = 13, min. = 0), nearly tripling mainstream EFL participants' mean score ($\bar{x}$ = 2.05 words, SD = 2.31, max. = 8, min. = 0). That is, in extrapolated values, while CLIL learners can recall a total of 261 word families, regular EFL learners can only produce 95. Table 5.13 shows the descriptive statistics for each group.

Table 5.13

CLIL and mainstream EFL learners' academic PVLT results

| | N | Raw data | | | | | In extrapolated values | | % of participants mastering the band |
		No. of items	Min	Max	Mean	SD	No. of items	Mean	
CLIL	82	18	0	13	5.63	2.72	836	261	-
EFL	56	18	0	8	2.05	2.31	836	95	-

As for the analysis of the mean scores' homogeneity in both groups regarding genders and schools, differences arise. Whereas in the case of CLIL learners, the variables gender (F = 0.071, p = 0.791) and school (F = 0.554, p = 0.647) are not significant within the group, in the mainstream EFL group, there are no significant gender differences (F = 0.052, p = 0.821), but learners' results differ significantly depending on the school analysed (F = 3.076, p = 0.036).

Frequency distribution: Figure 5.10 shows that distributions are right-skewed, although there are clear differences between them. Starting with CLIL learners, the number of participants per band increases gradually until reaching the 5-6 interval, at which we find a peak. After that, there is a decreasing tendency, and the number of participants progressively falls. As for mainstream EFL learners, the distribution follows a decreasing pattern, and the peak is found at the 0-1 interval, since half of the learners recall zero or one term. In this case, none of the participants, neither CLIL nor mainstream EFL learners, productively master the academic band.

	[0-1]	[2-3]	[4-5]	[5-6]	[7-8]	[9-10]	[11-12]	[13-14]	[15-16]	[17-18]
CLIL	6.10%	15.85%	24.39%	30.49%	17.07%	1.22%	3.66%	1.22%	0.00%	0.00%
EFL	50.00%	23.21%	12.50%	12.50%	1.79%	0.00%	0.00%	0.00%	0.00%	0.00%
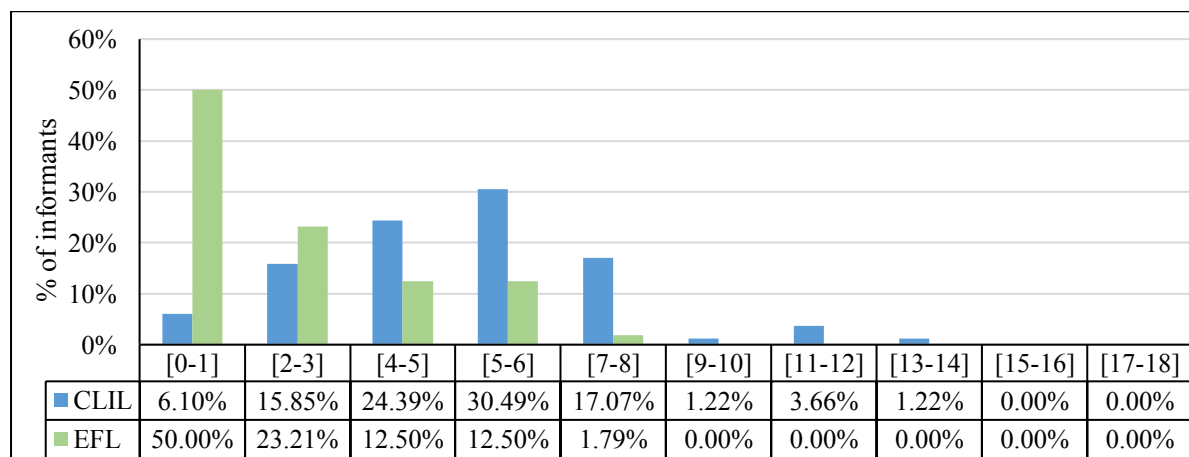
Figure 5.10. CLIL and EFL learners' academic PVLT results frequency distribution.

Inferential analysis: the test shows that the difference between CLIL and mainstream EFL participants is statistically significant ($t(136)$ = -8.067, $p < 0.0004$, $d = 1.41$). The significance of the difference is consistent no matter the school examined.

In short, this sub-section has provided a detailed analysis of the receptive and productive knowledge of the 2K and academic bands of CLIL and mainstream EFL learners showing a clear advantage for the CLIL learners. Additionally, the longitudinal analysis of receptive vocabulary knowledge yields another relevant finding: the larger IAoE CLIL learners receive throughout an academic year is not translated into a more extensive growth of the recognition of the 2K and academic vocabulary bands. The various reasons that can be attributed to such findings will be discussed in Chapter Six.

3.2. Vocabulary learning strategies

In our theoretical framework, the teaching method (Oxford, 1989; Oxford & Nyikos, 1989; Politzer, 1983), the type of tasks (Bialystok, 1981; Bacon, 1992; Vandergrift, 1997; Manchón, 2001; Trenchs, 1996; Cohen, Weaver, & Li, 1998) and the teaching approach have been shown as some of the factors that may influence L2 learners' selection of VLSs. This section aims to determine to what extent CLIL and regular EFL learners differ in their selection of VLSs and, if so, whether CLIL fosters the selection of the VLSs that are positively related to lexical knowledge.

3.2.1. Vocabulary learning strategies selection at T2

Overall use of VLSs

Data description: by the end of CSE, CLIL and mainstream EFL learners differ in their selection of VLSs, with CLIL learners ($\bar{x}$ = 2.54) making non-significant (U(N_{CLIL}=82, N_{EFL}=55) = 2203.00, z = -0.229, p = 0.819) greater use of strategies than their regular EFL counterparts ($\bar{x}$ = 2.50). A closer look at the data yields some discrepancies between both groups, such as their use of the different kinds of strategies and their preferred and least widely use VLSs that should be explored in greater detail.

Selection of the different kinds of VLSs

Data description: CLIL and regular EFL groups disagree on their preferred kind of VLSs (see Table 5.14), which are, for CLIL learners' guessing from context' strategies ($\bar{x}$ = 2.82), while for mainstream EFL learners, they are 'linking' strategies ($\bar{x}$ = 2.87). Nevertheless, both groups agree on the least preferred strategies, the kinaesthetic strategies, although mainstream EFL learners present a slightly higher use ($\bar{x}$ = 1.52) in comparison to their CLIL counterparts ($\bar{x}$ = 1.45).

All in all, CLIL learners make greater use of 'lexical analysis' (dif.= 0.22), 'guessing from context' (dif. = 0.08) and social strategies involving interaction with teachers (dif. = 0.12) and students (dif. =

0.15). On the other hand, mainstream EFL learners show greater use of 'mental imagery' (dif. = 0.10), 'repetition' (dif. = 0.16), 'linking' (dif. = 0.19) and 'kinaesthetic' (dif. = 0.07) strategies.

Inferential analysis: CLIL and EFL learners differ significantly in their use of 'lexical analysis' (U(N_{CLIL}=82, N_{EFL}=55)= 1752.500, z = -2.220, p = 0.026).

Table 5.14

Differences between CLIL and mainstream EFL learners' selection of types of strategies.

Group	CLIL learners ($\bar{x}$)	EFL learners ($\bar{x}$)	*p-value*
Lexical analysis strategies	2.82	2.6	0.026*
Mental imagery strategies	2.35	2.25	0.455
Repetition strategies	2.36	2.52	0.336
Linking strategies	2.68	2.87	0.187
Kinaesthetic strategies	1.45	1.52	0.536
Guessing from context strategies	2.74	2.66	0.509
Social strategies involving interaction with teachers	2.42	2.30	0.399
Social strategies involving interaction with students	2.67	2.52	0.982

Individual use of VLSs

Description: as displayed in Table 5.15, CLIL and regular EFL informants' preferred individual strategy does not match: mainstream EFL participants prefer the use of 'word lists' ($\bar{x}$ = 3.43), whereas, in the case of CLIL participants, the preferred VLSs is 'analysis of affixes and roots' ($\bar{x}$ = 3.17). As for the second preferred VLSs, it is, for both groups, 'checking for L1 cognates', although CLIL learners ($\bar{x}$ = 3.15) resort to this strategy more frequently than their regular EFL counterparts ($\bar{x}$ = 3.00). To complete the analysis of the top-three strategies, the two groups, again, do not share the third most widely selected strategy, which is 'use of English-language media' ($\bar{x}$ = 2.78) for CLIL participants, and 'grouping words together to study them' ($\bar{x}$ = 2.82) for mainstream EFL learners.

Differences regarding VLSs use go beyond the top-three strategies. Generally speaking, in comparison to regular EFL learners, the CLIL group shows greater use of ten strategies: 'analysis of the part of speech' (dif. = 0.30), 'analysis of affixes and roots' (dif. = 0.37), 'checking for L1 cognates' (dif. = 0.15), 'using a bilingual dictionary' (dif. = 0.14), 'asking the teacher for paraphrasing or a synonym of a new word' (dif. = 0.24), 'using new word in a sentence' (dif. =0.12), 'asking students for meaning' (dif. = 0.15), 'connecting the word to its synonyms and antonyms' (dif. = 0.23), 'using English-language media' (dif. = 0.29), and 'asking the teacher for an L1 translation (dif. = 0.02).

Inferential analysis: significant differences are only found in the selection of the strategies 'analysis of affixes and roots' (U(N_{CLIL}=82, N_{EFL}=56) = 1785.*00*, z = -2.346, p = 0.019), 'analysis of

the part of speech' (U(N$_{CLIL}$=82, N$_{EFL}$=56) = 1847.00, z = -2.049, p = 0.040), and 'use of word lists' (U(N$_{CLIL}$=82, N$_{EFL}$=56) = 1670, z = -2.895, p = 0.004). Table 5.15 shows CLIL and regular EFL learners' mean use of the strategies examined at T2.

Table 5.15

CLIL and mainstream EFL learners' mean use of strategies at T2.

Group[5]	Strategy	CLIL learners' mean of use	EFL learners' mean of use	*P*-value
LA	Analysis of part of speech	2.91	2.61	0.04*
	Analysis of affixes and roots	3.17	2.80	0.019*
	Using a new word in a sentence	2.56	2.44	0.462
	Grouping words together to study them	2.82	2.86	0.864
	Connecting the word to its synonyms and antonyms	2.30	2.07	0.161
	Using English-language media	3.11	2.82	0.051
MI	Analysis of any available picture or gesture	2.70	2.71	0.942
	Studying word with a pictorial representation of its meaning	2.40	2.55	0.280
	Connecting word to a personal experience	1.65	1.78	0.461
RP	Verbal repetition	2.20	2.39	0.215
	Written repetition	2.52	2.64	0.520
LI	Word lists	2.87	3.38	0.004*
	Using a bilingual dictionary	2.50	2.36	0.310
KI	Using physical action when learning a word	1.45	1.52	0.536
GC	Skipping or passing on a new word	2.34	2.34	0.951
	Checking for L1 cognates	3.15	3.00	0.352
ST	Asking the teacher for an L1 translation	2.66	2.64	0.982
	Asking the teacher for paraphrasing or a synonym of a new word	2.20	1.96	0.162
SS	Asking students for meaning	2.67	2.52	0.314

---

[5] LA = Lexical analysis strategies, MI = mental imagery; RP= repetition strategies; LI = linking strategies; KI = kinaesthetic strategies; GC = Guessing from context strategies; ST = social strategies involving interaction with the teacher; SS = social strategies involving interaction with other students.

3.2.2. Evolution in the selection of vocabulary learning strategies

Overall selection of VLSs

Data description: the longitudinal analysis shows some differences in the evolution of VLS use by CLIL and EFL learners. CLIL learners' VLSs use scarcely increases by 0.02 percentual points (Z= -0.660, p = 0.509), while regular EFL learners' use of VLS shows a significant increase (Z= -1.990, p = 0.047) of 0.07 percentage points.

Selection of the kinds of VLSs

Description of the differences at T1: the most remarkable differences are found in the use of 'guessing from context' (dif. = 0.13), 'linking' (dif. = 0.11) and 'lexical analysis' strategies (dif. = 0.09).

Inferential analysis of the differences at T1: the results show that none of the differences is representative enough to be considered statistically significant (p > 0.05).

Description of the evolutions by group: little variation is observed in the use of the different types of VLS. As illustrated in Figure 5.11, CLIL learners increase their use of 'lexical analysis', 'linking', and 'kinaesthetic' strategies and 'social strategies involving interaction with teachers'. Similarly, in the mainstream EFL group, increases are observed in the use of 'lexical analysis', 'mental imagery', 'kinaesthetic' and 'guessing from context' strategies.

Inferential analysis of the variations: the statistical analysis shows only a significant increase in the use of 'kinaesthetic' strategies (z = -1.211, p = 0.027) by CLIL learners, while no significant variation is found in the regular EFL group.
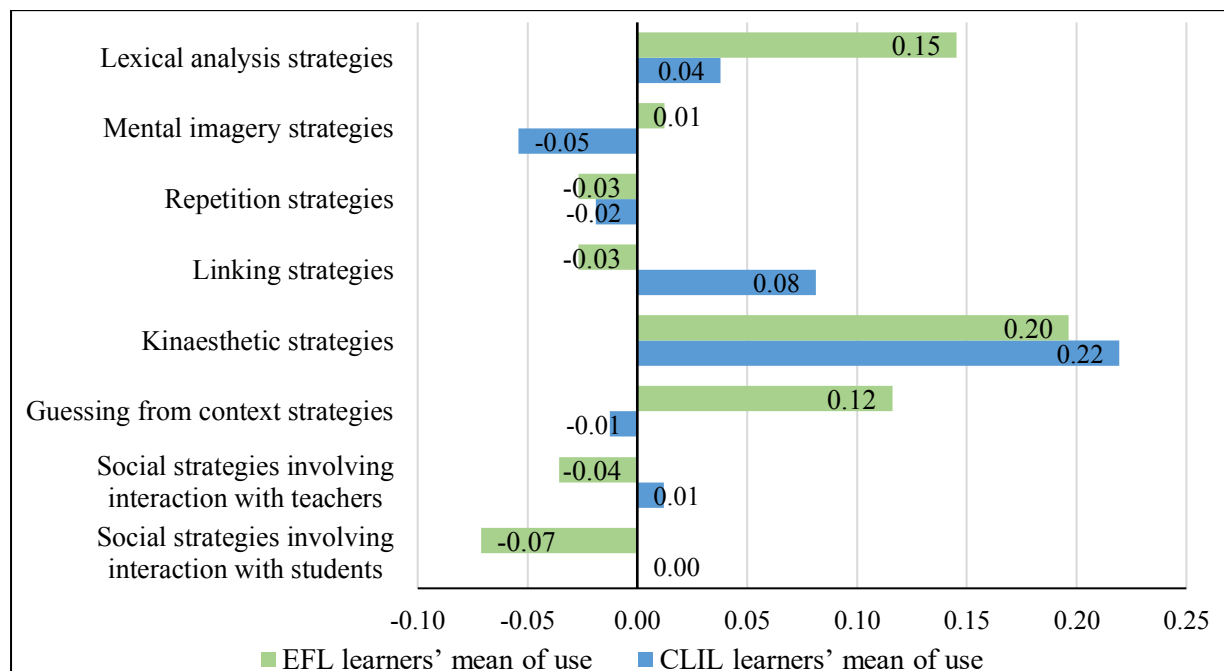
Figure 5.11. Variations in the selection of the different kinds of VLSs.

Individual use of VLSs

Differences between both groups at T1: CLIL and regular EFL learners differ in the evolution of the use of the following strategies: 'analysis of affixes and roots', 'connecting the word to its synonyms and antonyms', 'connecting the word to a personal experience', 'analysis of any available picture or gesture', and 'written repetition' strategies. Among them, the most remarkable difference is found in the evolution of the use of the 'analysis of the affixes and roots' (dif. = 0.30), 'written repetition' (dif. = 0.30) and 'verbal repetition' (dif. = 0.26) strategies.

Inferential analysis of the differences at T1: CLIL and regular EFL learners differ significantly in their use of 'analysis of affixes and roots' ($z = -4.351$, $p < 0.0004$), 'connecting the word to its synonyms and antonyms' ($z = -2.437$, $p = 0.015$), and 'use of English-language media' ($z = -3.960$, $p < 0.0004$). strategies.

Description of the evolution by group: both groups mainly share the same patterns of variation. In general, CLIL learners increase their use of a total of ten strategies at T2 (see Figure 5.12), whereas the other seven strategies decrease and the use of 'asking students for meaning' remains stable. As for regular EFL learners, they resort more frequently to ten strategies, whereas their selection of seven strategies decreases and the use of 'verbal repetition' and 'using a bilingual a dictionary' strategies is similar at both times.

Inferential analysis: concerning CLIL participants, the statistical analysis shows significant variations in the use of the strategies 'using English-language media' ($z = -3.074$, $p = 0.002$) and 'use of affixes and roots' ($z = -2.440$, $p = 0.015$). These changes are not homogeneous among the schools examined, but, concerning the strategy 'using English-language media', its use increases significantly only in school 2 ($z = -3.051$, $p = 0.002$), and the increased use of the strategy 'use of affixes and roots' only takes place in school 1 ($z = -2.07$, $p = 0.038$).

For their part, in the case of EFL learners, only the use of the strategies 'using physical action when learning a word' ($z = -2.214$, $p = 0.027$), 'written repetition' ($z = -2.915$, $p = 0.044$) and 'using English-language media' ($z = -2.587$, $p = 0.010$) increases significantly. However, these increments are no widespread among the learners in the different schools: in the case of the strategies 'written repetition' ($z = -2.150$, $p = 0.032$) and 'use of English-language media' ($z = -1.952$, $p = 0.051$), their use grows significantly only in school 1, whereas the use of the strategy 'using physical action when learning a word' has a significant increase only in school 4 ($z = -2.138$, $p = 0.033$).

Inferential analysis of the differences between groups: the only significant difference in the variations between both groups is found in the 'analysis of the affixes and roots' strategy ($z = -2.378$, $p = 0.017$), as EFL participants increases its use significantly more than the CLIL group at T2. Figure 5.13 shows CLIL and EFL participants' variations in their selection of strategies.
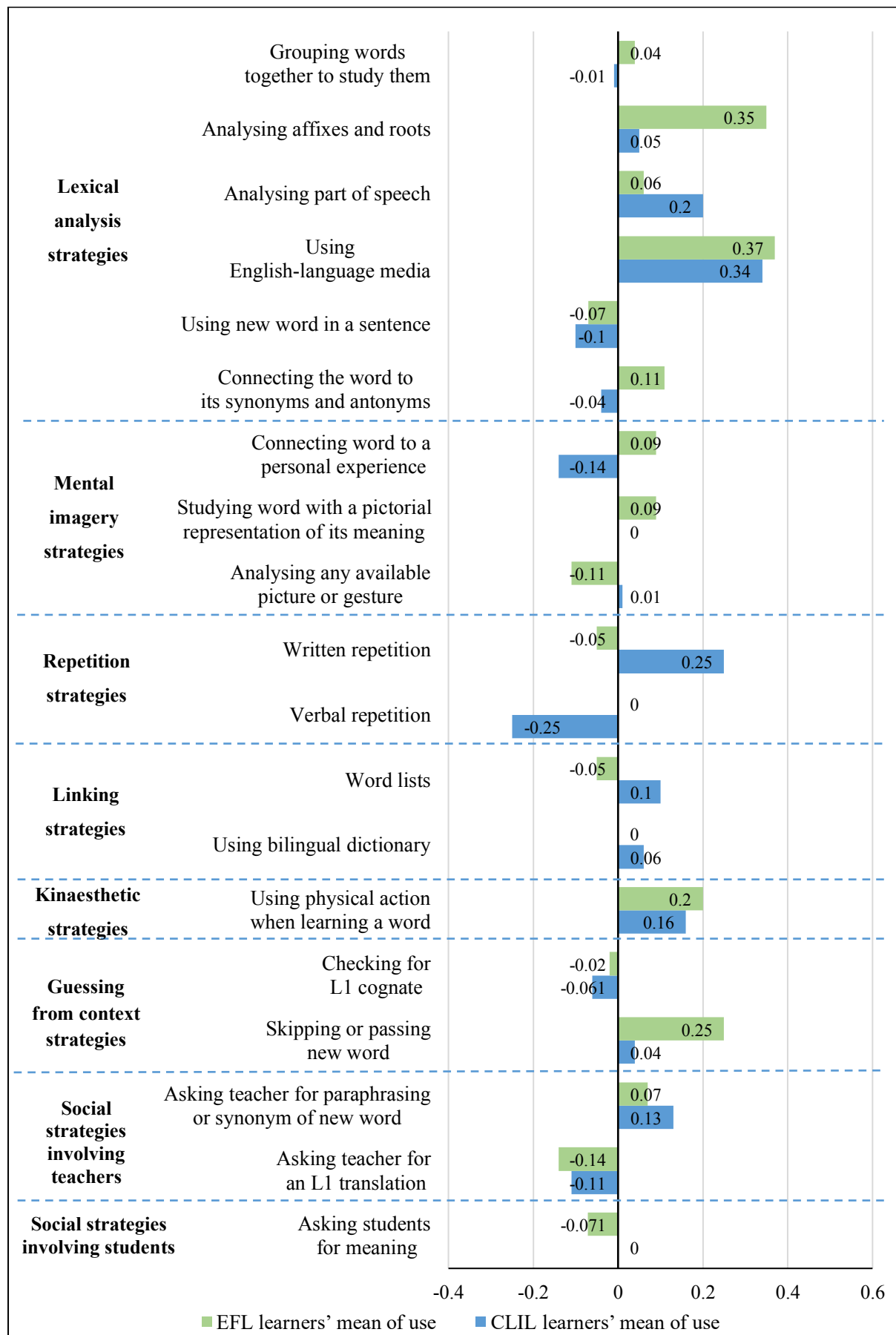
Figure 5.12. Variations in the selection of VLSs.

This section has explored differences between CLIL and regular EFL learners' choice of VLSs to understand how the implementation of a CLIL approach may affect the selection of VLSs. This analysis has identified several differences: for example, CLIL learners significantly use more frequently the strategy 'analysis of affixes and roots' at both T1 and T2. At the same time, mainstream EFL participants select significantly more often the strategy' use of word lists'. Interestingly, some of the strategies used more frequently by CLIL learners are positively related to receptive vocabulary acquisition, as shown in section 2.3. These findings will be discussed in the following chapter.

Several implications can be drawn from this analysis. Among other findings, the results obtained seem to show that CLIL students reflect more on language properties than mainstream EFL learners. In other words, CLIL learners make greater use of some strategies that seem to reflect metalinguistic awareness, such as 'analysis of part of speech', 'analysis of affixes and roots' and 'connecting the word to its synonyms and antonyms'. By contrast, mainstream EFL learners make greater use of repetition strategies. The use of repetition strategies has been proved to be not as beneficial when learning vocabulary (Castellano-Risco, 2018; Schmitt, 1997). Therefore, it seems that CLIL reduces the selection of those less-useful strategies. These and other findings will be discussed in depth in the following chapter.

4. IAoE: comparison among learners with different L2 exposure

In the previous chapter, four groups of participants were identified in my sample attending to the IAoE they had received. Among them, three are made up of CLIL learners, while the fourth comprises regular EFL learners. This section presents the results considering the IAoE learners have received and looking for differences among groups. To do so, the lexical knowledge and the selection of VLSs of these four groups (early CLIL [CLIL 1], standard CLIL [CLIL 2], late CLIL [CLIL 3] and EFL) is going to be compared.

4.1. Lexical knowledge

4.1.1. Vocabulary recognition: high-frequency terms

Data description: results show that the greater the amount of input, the more extensive the vocabulary knowledge (see Table 5.16). Early CLIL learners, who had joined the CLIL programmes in the 1st grade of Primary Education, present the most extensive vocabulary knowledge of the 2K band, whereas the group with the lowest exposure to English, i.e., the mainstream EFL group, shows the lowest recognition rate.

Table 5.16

2K VLT results according to IAoE

	N	IAoE	Raw data					In extrapolated values		% of participants scoring >26
			No. of items	Min	Max	Mean	SD	No. of items	Mean	
Early CLIL	23	3,332	30	17	30	24.96	3.51	2,000	1,664	50
Standard CLIL	25	2,715	30	12	30	24.68	3.90	2,000	1,645	48
Late CLIL	34	2,315	30	11	29	22.82	4.43	2,000	1,521	23.53
Regular EFL	56	1,332	30	2	27	16.09	6.58	2,000	1,042	8.89

This trend is also observed when comparing the percentage of participants mastering the 2K band. As shown in Table 5.16, half of the early CLIL participants and nearly half of standard CLIL learners master the 2K band receptively. However, this figure falls to 23.53% in the case of late CLIL participants and to less than 10% in the regular EFL group. In light of these results, it is observed that those participants with the largest amount of L2 exposure are those who more frequently master the band, although differences are nearly negligible in the case of early and standard CLIL participants.

Frequency distribution: as represented in Figure 5.13, the distribution of the three CLIL groups is quite similar. They present a relatively homogeneous shape, and none of the groups presents results lower than 30% of the band. However, two main discrepancies are observed: (1) late CLIL participants show a greater percentage of the population scoring lower than 70% when compared to the other two groups; and (2) the modal class interval is found in the 22-24 interval for standard and late CLIL learners, but it is in the 25-27 interval for early CLIL learners. As for the distribution of the regular EFL group, it differs greatly from that of the other three groups: the modal class interval is found in the 19-21 interval, and it presents a higher percentage of the population scoring less than 50% of the items in comparison to the CLIL groups.

	[0-3]	[4-6]	[7-9]	[10-12]	[13-15]	[16-18]	[19-21]	[22-24]	[25-27]	[28-30]
■ CLIL 1	0.00%	0.00%	0.00%	0.00%	0.00%	8.70%	13.04%	17.39%	39.13%	21.74%
■ CLIL 2	0.00%	0.00%	0.00%	0.00%	4.00%	0.00%	12.00%	32.00%	28.00%	24.00%
■ CLIL 3	0.00%	0.00%	0.00%	5.80%	2.90%	5.80%	14.70%	29.40%	26.40%	14.70%
■ EFL	1.79%	5.36%	7.14%	8.93%	14.29%	17.86%	19.64%	8.93%	7.14%	8.93%
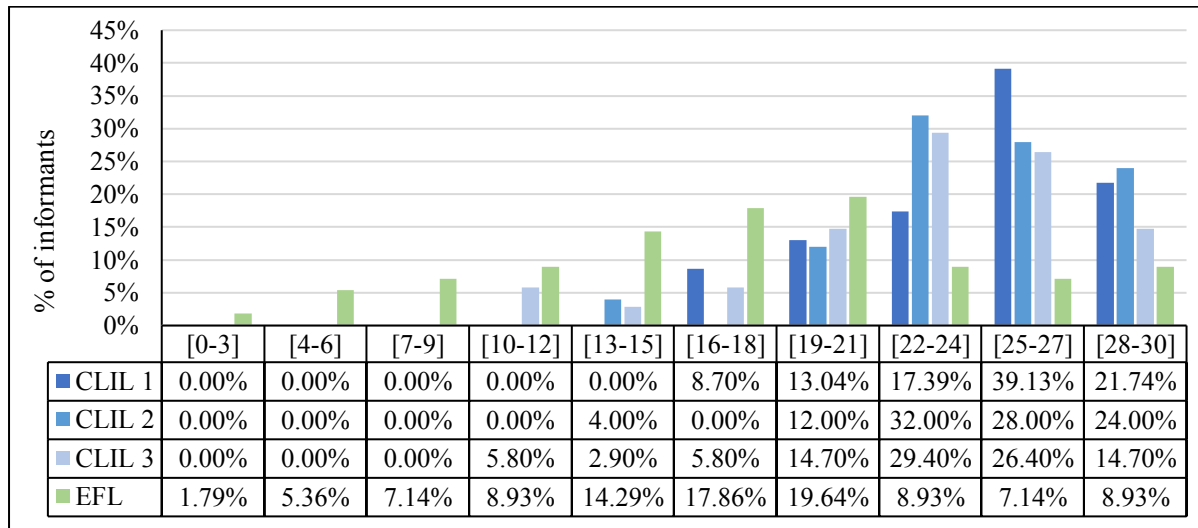
Figure 5.13. Frequency distribution of CLIL subgroups' 2K VLT results.

Inferential analysis: in order to determine whether the differences between CLIL groups with different exposure to English are statistically significant, a *t*-test is carried out and the p-values obtained from multiple comparisons are adjusted using the False Discovery Rate (FDR) method. Non-significant differences are found when comparing early and standard CLIL groups ($t(46) = 0.258$, $p = 0.797$, $d = 0.07$), standard and late CLIL groups ($t(57) = -1.677$, $p = 0.10$, $d = -0.44$) and early and late CLIL groups ($t(55) = -1.939$, $p = 0.058$, $d = -0.53$). In contrast, the analysis shows significant differences between early CLIL and regular EFL learners ($t(71.676) = -7.749$, $p < 0.0004$, $d = 1.51$), standard CLIL and regular EFL learners ($t(72.608) = 7.304$, $p < 0.0004$, $d = 1.42$) and late CLIL and regular EFL learners ($t(87) = -5.788$, $p < 0.0004$, $d = 1.03$) are statistically significant in favour of the CLIL groups.

4.1.2. Vocabulary recognition: academic terms

Data description: Table 5.17 illustrates that the participants who present the highest mean score are not those with the most substantial CLIL experience (i.e., early CLIL learners), but those who had started CLIL in 4th to 6th grades of primary education.

This analysis of the frequency distribution links with the percentage of learners mastering the academic band. In this case, unlike in the analysis of the mean score, there is a direct relation to foreign language exposure. Learners with the most extensive experience in CLIL master the band more frequently than other learners less exposed to CLIL. Concretely, 39.19% of early CLIL learners show a full receptive mastery of the academic band whereas this figure goes down to 24%, 26.47% and 5.36% in standard CLIL, late CLIL and regular EFL groups respectively.

Table 5.17

Academic VLT results according to IAoE

| | N | IAoE | Raw data | | | | | In extrapolated values | | % of participants scoring >26 |
			No. of items	Min	Max	Mean	SD	No. of items	Mean	
Early CLIL	23	3,332	30	0	30	22.74	6.98	570	432	39.19
Standard CLIL	25	2,715	30	15	30	23.36	4.21	570	444	24
Late CLIL	34	2,315	30	12	30	21.50	4.87	570	408	26.47
Regular EFL	56	1,332	30	0	29	13.11	7.84	570	249	5.36

Frequency distribution: none of the groups shares the modal class interval; nearly a third of early CLIL learners' results are in the 28-30 interval, and another third of the population is in the 19-21 interval (see Figure 5.14). As for the standard CLIL group, these participants centralise most of their results in the 22-24 and 19-21 intervals. For their part, the late CLIL group has nearly a fourth (24%) of their participants ranging from 19 to 21. Finally, in the regular EFL group, the modal class interval is situated in the 13-15 interval.

Besides, as IAoE increases, so does the percentage of participants in the last interval (28-30). In this sense, whereas 30.4% of the early CLIL learners know more than 90% of the academic items, this percentage decreases to 20%, 14.70% and 3.57% in standard CLIL, late CLIL, and regular EFL groups respectively.

	[0-3]	[4-6]	[7-9]	[10-12]	[13-15]	[16-18]	[19-21]	[22-24]	[25-27]	[28-30]
CLIL 1	4.30%	0.00%	0.00%	4.30%	0.00%	4.30%	30.40%	4.30%	21.70%	30.40%
CLIL 2	0.00%	0.00%	0.00%	0.00%	4.00%	8.00%	24.00%	24.00%	20.00%	20.00%
CLIL 3	0.00%	0.00%	0.00%	2.90%	8.80%	20.60%	23.60%	8.80%	20.60%	14.70%
EFL	14.29%	3.57%	8.93%	12.50%	16.07%	14.29%	10.71%	7.14%	8.93%	3.57%
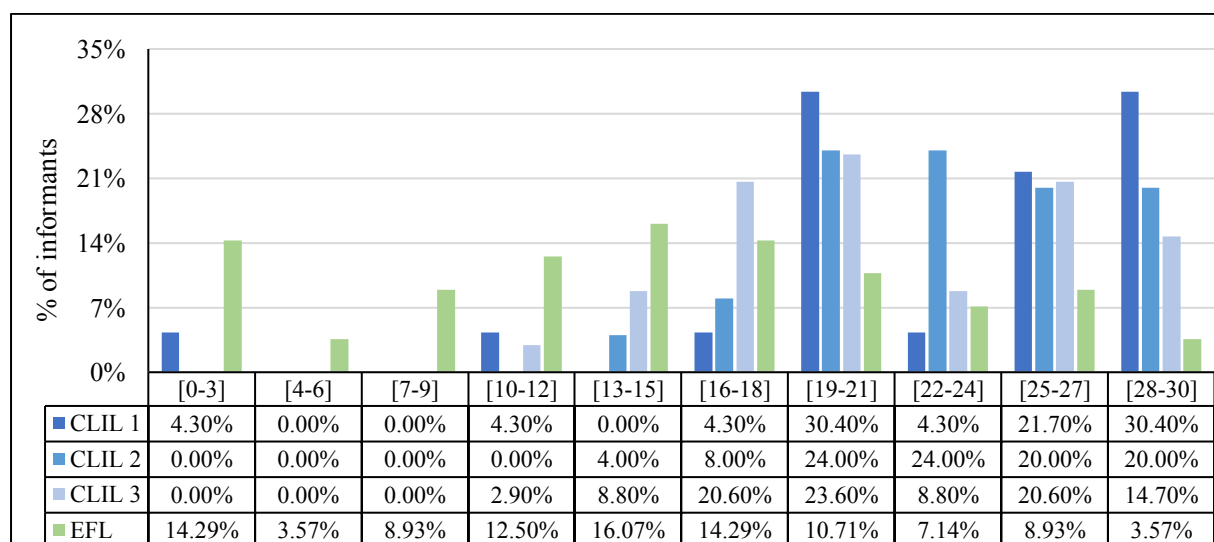
Figure 5.14. Frequency distribution of CLIL subgroups' academic VLT results.

Inferential analysis: non-significant differences are found when early and standard ($t(46) = 0.374$, $p = 0.710$, $d = 0.11$), standard and late ($t(57) = -1.53$, $p = 0.132$, $d = 0.40$) and early and late CLIL

groups ($t(55) = -0.791$, $p = 0.432$, $d = 0.20$) are compared. In contrast, the comparisons of CLIL and mainstream EFL groups' rates of recognition of the academic band yield statistically significant differences between early CLIL and mainstream EFL groups ($t(77) = -5.115$, $p < 0.0004$, $d = 1.30$), standard CLIL and regular EFL participants ($t(79) = -6.137$, $p < 0.0004$, $d = 1.63$) and late CLIL and mainstream EFL learners ($t(87.93) = -6.264$, $p < 0.0004$, $d = 1.29$).

4.1.3. Vocabulary recall: high-frequency terms

Data description: Table 5.18 presents the four groups' results in the 2K band of the PVLT. As can be observed, the learners with the most extensive exposure to English show the best results. As for the remaining groups, the other two CLIL groups —standard and late CLIL— obtain similar scores while the regular EFL group's results are well below. In extrapolated values, early CLIL participants have reached the 1K band, whereas participants from the other three groups do not reach this point.

Table 5.18

2K PVLT results according to IAoE

| | N | IAoE | Raw data | | | | | In extrapolated values | | % of participants scoring >26 |
			No. of items	Min	Max	Mean	SD	No. of items	Mean	
Early CLIL	23	3,332	18	3	16	9.82	3.20	2,000	1,092	8.69
Standard CLIL	25	2,715	18	0	16	8	4.14	2,000	889	4
Late CLIL	34	2,315	18	0	14	8.06	3.44	2,000	895	-
Regular EFL	56	1,332	18	0	12	4	3.17	2,000	444	-

Frequency distribution: the difference between the early CLIL participants and the rest of groups is also observed in the analysis of the frequency distribution of the results (Figure 5.15): nearly 70% of the early CLIL population are able to produce more than 50% of the words, whereas, in the remaining groups, about half of the groups do not hit 50% of the answers. Besides, as displayed in previous sections, there is a glaring difference between the regular EFL group and the other groups: the distribution of the EFL results shows a clear decreasing pattern, with no population recalling more than 12 words, and illustrate that their results are quite lower, as the modal class interval is found in the 1-2 and 2-3 intervals.

Finally, early CLIL learners' greater productive knowledge of the 2K band is also confirmed when exploring the percentage of participants mastering the 2K band, i.e., scoring higher than 86.66% of the test. Concretely, 8.69% of early CLIL learners demonstrate to master the 2K band productively. In contrast, only 4% in the case of standard CLIL participants and none of the late CLIL and EFL learners reaches the level required.

	[0-1]	[2-3]	[4-5]	[5-6]	[7-8]	[9-10]	[11-12]	[13-14]	[15-16]	[17-18]
CLIL 1	0.00%	4.35%	4.35%	13.04%	8.70%	26.09%	26.09%	8.70%	8.70%	0.00%
CLIL 2	8.00%	8.00%	8.00%	12.00%	12.00%	28.00%	4.00%	16.00%	4.00%	0.00%
CLIL 3	2.94%	2.94%	14.71%	29.41%	0.00%	23.53%	14.71%	11.76%	0.00%	0.00%
EFL	25.00%	25.00%	19.64%	16.07%	8.93%	0.00%	5.36%	0.00%	0.00%	0.00%
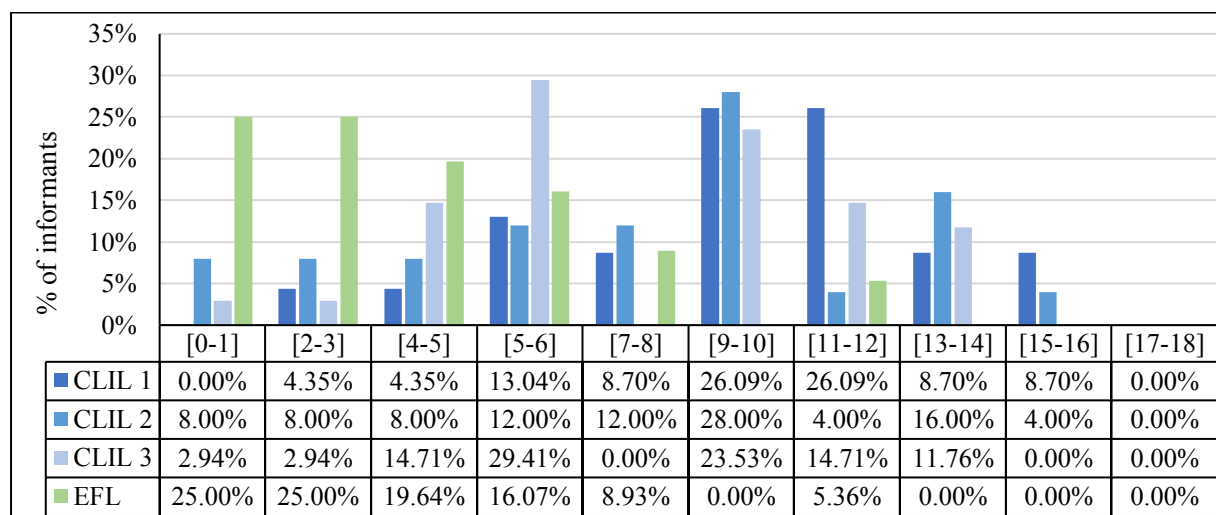
Figure 5.15. Frequency distribution of CLIL subgroups' 2K PVLT results.

Inferential analysis: in light of the results, there are no significant differences in the productive knowledge of the 2K band between early and standard CLIL (U(N_{CLIL1}=23, N_{CLIL2}=25)= 211.50, Z = -1.575, p = 0.143) groups, early and late CLIL (U(N_{CLIL1}=23, N_{CLIL2}=25)= 282.00, Z = -1.778, p = 0.075) participants or *CLIL 2* and *CLIL 3* (U(N_{CLIL1}=23, N_{CLIL2}=25)= 417.50, Z = -0.115, p = 0.953) groups. However, there are significant differences when comparing early (U(N_{CLIL1}=23, N_{EFL}=56)= 143.50, Z = -5.413, p = 0.000), standard (U(N_{CLIL2}=25, N_{EFL}=56)= 325.50, Z = -3.838, p = 0.000) and late (U(N_{CLIL3}=34, N_{EFL}=56)= 375.00, Z = -4.814, p = 0.000) CLIL groups to the mainstream EFL group. This result supports previous outcomes in which no differences within CLIL learners regarding receptive knowledge of vocabulary are found, no matter their IAoE, but with significant differences when comparing each of the three subgroups' results with the mainstream EFL learners.

4.1.4. Vocabulary recall: academic terms

Data description: similar to previous CLIL sub-groups comparisons, early CLIL participants perform better than the other three groups (see table 5.19) and, again, it is the EFL group which shows the lowest results. However, an increased IAoE does not seem to systematically result in a more extensive productive knowledge of the academic band, as those learners with the least experience in CLIL —late CLIL participants— have a higher mean score in comparison to those learners who joined the CLIL experiences in 4th, 5th or 6th grade of Primary Education —standard CLIL participants.

Table 5.19

Academic PVLT results according to IAoE

	N	IAoE	Raw data					In extrapolated values		% of participants scoring >26
			No. of items	Min	Max	Mean	SD	No. of items	Mean	
Early CLIL	23	3,332	18	2	13	6.78	3.20	836	315	-
Standard CLIL	25	2,715	18	0	10	5	2.69	836	247	-
Late CLIL	34	2,315	18	0	9	5.32	2.18	836	232	-
Regular EFL	56	1,332	18	0	8	2.05	2.31	836	95	-

Frequency distribution: the first thing that attracts one's attention is that the four graphs are skewed to the right, i.e., the tail of the distribution on the right-hand side is longer than on the left-hand side (see Figure 5.16). However, despite the similarity observed at first glance, the groups present some discrepancies among them. The early CLIL group has the largest percentage of the population scoring higher than 50% (21.74%), whereas, in the other groups, this percentage does not reach 5% of the total population. Moreover, the four clusters do not share the same modal class interval, which is, for the early and late CLIL groups, the 5-6 interval, for standard learners, the 4-5 interval, and for the regular EFL group is the 0-1 interval.

	[0-1]	[2-3]	[4-5]	[5-6]	[7-8]	[9-10]	[11-12]	[13-14]	[15-16]	[17-18]
CLIL 1	0.00%	17.39%	21.74%	26.09%	0.00%	17.39%	13.04%	4.35%	0.00%	0.00%
CLIL 2	12.00%	16.00%	28.00%	24.00%	16.00%	4.00%	0.00%	0.00%	0.00%	0.00%
CLIL 3	5.88%	14.71%	23.53%	38.24%	14.71%	2.94%	0.00%	0.00%	0.00%	0.00%
EFL	50.00%	23.21%	12.50%	12.50%	1.79%	0.00%	0.00%	0.00%	0.00%	0.00%
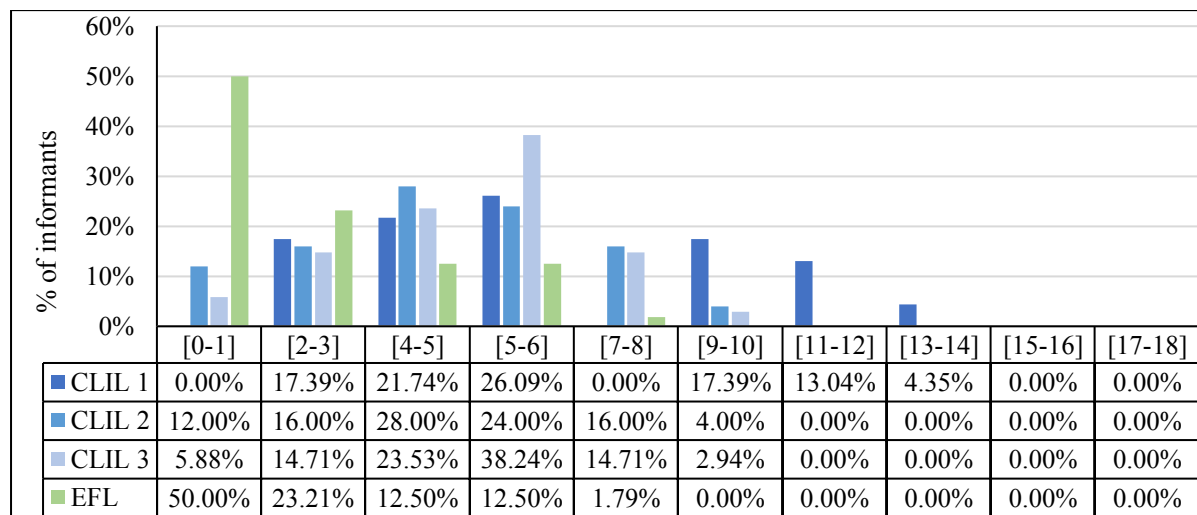
Figure 5.16. Frequency distribution of CLIL subgroups' academic PVLT results.

Statistical analysis: on the one hand, there are no significant differences between the productive knowledge of the academic band neither between early and standard (U(N_{CLIL1}= 23, N_{CLIL2}= 25)= 208.50, Z = -1.634, p = 0.103) CLIL groups, nor early and late (U(N_{CLIL2}= 25, N_{CLIL3}= 34)= 384.00, Z = -0.631, p = 0.528) CLIL participants nor standard and late (U(N_{CLIL1}= 23, N_{CLIL3}= 34)= 299.00, Z = -1.501, p = 0.133) CLIL groups. On the other hand, in comparison to the EFL group, early (U(N_{CLIL1}=23, N_{EFL}=56)= 144.00, Z = -5.5462, p = 0.000), standard (U(N_{CLIL2}=25, N_{EFL}=56)= 303.50,

$Z = -4.108, p = 0.000$), and late ($U(N_{CLIL3}=34, N_{EFL}=56)= 321.50, Z = -5.306, p = 0.000$) CLIL learners present a significant larger productive knowledge of the academic band.

4.2. Selection of vocabulary learning strategies

In Chapter Three, certain factors that could influence the selection of language learning strategies were highlighted, among which the type of instruction and the duration of the L2 learning process were included. This section presents the results related to the choice of VLSs by learners who differ in the intensity of their exposure to English input.

Overall use of VLSs

Data description: the data shows an expected pattern of increased VLSs use as IAoE increases. In other words, the early CLIL group makes greater use of VLSs ($\bar{x} = 2.64$) in comparison to the standard CLIL ($\bar{x} = 2.51$), late CLIL ($\bar{x} = 2.48$) and regular EFL ($\bar{x} = 2.50$) groups.

Statistical analysis: differences among the different groups are not significant ($p. > 0.05$).

Use of the different kinds of VLSs

Data description: a clear pattern of increase or decrease of VLSs use as the IAoE varies is not observed. On several occasions, CLIL learners who differ mostly on their exposure to the FL show a similar use of the different kinds of VLSs, whereas the use VLSs by CLIL learners with more similar exposure to English differ in greater proportion. All in all, as with the analysis of the overall VLSs use, the least widely used strategies correspond to the 'kinaesthetic' group and is common to the four learning groups. On the contrary, the three CLIL groups do not share their preferred strategy cluster: for the early and late CLIL groups, the 'lexical analysis' strategies are preferred, while for the standard CLIL group the 'guessing from context' group is the preferred one and for the EFL group the preferred strategies are the 'linking' ones.

Statistical analysis: the analysis of the strategies used by the different groups of learners shows non-significant results, either among the CLIL subgroups or among the three CLIL sub-groups and the EFL one.
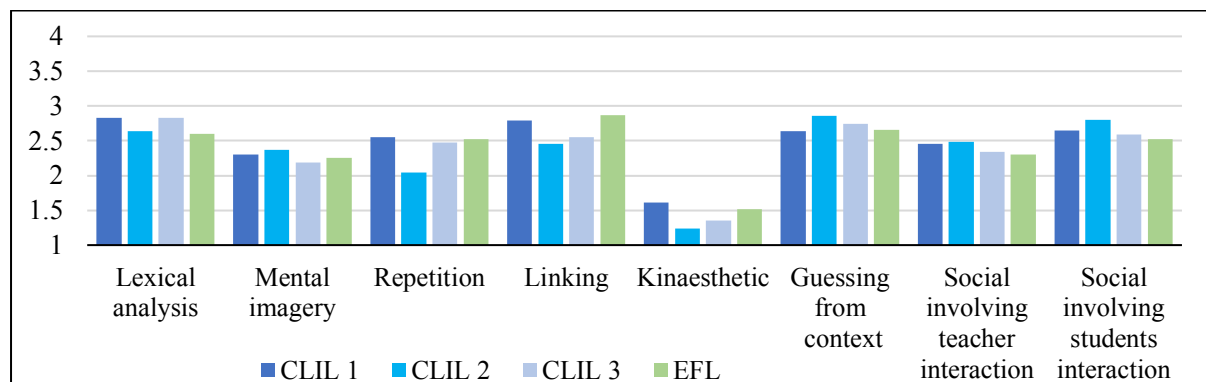
Figure 5.17. Mean use of each strategies cluster.

Individual use of VLSs

Data description: given the difference in IAoE, the strategies use among groups was expected to be different. However, as with the overall use of VLSs, CLIL subgroups make similar use of VLSs, being the similarities between early and standard CLIL groups especially outstanding (see Figure 5.18).

Statistical analysis: no statistically significant ($p > 0.05$) differences are found between early and standard CLIL learners. In contrast, this is not the case for late CLIL learners, who present significant differences compared to the other two groups. In comparison to early CLIL learners, late CLIL participant make significant lower use of 'asking the teacher for an L1 translation' ($U(N_{CLIL1}=21$, $N_{CLIL3}=34)= 815.00$, $Z = -2.983$, $p = 0.003$) and 'using of bilingual dictionary' ($U(N_{CLIL1}=23$, $N_{CLIL3}=34)= 854.00$, $Z = -2.270$, $p = 0.023$). Differences between standard and late CLIL groups are also found in the selection of the strategy 'analysis of the part of speech' ($U(N_{CLIL2}=25$, $N_{CLIL3}=34)= 893.00$, $Z = -2.046$, $p = 0.041$), which is more selected by standard CLIL learners.

As for the comparison between the CLIL groups and the EFL one, early CLIL learners make significant greater use of the strategies 'analysis of affixes and roots' ($U(N_{CLIL1}=23$, $N_{EFL}=56)= 419.500$, $Z = -2.573$, $p = 0.010$), 'use of a bilingual dictionary' ($U(N_{CLIL1}=21$, $N_{EFL}=56)= 468.00$, $Z = -2.122$, $p = 0.046$), and 'asking the teacher for an L1 translation' ($U(N_{CLIL1}=21$, $N_{EFL}=56)= 459.50$, $Z = -2.122$, $p = 0.034$). For their part, standard CLIL learners select the strategy 'analysis of the part of speech' ($U(N_{CLIL2}=25$, $N_{EFL}=56)= 462.00$, $Z = -2.554$, $p = 0.011$) significantly more often. Finally, the comparison between late CLIL and mainstream EFL learners' results shows that the strategy 'use of English-language media' ($U(N_{CLIL3}=34$, $N_{EFL}=56)= 692.00$, $Z = -2.279$, $p = 0.023$) is significantly more selected by CLIL learners, whereas mainstream EFL learners resort to the strategy 'use of word lists' ($U(N_{CLIL3}=34$, $N_{EFL}=56)= 686.00$, $Z = -2.396$, $p = 0.017$) significantly more frequently.
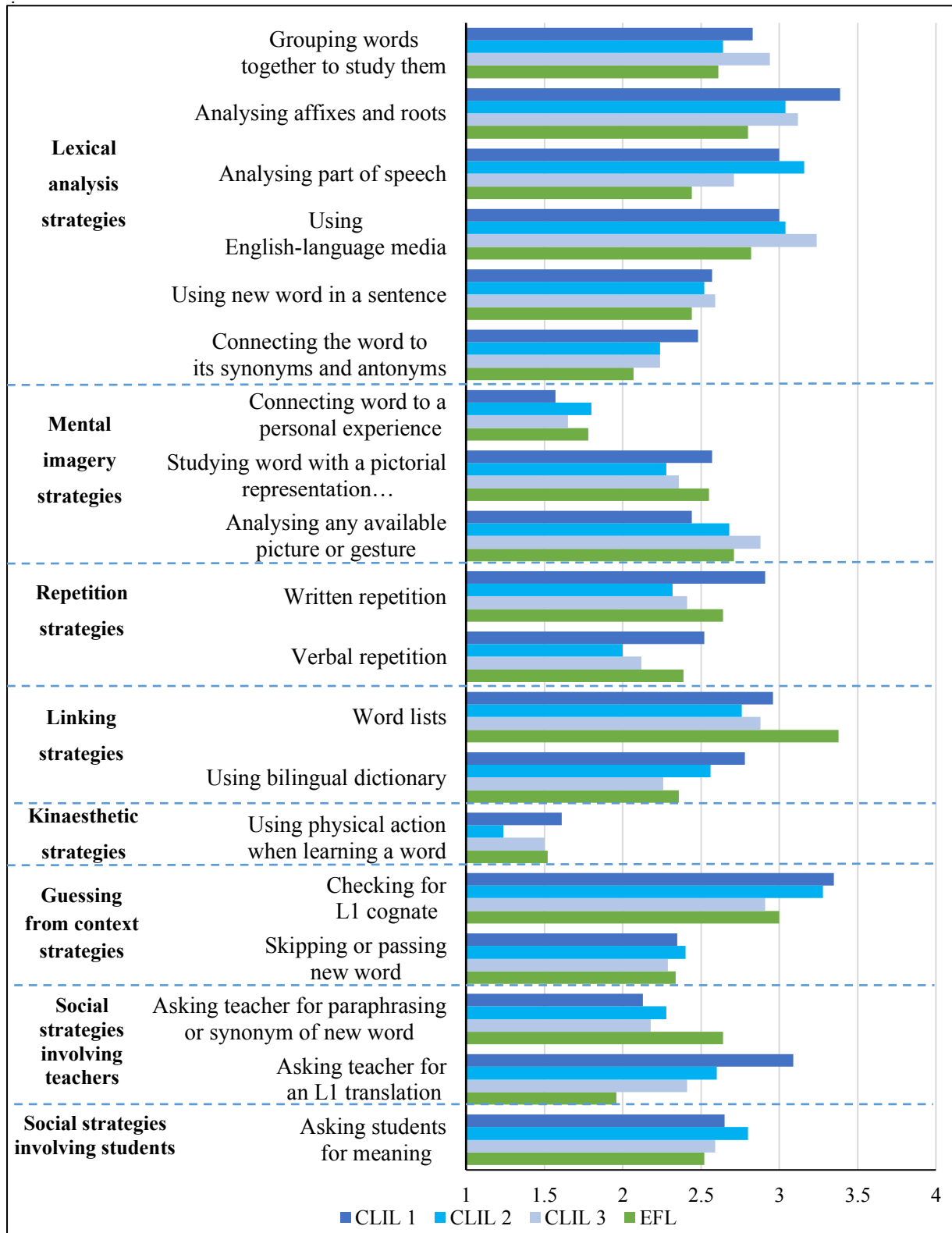
Figure 5.18. Mean use of VLSs specified in CLIL subgroups.

This chapter has reported the results of the study. In section one, the vocabulary learning experience of the participants has been detailed. Then, in section two, a comparison of CLIL and mainstream EFL learners' receptive and productive knowledge of the 2K and academic vocabulary

bands has been presented. After that, in section three, the selection of VLSs by CLIL and regular EFL learners have been examined and contrasted. These comparisons have shown that CLIL and mainstream EFL learners differ in their lexical knowledge and present different preferences concerning VLSs usage. Therefore, in section four, the impact of IAoE on the accounted differences between CLIL and regular EFL learners has been explored.

The analysis of the data has identified some patterns in both, vocabulary knowledge of the 2K and academic bands and VLSs usage. Starting with vocabulary knowledge, the data shows that learners at the end of Secondary Education are still in the process of mastering the most common 2K English words, which are the number of word families considered enough to engage in daily conversations. Similarly, they are still far from controlling the academic terms. Besides, their recalling ability is well below the reproduction one. Finally, significant vocabulary growth has also been checked to take less than an academic year. Moving on to the second aspect, i.e., the analysis of the selection of VLSs, in general, secondary-school learners have been shown to make greater use of VLSs related to the lexical analysis of the words and written repetition. In contrast, they are quite reluctant to resort to movement to retain the meanings of new lexis. The two variables under study have been related, and, as a result, the selection of some VLSs has been found to have a positive impact on the participants' vocabulary knowledge.

After that, differences between L2 learners following two different educational approaches — CLIL and EFL— have been examined. Focussing, first, on the differences in vocabulary knowledge, along the dimensions and bands explored, on the one hand, a better performance by CLIL participants has recurrently been observed, at the same time that, on the other hand, a more extensive receptive vocabulary growth by mainstream EFL learners has been documented. Concerning the selection of VLSs, some significant differences in their choice have been observed, being the most relevant fact that, in general, CLIL learners tend to make greater use of those strategies distinguished for having a positive impact on vocabulary knowledge.

Traditionally, differences in vocabulary knowledge in favour of CLIL participants have been attributed to variations in AoE. For this reason, the last section has been devoted to analysing the impact of this variable on vocabulary learning. In general, the data has shown that those learners with greater exposure to English obtain better results in both the VLT and PVLT. However, in none of the bands measured, this difference is great enough to be considered significant. Therefore, it cannot be assumed that a larger IAoE is directly correlated to a more extensive knowledge of vocabulary. Still, other factors could be interfering in the vocabulary learning process.

Once the data has been treated, the following chapter will deal with the interpretation of these findings and their comparison to other studies with similar characteristics.

CHAPTER SIX:
DISCUSSION

1. Introduction

The main goal of the present thesis is to analyse the impact of some input-related factors on secondary-school learners' vocabulary learning. Thus, it is structured around the analysis of how two factors, namely, language teaching approach (CLIL and mainstream EFL approaches) and IAoE, affect the development of lexical competence, specifically with respect to two elements: the selection of VLSs and the knowledge of high-frequency (2K) and academic terms.

Since the emergence of CLIL in the last decade of the 20th century, there has been a vast body of research investigating the presumed linguistic benefits of CLIL over traditional EFL approaches (Agustín-Llach, 2009; Dalton-Puffer, 2007; Hüttner & Rieder-Bünemann, 2010; Lorenzo, Casal & Moore, 2009; Lorenzo & Rodríguez, 2014; Martínez Agudo, 2020; Ruiz de Zarobe, 2008; Xanthou, 2010, 2011; Sylvén, 2019). Nevertheless, in my view, there is another issue whose exploration has been neglected: how CLIL itself, as an educational approach, affects L2 processing. The implementation of CLIL involves bringing to the content class a set of teaching methods and principles that promote learners' active cognitive engagement. The L2 learning context in CLIL is different from a more traditional EFL class: on the one hand, CLIL is a content-based approach, in which the focus is on the subject matter and on how it is conveyed. On the other hand, regular EFL approaches are instruction-based, that is, the content of the EFL subject is the foreign language *per se*, and its main aim is the appropriate development of L2 communicative competence. This radical difference between the approaches may result in changes in L2 processing, whose exploration may be essential to have a complete view of the novelties of CLIL.

Focussing on the field of interest of this dissertation —lexical competence—, most research has concentrated on the effects of CLIL on lexical knowledge (Agustín-Llach & Canga Alonso, 2016;

Arribas, 2016; Castellano-Risco, 2018a; Jiménez Catalán & Ruiz de Zarobe, 2009; Sylvén, 2019; Xanthou, 2010, 2011), while other components of lexical competence, such as lexical processing have been disregarded. In this respect, this PhD dissertation was planned from a holistic perspective, aiming to explore the effect of CLIL not only on lexical knowledge, as had been traditionally done, but also on the second component of lexical competence: lexical processing. In my view, this inclusion may provide more detailed insights about how CLIL affects language learning and may help researchers determine more clearly the effects of this educational approach.

Besides, this study includes a second element of analysis: the role of the IAoE on lexical competence. The greater exposure is usually regarded as one of the main reasons for CLIL success. However, to the best of my knowledge, there is no research empirically proving that the reason why CLIL learners show a better lexical command is exclusively that they are exposed to a greater amount of input. This piece of research was designed to shed some light on this issue by exploring the differences among CLIL learners with different amounts of exposure to English.

2. Lexical knowledge and language teaching approach

The importance, both in quantitative and in qualitative terms, of CLIL experiences in Spain (Eurydice, 2017) and, more generally, in Europe, has raised the question of whether students can cope with the language used in class and, if so, the extent to which its implementation has really contributed to improving learners' proficiency in the L2 when compared to their EFL peers. The literature on CLIL (Agustín-Llach, 2009; Hüttner & Rieder-Bünemann, 2010; Lorenzo et al., 2009; Lorenzo & Rodríguez, 2014; Martínez Agudo, 2020; Ruiz de Zarobe, 2008; Xanthou, 2010, 2011; Sylvén, 2019) has already demonstrated some of its advantages concerning various linguistic aspects over more traditional EFL approaches. The particular focus of the present thesis was to concentrate on vocabulary and to analyse the putative benefits that a CLIL approach may bring to L2 students. This analysis obviously needed to start by considering the recognition of the form and meaning of high-frequency words. Vocabulary research has repeatedly shown (Matthews & Cheng, 2015; Yen Dang, 2020; Yen Dang, Webb & Coxhead, 2020) that a solid foundation in managing the 2K band is necessary to be able to use language in different contexts and with different purposes. For this reason, the present study was designed to explore not only the receptive knowledge of general vocabulary but also learners' ability to recall high-frequency (2K) words.

The particular demands that CLIL places on learners are also related to understanding and expressing disciplinary content through the L2 (Coyle et al., 2010; Llinares et al., 2012). Studies suggest that CLIL promotes the development of academic language skills (Dalton-Puffer, 2007; Lorenzo & Rodríguez, 2014; Nikula, 2007; Nightingale & Safont, 2019; Pascual Peña, 2015; Yi Lu & Jeong, 2018) to a greater extent than more traditional language approaches. Appropriate development of academic

language skills should imply the management of specific lexis related to the articulation of the discourse sufficient to comprehend and express the content being taught (the so-called "language OF learning" in CLIL; Coyle, 2007, 2010). Nevertheless, to the best of my knowledge, little research (Merikivi & Pietilä, 2014; Olsson, 2015) has been carried out to ascertain whether academic vocabulary is also further developed in CLIL contexts. For this reason, in addition to the traditional analysis of the knowledge of high-frequency terms, this study was planned to measure also the recognition and production of academic terms. Section two of Chapter Five presented the results related to CLIL and EFL participants' differences in lexical knowledge.

Starting first with the knowledge of the 2K band, the exploration of learners' receptive knowledge of this band indicates that, by the end of CSE, learners identify a total of 1,378 words out of the 2K most frequent terms. Results also reveal a significant difference between both groups: CLIL learners, who recognise about 1,600 words, and EFL learners, who only know receptively about 1,000 words. This difference has been acknowledged at both T1 and T2, although the gap has been reduced in extrapolated values.

In practice, this means that CLIL learners are relatively close to mastering the 2K band, and, consequently, to controlling 80% of the running words in any academic text (Schmitt, 2010). Therefore, these learners would be expected to have little difficulty in reading academic texts, especially if they are instructed to use some kind of compensation strategies to infer the meaning of unknown words. In contrast, the EFL group is likely to have many problems when facing the same reading tasks, as they only recognise approximately 40% of any piece of academic writing.

These findings present some discrepancies with other studies exploring the receptive knowledge of high-frequency lexical items. In the case of mainstream EFL learners' receptive knowledge, the findings are in line with Canga Alonso(2015a), but they are substantially lower than the ones described for EFL learners (Canga Alonso & Arribas, 2014, Sylvén, 2019). As for the CLIL learners in our sample, their receptive knowledge is higher than the one reported in other studies carried out in Spain (1607 words in the present study vs 1300 words in Arribas, 2016) and Finland (1,607 word families in the present study vs 841 items in Merikivi & Pietilä, 2014). Still, these results are considerably lower than the 2K-band recognition rates of Swedish CLIL learners, where learners recognise, in extrapolated values, an average of 1,933 words out of the 2K most frequent English items (Sylvén, 2019). In this respect, and in the absence of further analysis, it may be argued that the learners' L1 may explain these differences. Unlike Spanish and Finish, Swedish is a Germanic language, and it is said that cognates are more likely to occur between related languages. The plausibly closer resemblance between Swedish and English and the higher presence of cognates may help Swedish learners recognise a more significant number of high-frequency English words.

On the other hand, the findings of the present study confirm previous research into the differences between CLIL and EFL learners' knowledge (Agustín-Llach, 2012; Agustín-Llach & Canga Alonso, 2016; Arribas, 2016; Canga Alonso, 2015a; Castellano-Risco, 2018a; Jiménez-Catalán & Ruiz de Zarobe, 2009; Sylvén, 2019; Xanthou, 2010, 2011).

The present study makes a significant contribution to the field by providing evidence that 1) quantifies the size of such difference, and 2) confirms that the difference between CLIL and EFL remains at the end of CSE. Regarding (1), the present dissertation is one of the first studies in this field that statistically proves that the difference between CLIL and EFL learners' receptive knowledge of the 2K band is large. As for (2), this study provides data showing that the difference in recognition of the 2K band is consistent in time: CLIL learners outperform EFL learners both at T1 and T2, although EFL learners' recognition rate increases in greater proportions.

Regarding the productive knowledge of the band, traditionally, CLIL learners show a better lexical production command (Canga Alonso & Arribas, 2014; Celaya and Ruiz de Zarobe, 2008; Gallardo del Puerto & Gómez-Lacabez, 2016; Merikivi & Pietilä, 2014) than do regular EFL learners. Most studies (Celaya and Ruiz de Zarobe, 2008; Gallardo del Puerto and Gómez-Lacabez, 2016; Olsson & Sylvén, 2019) have used written assignments to collect the data, and, as a result, have focussed on learners' overall lexical profiles, rather than an analysis of the production of the lexis included in specific bands. In this respect, to the best of my knowledge, few studies have addressed differences in CLIL and EFL learners' productive knowledge with the use of specific productive vocabulary tests exclusively (Canga Alonso & Arribas, 2014; Merikivi & Pietilä, 2014). The present dissertation contributes to this analysis.

In general, the results of this study indicate that learners can recall an average of 601 words out of the 2K most frequent ones. Nevertheless, as with the receptive knowledge of the band, considerable differences have been found between educational approaches. CLIL learners can productively recall nearly 1,000 lexical items, in contrast to EFL learners, who can recall about 400 items.

In practice, these results suggest that learners, regardless of their learning context, are far from being able to write pieces of academic texts without support, as they are still unable to recall a significant amount of the most basic vocabulary. CLIL learners' productive knowledge of high-frequency terms is higher than that reported in previous studies carried out in Spain (948 items in the present study vs 813 word families in Canga Alonso & Arribas, 2014) or Finland (948 word families in the present study vs 646 items in Merikivi & Pietilä, 2014), whereas the productive knowledge of the EFL group is well below that reported by other studies using similar measuring tools and exploring samples having similar characteristics (444 word families vs 640 in Canga Alonso & Arribas, 2014; 645 Moreno Espinosa, 2010; or 499 in Merikivi & Pietilä, 2014). In this respect, this research is in line with previous studies (Merikivi & Pietilä, 2014; Sylvén, 2019) that explored the differences between CLIL and EFL learner's productive knowledge of the 2K band and complements previous analyses with the quantification of

the size of the difference. This study concludes that not only CLIL and EFL learners' recalling capacity differs significantly, but that the difference between both groups is statistically regarded as large.

With respect to knowledge of the 2K band, my findings are in line with previous research into the differences between CLIL and regular EFL learners' receptive knowledge (Agustín-Llach, 2012; Agustín-Llach & Canga Alonso, 2016; Arribas, 2016; Canga Alonso, 2015a; Castellano-Risco, 2018a; Jiménez Catalán & Ruiz de Zarobe, 2009; Sylvén, 2019; Xanthou, 2010, 2011) and productive knowledge (Canga Alonso & Arribas, 2014; Merikivi & Pietilä, 2014). Besides, the present study also contributes to the field by incorporating to the analysis a calculation of the effect size.

Three main reasons can be stated to explain the better performance of CLIL learners. First, to become enrolled in the bilingual section in CSE, CLIL learners had to pass a language test; thus, they may have had a higher lexical level from the outset. Second, CLIL implies learning contents through a foreign language —English, in this case. This results in learners being exposed to a broader range of vocabulary, as the focus is not only on general but also on academic English and the specific vocabulary related to different disciplinary areas. This may result in more varied input, which may foster lexical acquisition. Finally, CLIL learners are exposed to a greater number of hours being instructed in English, since they receive EFL instruction in addition to academic subjects taught in English. This difference in the quantity of exposure to the L2 could also explain the better results obtained by CLIL learners. This latter reason will be explored in greater detail below.

With respect to the academic band, by the end of CSE, learners recognise nearly three-fifths of the academic word list. Nevertheless, as with the knowledge of high-frequency words, significant differences are found depending on the educational approach to which learners are exposed: while CLIL learners recognise an average of 426 items out of the 570 making up the AWL (Coxhead, 2000), mainstream EFL learners know, in the receptive dimension, 249 words. However, the significance of this difference differs over time, and, as with the growth of the knowledge of high-frequency words, the difference between both groups drops in extrapolated values from T1 to T2.

These results corroborate previous findings concerning CLIL and EFL differences in the receptive knowledge of academic terms (Merikivi & Pietilä, 2014). However, in the present study, both CLIL and EFL learners' understanding of the academic band is higher. Moreover, my study differs itself from previous research by incorporating two new methodological approaches to enrich the analysis: a longitudinal analysis, and a calculation of the effect size of the difference. In this respect, the results confirm that 1) the difference in academic lexical knowledge is large according to Plonsky and Oswald's interpretation (2014), and 2) that the difference remains significant during the academic year the data was collected.

Regarding the productive knowledge of the academic band, on average learners can recall 194 out of the 836 academic word families included in the UWL, albeit with large differences among groups:

when EFL learners recall about 95 words, CLIL learners recall an average of 261 terms, almost three times as many.

In practice, this means that both groups can recall only a limited number of academic terms, and, consequently, they will probably encounter great difficulties when trying to develop pieces of academic texts. My analysis of the productive knowledge of the 2K band anticipated potential problems when facing writing tasks, and, in this respect, learners' performance in the academic PVLT supports this hypothesis.

The observed differences between CLIL and EFL learners' lexical knowledge of the academic band complement previous studies into CLIL and EFL learners' lexical knowledge: this finding proves that CLIL learners not only show better control of high-frequency terms, as had already been confirmed, but also of academic lexis. As noted above, two reasons are usually given to justify CLIL learners' better performance in the 2K band: a larger and more varied amount of input, and a better starting level. However, to the best of my knowledge, no studies have yet attested to the relationship between a better starting point and a better command of academic terms; therefore, in my view this latter explanation cannot be the only reason for such results.

One possible explanation for these results is the nature of the learning contexts (CLIL vs EFL): although both groups study English in a formal setting, CLIL learners attend a content-based programme while EFL learners follow an instruction-based programme. This results in a different approach to the language: Whereas in CLIL the language is a vehicle to communicate and learn new contents, in regular EFL settings language is both the aim and the content of the subject. Thus, in this regard, the main difference between contexts is the frequency of occurrence of academic terms: CLIL learners will have greater contact with this vocabulary, resulting in its better command.

This does not mean that the EFL group does not develop their academic language; it only means that they develop it in a lower proportion. The regional curriculum (Junta de Extremadura, 2015) includes some aims within the organisation of the EFL subject closely related to academic language, such as the formulation of hypotheses in an L2, the mastery of lexis related to scientific studies, and the use of some specific discourse markers. These objectives mean that EFL learners are expected to acquire some academic knowledge, if only in an implicit way.

In light of these findings, four main conclusions regarding the lexical knowledge of CLIL and EFL learners can be drawn. First, at the end of CSE, these groups of learners are still in the process of acquiring the 2K and academic bands, although there are significant differences between the groups. As mentioned above, the mastery of the 2K band is the cornerstone upon which learners' lexical development is built. For this reason, I recommend helping learners achieve this goal as soon as possible, employing meaningful and lexical-focused activities and extensive supported-reading practice.

Second, learners can recognise a number of items larger than the ones they can recall, in both the 2K and academic bands. At least two reasons can explain these differences. On the one hand, the literature on the receptive-productive dichotomy reports that receptive knowledge is typically more extensive (Fan, 2000; Laufer, 1998; Laufer & Paribakht, 1998; Melka, 1997; Takala, 1984). Also, other reasons, such as the test format, may contribute to this difference. While the PVLT and the VLT share many of the items measured, each test assesses lexical knowledge through different tasks: the VLT is a meaning recognition test in which a set of choices is given, and test-takers are asked to select the correct option. The PVLT, by contrast, is a completion test; that is, only a prompt of the expected word is given, and participants need to identify the word. In practice, the completion task is much more cognitively demanding, as test-takers need to understand the context first and then recall a word that fits in. Such difficulty could result in worse performances.

Third, CLIL and EFL learners know fewer academic terms in comparison to 2K terms. In my view, such a difference may be caused by the nature of the academic band. Unlike the rest of the bands measured in the VLT and PVLT, this band is not based on frequency. It includes a wide variety of words belonging to different frequency bands; therefore, it cannot be situated at a specific level. Depending on the lists and on the studies, words contained in that list are usually located within the 3K and 5K band window (Schmitt, 2010). As these words will probably be less frequent in the input presented to learners, it is only normal that they produce lower test results compared to words in the 2K band.

Finally, the findings of this study confirm that CLIL participants systematically outperform regular EFL participants in both the VLT and PVLT. Unlike previous studies, the effect size of the differences has been calculated with a longitudinal analysis of the growth in the case of receptive vocabulary.

Focussing on this latter aspect, at first sight the results seemed to point to conflicting findings, as EFL learners showed a greater receptive knowledge growth rate. In my view, the most plausible explanation is the existence of a sort of 'ceiling effect'. VLT and PVLT split vocabulary knowledge into levels or bands. As a specific band has been or is close to being mastered, it becomes more challenging to find significant vocabulary growths. Regular EFL learners' recognition of high-frequency items at T1 was well below that of CLIL learners, so one could argue that it would be easier for EFL learners to improve their vocabulary level in larger proportions. However, a closer look at the data reveals that this explanation is not corroborated in the evolution of the rate of full recognition of the band. As is the case with the receptive knowledge of the 2K and academic bands, the development of complete mastery of the bands is more remarkable for the CLIL learners. This indicates that, although the mean score growth is larger in the regular EFL group, there is also an under-estimated positive evolution of CLIL learners. Their full receptive mastery of the bands improves in larger proportions.

For this reason, I would argue that the design of the study may not allow us to fully appreciate lexical growth. In this respect, it may be relevant to consider the option of administering other bands in addition to the 2K and academic bands to explore vocabulary growth, despite the apparent drawback of the time required to administer all the bands. I would opt to use the latest version of the VLT (Webb et al., 2017). This new proposal includes five vocabulary levels ranging from the 1K to the 5th 1K band. Thus, it contains the first 1K band, which accounts for approximately 65-85% of spoken and written English (Webb & Nation, 2017), and whose recognition is central to becoming able to understand daily conversations (Meara, 2010). Moreover, the inclusion of the 4th and 5th 1K bands facilitates a profile of the knowledge of the 5K most frequent words, which are considered the most critical lexical items for SLA learners. Exploring the recognition rate of the five thousand most frequent English words would help us determine how vocabulary level may vary along time.

All in all, CLIL learners seems to present a higher recognition and productive knowledge of high-frequency and academic terms, corroborated by the analysis of learners' mean scores, effect sizes of the differences and rate of full recognition of the bands. As I have already mentioned, CLIL learners' better performance has been related to differences in the quantity (IAoE) and quality of input, and their better starting comprehension level. However, CLIL learners are usually regarded as homogeneous L2 learners, with little attention traditionally paid to their language learning features or background. The CLIL learners in this PhD study are not homogeneous, as, among other differences, they had joined the CLIL programme at different ages, and, consequently, had different L2 exposure. This difference in IAoE between some of the CLIL groups is, in fact, greater than the one existing between some CLIL groups (such as the late CLIL learners) and the EFL one. Given that RQ3 directly addresses the role of IAoE on lexical competence, and that there is a section in Chapter Five examining the relationship between IAoE and lexical competence, this discussion will be taken up in the section 4.

3. Language teaching approach and selection of VLSs

As stated above, the final aim of this thesis is to understand how CLIL, as a language teaching approach, affects lexical competence development. In this dissertation, lexical competence is regarded as more than the mere knowledge of L2 words, as it also involves the way lexis is processed in the mind. The previous section has noted the benefits of CLIL in one of the components of lexical competence: the recognition and production of high-frequency and academic terms. This section discusses whether CLIL produces any change in the way new L2 words are processed.

The integration of language and disciplinary content within CLIL is quite challenging in practice, and its success requires a third element: cognitive engagement (Coyle et al., 2010). In CLIL, learners are asked to deal with new content and new language simultaneously. This demand is expected to result in the use of new learning methods and thinking processes that allow learners to cope with both types

of knowledge at the same time. In the case of content learning, there is abundant literature about how CLIL modifies content development. For example, Mehisto et al. (2008) illustrate the difference of CLIL from other disciplinary teaching approaches with Bloom's taxonomy of educational objectives: Unlike other disciplinary subjects, in CLIL it is common in an L2 to find activities in which learners are asked to establish connections among facts, to give reasons to support their decisions or thoughts, and to produce original work. These activities promote the Higher Order Thinking Skills (Bloom, 1984), which involve a new way of learning and a larger cognitive engagement.

However, to the best of my knowledge, there is little research on how CLIL affects language processing. In this respect, LLSs and VLSs, as reflections of the learning processing taking place in the mind, could provide some insights. The selection of LLSs has been proven to be affected by the language teaching approach and the duration of the L2 learning process (Griffiths, 2003; Oxford & Nyikos, 1989). In this respect, the little research available relating CLIL and LLSs corroborates that CLIL learners make different use of LLSs than regular EFL learners (Milla & Gutierrez-Mangado, 2019; Psaltou-Joycey et al., 2012). However, despite the importance of lexis in CLIL, to the best of my knowledge practically no attention has been paid to the use of VLSs by CLIL learners. This research was designed to explore the differences in the selection of VLSs of CLIL and regular EFL learners. The results were presented in section 3 of Chapter Five.

To establish the framework of the study, the last part of the methodology section dealt with a reconsideration of Schmitt's taxonomy, based on evidence that the groups proposed did not show internal coherence. Ehrman, Leaver and Oxford (2003) and Cohen (2007, 2014) highlighted that the usefulness of a strategy is determined by how it relates to other strategies. Therefore, considering the objectives of this dissertation, the exploration of how strategies were combined was central to the interests of this study. A factor analysis served this purpose. Results revealed the existence of eight categories: 'lexical analysis', 'mental imagery', 'repetition', 'linking', 'kinaesthetic', 'guessing', 'social strategies involving interaction with teachers' and 'social strategies involving interaction with classmates'. This new classification is based on the underlying connections between the different VLSs and allows an analysis of learners' use of other categories.

Once the conceptual framework was clarified, the relationship between participants' selection of VLSs and vocabulary knowledge was analysed. In this analysis, some VLSs have been positively related to better vocabulary mastery. First, the combined use of 'lexical analysis' strategies has been found to have a positive impact on the mastery of receptive knowledge of the 2K and academic bands. Second, the use of the strategy 'analysis of affixes and roots' has been positively related to lexical development of the receptive knowledge of the 2K and academic bands. Third, the use of the strategy 'connecting to its synonyms and antonyms' has also been found to be positively related to receptive knowledge of the 2K band. Finally, a better command of the receptive academic vocabulary is linked

to higher use of the strategies 'asking the teacher for an L1 translation' and 'asking students for meaning'.

Several reasons may explain the positive impact of these specific strategies on lexical development. First, some of these VLSs imply a close relationship with the L2 language and heightened metalinguistic awareness. Metalinguistic awareness is defined as "the ability to attend to, and reflect upon, the properties of language" (Davidson & Raschke, 2009, p. 1). The development of this capacity can result in an opportunity for enhancing vocabulary knowledge, as it may help learners identify patterns of use and infer new meanings (to which 'lexical analysis' strategies, in general, and the strategy 'analysis of affixes and roots', in particular, may contribute), and reflect on the properties of the terms to identify their similarities and discrepancies in aspects such as meaning, register, associations with other terms or collocations (which may foster the use of the strategy 'connecting to its synonyms and antonyms').

A second possible explanation, closely related to the previous one, is that the use of these specific VLSs involves a certain degree of cognitive processing. To create connections with synonyms and antonyms, or to analyse word parts, L2 learners need to establish some sort of connection among lexical terms, for which they need to be cognitively engaged with the task. In this respect, some research (Bloom, 1984; Craik & Lockhart, 1972; Craik & Tulving, 1975) suggests that the higher the cognitive implication, the greater the knowledge.

Finally, in the case of the social strategies in which learners look for an L1 translation ('asking the teacher for an L1 translation' and 'asking students for meaning'), their positive influence in the receptive knowledge of academic terms may be related to the essence of academic vocabulary. A high percentage of the items belonging to academic word lists (either the UWL or the AWL) have a Latin or Greek origin and may resemble their Spanish equivalent. In this sense, the creation of links between L1 and L2 terms may be positive, as Spanish learners may resort to their mother tongue to retain the meaning and form of these English words.

Thus, in light of these results, vocabulary knowledge seems to be fostered by the use of actions that involve a cognitive implication, e.g., splitting words into parts and inferring meanings from those different parts; and some kind of word knowledge, i.e., in order to infer meanings from word parts one must know what the suffix or affix means and have a basic knowledge of the root word. It makes sense that, as vocabulary knowledge is developed, the L2 learner will be able to create more connections between words resulting from an easier access to them and to previous learning experiences. However, a low number of significant correlations between VLSs and vocabulary knowledge has been found, especially when it comes to productive vocabulary. A plausible reason for this may be the low level of performance in the vocabulary tests and the homogeneity of results. In those bands in which the results are more varied, more connections arise; whereas in bands where vocabulary knowledge is more limited and homogeneous (i.e., 2K and academic productive vocabulary), no significant correlations are found.

Thus, further research is necessary to incorporate into the sample more participants with different vocabulary levels.

Moving on to the comparison between CLIL and mainstream EFL learners' VLS use, and starting with the overall VLSs use, CLIL learners have been found to use them more. However, the difference with EFL is not significant, but it could be relevant for this study, as it may reflect a pattern of use already shown by Psaltou-Joycey et al. (2012). Different reasons may explain this finding. First, the literature (Oxford & Nyikos, 1989; Park, 1997; Psaltou-Joycey & Kantaridou, 2009; Vrettou, 2009; Wharton, 2000) often reveals a greater and more efficient use of strategies by the most proficient learners. In this study, CLIL participants have demonstrated a larger lexical knowledge and it is only natural that they may be expected to make greater use of VLSs. Second, this greater use of VLSs by CLIL learners could be related to the methodology employed and the teacher's role. Nation (2001) suggests that teachers play a relevant role VLSs learning, as they can provide learners with new ways of learning. More recently, Oxford (2017) highlights that strategies are teachable; therefore, the teacher's role as an instructor is vital. In the case of the two groups here explored, EFL students' way of dealing with L2 vocabulary would have only been influenced by the EFL teacher, while CLIL learners are exposed to a number of teachers who use English to teach content in different disciplinary subjects. All those teachers can influence CLIL students' language learning process, fostering a more comprehensive range of techniques or VLSs, or demonstrating how to apply some specific strategies to content subjects, and thereby influencing CLIL learners' selection of strategies.

Concerning the combined use of different strategies, differences have emerged regarding the use of 'lexical analysis' strategies, which are significantly more employed by CLIL learners. A total of six VLSs are regarded as 'lexical analysis' strategies. They are closely related to lexis and lexical development, as they imply either (a) mastery of the L2 sufficient to infer meaning from authentic input ('using English-Language media') and to connect these newly learned words to other L2 lexical items ('connecting the word to its synonyms and antonyms'), or (b) some kind of work to understand word meaning ('analysis of affixes and roots', 'analysis of the part of speech', 'grouping words together to study them', and 'using a new word in a sentence').

The greater use of this kind of VLSs may be directly related to CLIL. 'Lexical analysis' strategies demand the type of attention and metalinguistic awareness that are expected to be promoted in this language teaching approach. CLIL fosters the use of HOTS (Higher Order Thinking Skills; Bloom, 1984), in which learners are asked to be cognitively engaged with the task (Coyle et al. 2010). For this reason, the combined use of VLSs involving greater cognitive attention (i.e., 'analysis of affixes and roots', 'analysis of the part of speech' and 'connecting the word to its synonyms and antonyms' strategies), may be better promoted in CLIL settings. Moreover, in CLIL, charts and other visual supports are common resources. These visual aids, based on the connections of ideas are common

resources, so CLIL learners can create links between words and look for connections such as synonyms and antonyms. However, there is a second possibility, and it is directly related to vocabulary knowledge. Some of the foregoing strategies demand a basic knowledge of lexis; that is, some kind of prior knowledge is needed in order to create links with other English words ('connecting the word to its synonyms and antonyms'), to understand the different affixes and roots and parts of speech ('analysis of affixes and roots' and 'analysis of the part of speech') and to understand media in an FL ('use of English-language media'). For this reason, it may be that mainstream EFL learners—with a demonstrated lower vocabulary knowledge— are incapable of making use of these strategies.

Finally, differences have also been found at the level of the selection of specific VLSs. Apart from non-significant differences in the overall use of VLSs, and the greater use of 'lexical analysis' strategies by CLIL learners, 'analysis of affixes and roots', and 'analysis of part of speech' strategies are more often selected by CLIL learners than their EFL peers. In contrast, they resort less frequently to 'word lists'.

CLIL learners' greater use of the strategies 'analysis of affixes and roots' and 'analysis of part of speech' could be due to several reasons. One of the most plausible explanations is that, as CLIL learners are more exposed to English and to a more varied instruction input, they have developed a metalinguistic awareness that they may be using to understand and learn vocabulary items. A second reason is related to the observed CLIL and EFL differences in vocabulary knowledge. These strategies have been related to larger vocabulary mastery, and CLIL participants have demonstrated a better vocabulary command. Therefore, it could be that vocabulary level could determine the use of this strategy. Finally, and closely related to the two previous arguments, their greater use may also be determined by the degree of cognitive engagement required by the VLSs. Both strategies imply HOTS mechanisms, as they indicate an in-depth analysis of the different lexical terms. CLIL, as a dual focus approach, also fosters HOTS tasks, in which learners are asked, in the content subjects, to compare, analyse and be critical while producing new work. Thus, it may be that these learners extrapolate these skills to L2 learning.

This latter reason may also explain the lower use of 'word lists' by CLIL learners. In this respect, there are different cognitive theories, such as the Depth of Processing Hypothesis (Craik & Lockhart, 1972; Craik & Tulving, 1975), which state that the use of methods or strategies based on the simple repetition of words or structures, such as word lists, does not seem to be as beneficial to language learning as other models, given that they do not require a high cognitive engagement. This fact, together with the idea that CLIL learners are familiar with HOTS tasks, may lead us to think that this lower use may be related to the language teaching approach. However, in this study, it has been observed that this strategy is one of the preferred ones for both CLIL and mainstream EFL learners. This finding is in line with other studies, which consistently reported word lists to be one of the preferred strategies for

learners (Schmitt, 1997; Xhaferi & Xhaferi, 2008; Castellano-Risco, 2018a). Therefore, it cannot be said that CLIL learners are reluctant to use it, but merely that they resort to it significantly less often than their EFL counterparts. It seems that those learners who are exposed to a lower amount of L2 input—mainstream EFL learners—are more willing to use word lists than those exposed to a greater amount of input. Perhaps, when CLIL learners started to learn English, they often used this strategy, but they reduced its use as they were gradually exposed to an increasing number of new vocabulary items. In other words, it may be that, as the language input increases, more and more varied strategies come into play, diminishing the use of those strategies that were initially employed by the learners.

Moving on to the evolution of the use of VLSs, one of the most remarkable fact is that VLSs are relatively stable in time in all learners, although it has been found that CLIL and EFL groups' selections evolve in different ways with respect to some specific strategies: 'analysis of affixes and roots', 'written repetition' and 'verbal repetition'. This makes the use of VLSs more homogenous at T2: whereas at T2, both groups of learners only differ in the use of three individual VLSs, at T1, they also differ in their use of the strategies 'use of English-language media' and 'written repetition'.

Concerning the greater homogeneity in the use of the VLS 'use of English-language media' at T2, in my view, it may be related to the lexical improvement regular EFL learners show at T2. As has been already explained, the use of this VLS implies at least a basic knowledge of the L2. At T1, regular EFL learners did not even recognise half of the 2K most frequent English words, impeding them from understanding even the most basic conversations. At T2, EFL learners have improved their lexical knowledge, and this may result in greater confidence to listen to English-language media.

There is a second reason that may also explain this more considerable homogeneity at T2. The greater use of both 'use of English-language media' and 'written repetition' by EFL and CLIL learners respectively is not a generalised phenomenon among all learners; when examining their evolution in the different schools, these significant variations are only observed in certain schools. If it is not a generalised phenomenon, some school-dependent factors, such as the influence of the teachers or the teaching and learning materials used, may account for these variations.

The issues here discussed seem to indicate that CLIL has some impact on the selection of VLSs. Although the differences can be regarded as marginal and are not found in the overall use of VLSs, significant differences are found at other levels, which may indicate an emerging different pattern of use. In fact, CLIL learners make significantly greater use of the kind of VLSs that are positively related to more extensive receptive knowledge of the high-frequency and academic items. As the usefulness of VLSs seems to be determined by how they are grouped (Chamot, 2004; Cohen, 2007, 2014; Gu, 2003; Cohen, 2007, 2014; Ehrman, Leaver & Oxford, 2004; Hsiao & Oxford, 2002; Oxford, 2002, 2003), this finding is quite relevant, as it shows that the implementation of a CLIL approach fosters lexical development by promoting the combined use of VLSs that are related to better lexical command.

Finally, both groups differ in the selection of specific VLSs: CLIL learners make more use of strategies that have been positively related to lexical knowledge.

4.   Instructed Amount of Exposure and lexical knowledge

CLIL is a language teaching approach that seems to present a series of benefits with regard to both content learning (Mehisto & Asser, 2007; Ouazizi, 2016; Serra, 2007; Surmount et al., 2016; Xanthou, 2011) and language learning (Agustín-Llach, 2009; Hüttner & Rieder-Bünemann, 2010; Lorenzo et al., 2009; Lorenzo & Rodríguez, 2014; Martínez Agudo, 2020; Nightingale & Safont, 2019; Pascual Peña, 2015; Ruiz de Zarobe, 2008; Xanthou, 2010, 2011; Sylvén, 2019; Yi Lu & Jeong, 2018). As shown above, traditionally, the differences between CLIL and mainstream EFL learners have been ascribed to, among other factors, differences in IAoE. Understanding the impact of this variable on the accounted differences between CLIL and EFL learners' lexical competence (i.e., vocabulary knowledge and selection of VLSs) is central to address criticism suggesting that the presumed benefits of CLIL are exclusively the result of greater L2 exposure (Bruton, 2011a, 2011b, 2013), and to have a better understanding of how CLIL works. However, the real effect of this factor has not been explored in isolation, as it interacts with other CLIL variables, such as the quality of the input, variations in the teaching methods, and the role of the teachers.

In this thesis, the role played by the IAoE among the different CLIL variables has been isolated by comparing learners of the same age but with different amounts of exposure within the same learning context (CLIL), and by contrasting each of the three identified kinds of CLIL experiences (varying in amount of exposure: early, standard and late CLIL learners) with regular EFL.

Regarding lexical knowledge, the results of the present study indicate that the three CLIL subgroups do not present statistically significant differences in their receptive and productive knowledge of the 2K and academic bands, despite having been exposed to different amounts of instructed input (up to 1,000 hours of difference). In contrast, regardless of the variations in the number of hours of instruction, differences between the various CLIL sub-groups and the mainstream EFL group always remain statistically significant. Interestingly, a difference of 1,000 hours of instruction within the same programme (late CLIL vs early CLIL) does not result in significant variations regarding general and academic vocabulary knowledge. In contrast, when different language teaching approaches (late CLIL vs EFL), are compared, differences in hours of instruction lead to significant lexical knowledge differences. These findings are in line with previous studies on the impact of age of onset in language learning (Agustín-Llach & Jiménez-Catalán, 2018; Muñoz, 2014) and vocabulary acquisition (Miralpeix, 2008; Muñoz, 2006).

One possible explanation for these results may be methodological. The statistical analysis of the difference between early and late CLIL groups' knowledge of the general, non-academic and academic bands yielded no significant differences between them but did indicate moderate effect sizes when explored. These magnitudes may be suggesting that the lack of significance may be caused by the limited samples contained in the different groups.

There may be a second explanation for these results: the nature of a CLIL approach and the role that language plays in it. CLIL encourages language learning through a focus on meaning, and learners are exposed to a wide variety of input related to the different academic subjects they have to learn in the foreign language. This re-conceptualisation of the language role in the classroom may result in a better lexical command.

Thirdly, the lack of significance in the differences among the CLIL groups may be related to maturational constraints. CLIL learners joined the programme at different ages, resulting in the differences in IAoE. Nevertheless, young learners usually have a slower rate of learning when compared to older learners (Agustín-Llach & Jiménez Catalán, 2018; Cadierno et al. 2020; García Mayo & García Lecumberri's volume, 2003; Miralpeix, 2007; Muñoz, 2006). Thus, the processing capacity of the learners who started CLIL programmes in the lower grades of primary education may not have been fully developed during the first years, resulting in a lower learning ratio of vocabulary items.

Moving to the exploration of the selection of VLSs, in general, few differences were found among the different subgroups. On the one hand, no significant differences in the overall selection of VLSs or the selection of different kinds of strategies were identified when the three CLIL groups are compared. This lack of difference among the CLIL subgroups is quite illustrative. It points to, once again, a considerable similarity in the use of VLSs within the CLIL subgroups, no matter their IAoE. However, when compared to the EFL group' selection of VLSs, again, no significant differences either in the overall selection of VLSs or in the different kinds of VLSs were found.

On the other hand, some differences in the usage of individual VLSs must be acknowledged: both the late CLIL and EFL groups present some significant differences in their selection of VLSs when compared to the other groups. In comparison to the early CLIL group, late CLIL learners make significantly lower use of the strategies like 'connecting the word to its synonyms and antonyms' and 'use of a bilingual dictionary'. Similarly, when comparing late CLIL learners' choices to those of the standard CLIL group, differences can be identified only in the usage of the strategy 'analysis of the part of speech', which is utilised significantly more often by standard CLIL participants. At first, I thought it could be related to school differences. However, no such differences were identified.

As for the comparison between the CLIL groups and the EFL one, EFL learners were found to make lower use of some lexical analysis strategies than their CLIL peers: early CLIL learners make greater use of the strategies 'analysis of affixes and roots', 'use of a bilingual dictionary', and 'asking

the teacher for an L1 translation'; standard CLIL learners select significantly more often the strategy 'analysis of the part of speech', and late CLIL learners resort more frequently to the strategy 'use of English-language media'. However, as with CLIL learners' differences, these differences are not generalised or common to the different groups, but seem isolated cases in specific schools.

For this reason, and, in the absence of other results or findings that could help to clarify this issue, I can only suggest two possible reasons that may justify these findings. First, the differences among groups may be due to the influence of other variables not controlled or studied in this PhD dissertation, such as participants' learning styles (Ehram & Oxford, 1990; Griffiths, 2004). Second, given the design of this study, the sample of which each CLIL group is comprised is relatively small, and this may lead to low statistical power. This may produce a lack of significance in the differences among groups.

All in all, the findings of this study regarding the impact of the IAoE on lexical development make it difficult to give a definitive answer on this issue. Regarding lexical knowledge, it cannot be stated that differences between CLIL and EFL learners are related exclusively to differences in IAoE, as other factors seem to be playing a role. One possible explanation may be the nature of a CLIL approach and the role that language plays in it. CLIL encourages language learning through a focus on meaning. Unlike traditional language teaching approaches in which language is treated as the main aim and content (i.e., the focus is on form), in CLIL, language is conceived as a vehicle for the transmission of content knowledge. Thus, the focus in CLIL is on language use, rather than on language metalinguistic tasks, and it affects the way language is learnt. There is also a difference in the kind of language needed. Whereas in other language teaching approaches the objective is to use the language in an everyday context, in CLIL, language is used in an academic context; therefore, this richer L2 exposure will naturally foster academic vocabulary learning. Another possible explanation is that CLIL enhances peer interaction and meaningful learning opportunities, in which learners have to build new content based on their prior knowledge, experiences and skills, which promotes language learning (Ellison, 2017; Mehisto et al., 2008). Finally, the extra motivation provided by CLIL and the positive attitudes developed towards the learning process (Fernández Fontecha, 2014; Lasagabaster, 2011; Sylvén & Thompson, 2015) constitute an additional element which has been connected to language proficiency (Bernaus & Gardner, 2008; Schmidt & Watanabe, 2001).

Regarding how lexical processing is affected by IAoE, my results do not allow me to establish a clear conclusion regarding how lexical processing is affected by IAoE. The lack of significant differences among groups could indicate either (1) a similar use of VLSs regardless of the type of instruction and IAoE, or (2) the test has low statistical power. For this reason, further research is needed.

CHAPTER SEVEN:

CONCLUSIONS

1. Conclusiones

Esta tesis doctoral ha tenido como objetivo explorar las diferencias en la competencia léxica de alumnos AICLE e ILE (Inglés como Lengua Extranjera) desde una perspectiva holística. Tradicionalmente, los estudios sobre AICLE y su impacto en el desarrollo léxico tenían, en mi opinión, una idea restringida de competencia léxica y se centraban exclusivamente en el producto del aprendizaje, es decir, el conocimiento del vocabulario. Esto ha dado lugar a algunos problemas al intentar identificar y aislar las causas de las diferencias entre ambos grupos de alumnos. Este estudio ha tratado de eludir el problema metodológico resultante de equiparar las experiencias de aprendizaje AICLE y no-AICLE en términos de desarrollo léxico (1) ampliando la concepción tradicional de competencia léxica e incluyendo, dentro de ella, el procesamiento léxico, y (2) analizando la cantidad de información a la que han estado expuestos los alumnos, en un intento de determinar si las diferencias entre alumnos AICLE e ILE están relacionadas exclusivamente con este factor o con los diferentes contextos de aprendizaje de idiomas.

A la luz de los resultados de este estudio, la implementación de un enfoque AICLE incide en el desarrollo de la competencia léxica, no solo en lo que respecta al conocimiento léxico, como ya se ha demostrado, sino también en la forma en que se procesa el vocabulario en la mente. En cuanto al conocimiento léxico, se ha observado una notable diferencia entre el conocimiento de vocabulario de los alumnos AICLE y el de los estudiantes de inglés como lengua extranjera a favor del primero tanto en la dimensión receptiva (Agustín-Llach y Canga Alonso, 2016; Canga Alonso, 2015b; Canga Alonso y Arribas García, 2014; Merikivi y Pietilä, 2014; Sylven, 2010; Xanthou, 2011), como en la productiva.

En lo referente a selección de estrategias, también se han encontrado diferencias. Los estudiantes AICLE hacen más uso de las estrategias de aprendizaje que los estudiantes ILE, probablemente debido al papel más variado que desempeñan los diferentes profesores involucrados en el programa AICLE, y seleccionan con más frecuencia algunas estrategias que se encuentran directamente relacionados con la ganancia de vocabulario.

Tradicionalmente, las diferencias entre ambos enfoques de enseñanza se han atribuido a la confluencia de varios factores, entre los que destaca la cantidad de exposición a la lengua extranjera. La última parte de este estudio ha intentado esclarecer el papel de la exposición sobre las diferencias encontradas previamente y, con este objetivo, la muestra se agrupó en cuatro grandes grupos atendiendo a la cantidad de exposición que habían recibido a lo largo de su vida educativa. El análisis del conocimiento léxico de los cuatro grupos ha demostrado que las diferencias de exposición de la lengua extranjera dentro del mismo enfoque de enseñanza de la lengua no afectan directamente el aprendizaje de vocabulario. Por el contrario, cuando las diferencias ocurren en diferentes enfoques de enseñanza de idiomas, dan como resultado variaciones significativas en el conocimiento léxico. Por su parte, el análisis de la selección de estrategias no ha arrojado hallazgos muy concluyentes, ya que no se han encontrado diferencias significativas en el uso de estrategias entre los cuatro grupos. Estos hechos plantean dos argumentos. En primer lugar, una entrada posterior en los programas AICLE, o el retraso en la implementación del programa, no parecen tener un impacto negativo sustancial en el desarrollo léxico L2 de los estudiantes, ya que los estudiantes pueden alcanzar el mismo nivel de gestión de términos de la lengua extranjera y parecen hacer uso de las mismas estrategias, independientemente de aspectos que se consideren determinantes como el género o la influencia del profesorado. Esto es consistente con investigaciones previas (Agustín-Llach y Jiménez Catalán, 2018; Miralpeix, 2007; Muñoz, 2014). En segundo lugar, parece que las diferencias en el conocimiento del vocabulario ya no se pueden atribuir únicamente a las diferencias en la cantidad de exposición a la lengua extranjera, sino a otros factores, como el uso de diferentes enfoques de enseñanza de idiomas, tales como el tipo y la calidad de exposición a la lengua extranjera o la metodología de enseñanza. Con todo, el enfoque AICLE parece afectar el aprendizaje de vocabulario en el sentido más general del término. Existe evidencia que respalda que AICLE parece modificar el conocimiento léxico y las estrategias que estos estudiantes seleccionan y utilizan para adquirir nuevos elementos léxicos de L2, aunque en este último elemento, el impacto es menos notable. Por lo tanto, estas diferencias deben tenerse en cuenta al tratar a los estudiantes AICLE y los estudiantes de inglés como lengua extranjera para adaptar la práctica docente a sus necesidades. Por un lado, los estudiantes AICLE parecen tener el conocimiento léxico L2 necesario para comprender conversaciones básicas en inglés. La práctica docente debe estar orientada a ayudarles a mantener y ampliar su conocimiento léxico. Por otro lado, los estudiantes expuestos a un enfoque más tradicional de ILE todavía están en el proceso de adquirir las palabras más frecuentes y los esfuerzos de los docentes deben ir orientados hacia el desarrollo y la adquisición de las primeras

bandas de vocabulario que les permitan sentirse mejor integrados en la lengua inglesa. Estos hallazgos tienen claras implicaciones para la práctica de la enseñanza de idiomas. En cuanto a los resultados relacionados con el conocimiento del vocabulario, en mi opinión, una de las implicaciones más claras es la selección adecuada de materiales. Hasta hace poco, la mayoría de los estudiantes AICLE usaban los mismos materiales para aprender inglés que otros alumnos expuestos a otros enfoques. De hecho, todavía está sucediendo en algunas escuelas. La confirmación de las diferencias entre el alumnado expuesto a distintos enfoques puede ayudar a los profesores a seleccionar los materiales didácticos más apropiados y adaptarlos al nivel de los alumnos. En el caso de la selección de estrategias, este estudio también ha identificado algunos tipos de estrategias que parecen correlacionarse positivamente con el aprendizaje de vocabulario. Dado que las estrategias pueden enseñarse, sería aconsejable que los materiales didácticos promuevan el uso de las estrategias más beneficiosas de forma explícita y que los profesores las conozcan.

2. Conclusions

This doctoral dissertation has attempted to explore differences in lexical competence between CLIL and mainstream EFL approaches from a holistic perspective. Traditionally, studies on CLIL and mainstream EFL learners' lexical differences had, in my view, a restricted idea of lexical competence and focussed exclusively on the product of learning, i.e., vocabulary knowledge. This has led to some problems when attempting to identify and isolate the causes of the differences between both groups of learners. This study has tried to circumvent the methodological problem arising from equating CLIL and non-CLIL learning experiences in terms of lexical development by (1) broadening the traditional conception of lexical competence and including, within it, lexical processing, and (2) by analysing the amount of input to which learners have been exposed, in an attempt to determine whether the differences between CLIL and regular EFL learners are related exclusively to this factor or the different language learning contexts.

In light of the results of this study, the implementation of a CLIL approach affects the development of the lexical competence, not only as regards lexical knowledge, as had already been proved, but also concerning the way vocabulary is processed in the mind. As for lexical knowledge, a notable difference between CLIL and mainstream EFL learners' vocabulary knowledge in favour of the former has been observed, in both, the receptive (Agustín-Llach & Canga Alonso, 2016; Canga Alonso, 2015b; Canga Alonso & Arribas García, 2014; Merikivi & Pietilä, 2014; Sylven, 2010; Xanthou, 2011), and productive dimensions. Concerning the selection of VLSs, differences have also been encountered. CLIL learners make more use of VLSs than mainstream EFL learners, probably due to the more varied role that the different teachers involved in the CLIL programme play and select more frequently some particular VLSs found directly related to vocabulary gain.

Traditionally, CLIL and regular EFL learners' differences have been related to the confluence of various factors, among which the IAoE stands out. The last part of this study has attempted to clarify the role of exposure on the accounted differences, and, with this aim, the sample was clustered into four main groups attending to the amount of exposure they had received along with their educational life.

The analysis of lexical knowledge of the four groups has shown that IAoE differences within the same language teaching approach do not directly impact vocabulary learning. In contrast, when IAoE differences occur in different language teaching approaches, they result in significant lexical knowledge variations. For its part, the analysis of the selection of VLSs has not yielded very conclusive findings, as no significant differences have been found in the use of VLSs among the four groups.

These findings raise two arguments. In the first place, a later entry (12 vs 6) into the CLIL programmes, or the delay in their implementation of CLIL, do not seem to have a substantial negative impact on learners' L2 lexical development, as learners can reach the same level of L2 terms management and they tend to make use of the same VLSs, regardless of aspects considered to be determining such as gender or teachers' influence. This finding is consistent with previous research on this issue in EFL (Agustín-Llach & Jiménez Catalán, 2018; Miralpeix, 2007; Muñoz, 2014). Secondly, it seems that differences in vocabulary knowledge can no longer be attributed solely to AoE differences, but to other factors such as differences related to using different language teaching approaches, such as the kind and quality of the L2 input, or the teaching methodology.

All in all, the CLIL approach seems to affect vocabulary learning in the most general sense of the term. There is evidence supporting that CLIL seems to modify the already-explored lexical knowledge of L2 learners and the strategies these learners select and use to acquire new L2 lexical items, although in this latter element, the impact is less noticeable. Therefore, these differences should be considered when treating CLIL and mainstream EFL learners to adapt the teaching practice to their needs. On the one hand, CLIL learners seem to have the necessary L2 lexical knowledge to understand basic English conversations. The teaching practice should be oriented to help them keep and expand lexical knowledge. On the other hand, mainstream EFL learners are still in the process of acquiring the most frequent words and teachers' efforts should be oriented towards the development and acquisition of the first bands of vocabulary that will enable them to feel better integrated into the English-speaking world.

These findings have clear implications for language teaching practice. As for the results concerning vocabulary knowledge, in my view, one of the clearest implications is the proper selection of materials. Until recently, most CLIL and mainstream EFL learners used the same materials in the EFL classes. In fact, it is still happening in some schools. The confirmation of CLIL and mainstream EFL learners' differences may help teachers select the most appropriate teaching materials and adapt them to the learners' level. In the case of the selection of VLSs, this study has also identified some kinds of VLSs that seem to correlate positively with vocabulary learning. Given that VLSs are teachable, it would be

advisable that teaching materials promote the use of these most beneficial VLSs explicitly and that teachers are aware of them.

3. Limitations

Findings need to be treated with caution, as there are several limitations to the study. First, as has been shown, the sample of this study (N= 138) is not very large. In fact, in some of the comparisons, the sample size could be weakening the results. For that reason, although the sample size is reasonable considering the context of the study, it would be advisable to get a larger and more diverse sample, which would include participants with a more varied range of features, such as nationalities, mother tongues, or SES. These additional features would enrich the analysis, in particular, that of VLSs.

Secondly, other variables should be considered when exploring the selection of VLSs, such as participants' learning styles or the influence of teachers and textbooks. Focussing on the teachers' effect, they have a key role in the EFL classroom, and it would be fascinating to analyse teachers' beliefs and speech and examine how these factors influence learners' choices. For this reason, I would recommend including other research instruments, such as observation and diaries compilation, with a twofold objective: to explore the influence of the teaching practice and to ensure the reliability of the usage reported by the learners. As for the impact of textbooks, it would also be relevant to examine which VLSs are suggested implicitly —by incorporating VLSs in exercises and tasks— and explicitly —for example, in tip sections—, to clarify if these VLSs are reflected in learners' selection.

The third primary concern is related to the appropriateness of the vocabulary tests chosen for this study. Both instruments, VLT and PVLT, were selected because they were considered reliable tools for measuring vocabulary knowledge of L2 learners of English and had already been used in previous studies analysing CLIL and mainstream EFL learners' lexical knowledge differences. However, after the administration of the tests, and the subsequent analysis of the results, some concerns about its suitability for this study emerged.

As for the appropriateness of bands and versions administered, only the knowledge of the 2K and academic bands was measured in this study. However, in light of the findings, the implementation of other bands is needed mainly for three reasons: first, some participants show a complete recognition of this band, so, to appreciate progression and the real differences among groups, more demanding levels should be administered. Besides, a ceiling effect in the 2K and academic bands receptive vocabulary growth has been observed, mainly caused by the greater command CLIL learners presented at T1. In this case, and considering the aims of this study, it would be important to explore the progression in the different bands, without focussing on specific levels. Finally, there is new evidence that both versions are not "parallel forms" (Bayazidi, 2017, p. 30; Kremmel & Schmitt, 2018), but they lead to

significantly different means, so different versions in the analysis of vocabulary gains should be avoided.

Quite close to the concern about the appropriateness of the bands selected for this study is the matter of the reliability of the tests. Although some studies have explored this issue, none of them has focussed on how the tests work for a secondary-school sample. Adolescents have their own features, and some of them, such as the maturational level, could make the test more cognitively demanding. Therefore, despite the large number of studies using this instrument with secondary-school learners, it would be advisable to check its reliability for adolescents and young learners.

Lastly, some authors (Kremmel & Schmitt, 2018) suggest that the lists that served as a basis to the VLT versions (2001) are outdated. The VLT and PVLT were developed making use of West's General Service List (1953) in the case of the 2K level, the Thorndike and Lorge's (1944) and Kučera and Francis's (1967) lists for the 3K, 5K and 10K levels, and the UWL (1984) and the AWL (Coxhead, 2000) in the case of the academic PVLT and VLT, respectively. Languages are 'living' entities that evolve with time. Words or expressions that were fashionable thirty years ago may no longer be, and new terms may have been incorporated into the language due to society's changes. Research should consider this evolution and apply the changes needed. In the case of the VLT and PVLT, the lists that served as a basis are obsoleted. New corpora and lists are available, such as the COCA/BNC (Nation, 2012a) or the new British National Corpus (BNC) lists published in 2018, and it would be advisable to update the tests considering these new materials.

4. Further Research

All these limitations can be considered, at the same time, starting points for further research. Starting with vocabulary knowledge, various proposals could be implemented. As a first approach, it is advisable to go further than the 2K and academic bands and explore the overall lexical competence of the learners, as some of the learners may present a greater command of other levels that would not be reflected in the command of the 2K and academic levels. Moreover, it seems relevant to examine the different materials to which learners are exposed to explore their impact on the final lexical competence learners demonstrate. Finally, to measure productive vocabulary knowledge, it would be better to make use of learners' productions on certain topics, together with the administration of the PVLT, as this methodology will ensure a better analysis of learners' lexical production.

Concerning the analysis of the VLSs further research proposals, it would be desirable to develop a taxonomy of language learners' VLSs in which the digital element, so present in today's classes and learners' world, would need to be considered. However, this is not the only change that should be contemplated, as this taxonomy should be inclusive and incorporate some improvements compared to

its predecessors. First, the development of this taxonomy should, in my opinion, follow a similar pattern to Schmitt's development, and focus on teachers' and students' thoughts at the same time that materials are examined. Moreover, regarding the sample, secondary-school learners should also be considered. In most western countries, young learners are FL learners *per se,* because they are required to attend EFL subjects at school. With the inclusion of both kinds of learners, it could be possible to achieve a more embracing taxonomy, and it could be better applied to subsequent research. Finally, the grouping of the different items should follow not only theoretical conceptualisations, but it should be necessarily supported with statistical analyses.

Regarding CLIL and mainstream EFL comparisons, it should be decided whether these comparisons are appropriate from a methodological perspective. Both groups differ in a wide range of aspects regarding their language learning background. For this reason, comparisons among learners may no longer be valid, and other kinds of comparisons should be carried out. For this reason and considering this and other CLIL and mainstream comparison studies, the comparison between learning contexts could be approached from two different perspectives. On the one hand, it would be advisable to design longitudinal studies in which both groups of learners' evolution was compared. On the other hand, it could also be possible to emphasise the analysis of different language learning elements, such as the interactions existing in the classrooms. In this sense, the development of a corpus based on L2 CLIL and mainstream EFL learners' production and interaction in the class could help to explore lexical development on a broader sense, and to explore from another perspective, the differences between CLIL and mainstream EFL learners.

Finally, I would also argue for exploring the impact of implementing CLIL programmes at different ages on learners' language development, an ignored issue up to date. In general, and in the Spanish context, the age of onset on CLIL programmes has been lowered, but the impact of these actions has not been examined scientifically. For this reason, it becomes relevant to explore the effects of these measures. To do so, samples who started CLIL programmes at different ages should be compared in many aspects, not only the L2 lexical development of these learners but including other components to get the most reliable data possible that could help the governing authorities to identify the most appropriate procedure for the full development of primary and secondary school learners.

Badajoz, November 2020

REFERENCES

Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics, 24*, 425–438.

Agustín-Llach, M. P. (2009). The Role of Spanish L1 in the Vocabulary Use of CLIL and non-CLIL EFL Learners. In M. P. Agustín-Llach & R. M. Jiménez Catalán (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 112–130). UK: Multilingual Matters.

Agustín-Llach, M. P. (2012*). Exploring the receptive vocabulary of young CLIL and non-CLIL learners*. Paper presented at Eurosla22.

Agustín-Llach, M. P., & Canga Alonso, A. (2016). Vocabulary growth in young CLIL and traditional EFL learners: evidence from research and implications for education. *International Journal of Applied Linguistics, 26*, 211–227.

Agustín Llach, M.P. & Canga Alonso, A. (2020). *Vocabulary Strategy Training to Enhance Second Language Acquisition in English as a Foreign Language.* Newcastle upon Tyne: Cambridge Scholars Publishing.

Agustín-Llach, M. P., & Jiménez Catalán, R. M. (2018). Teasing out the role of age and exposure in EFL learners' lexical profiles: A comparison of children and adults. *IRAL - International Review of Applied Linguistics in Language Teaching, 56*(1), 36–43. https://doi.org/doi.org/10.1515/iral-2015-0075

Aitchison, J. (2012). *Words in the Mind: An introduction to the Mental Lexicon*. Wiley-Backwell.

Alejo González, R., & Piquer Píriz, A. (2010). CLIL teacher training in Extremadura. A needs analysis perspective. In D. Lasagabaster & Y. Ruiz de Zarobe (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training* (pp. 219–242). Cambridge: Cambridge Scholars Publishing.

Alejo, R., & Piquer-Píriz, A. (2016a). Measuring the productive vocabulary of secondary school CLIL students: is Lex30 a valid test for low level school learners? *Vigo International Journal of Applied Linguistics, 16*, 31–53.

Alejo, R., & Piquer-Píriz, A. (2016b). Urban vs. rural CLIL: an analysis of input-related variables, motivation and language attainment. *Language, Culture and Curriculum, 29*, 245–262. https://doi.org/10.1080/07908318.2016.1154068

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: research reviews* (pp. 77–117). Newark, D.E.: International Reading Association.

Anderson, R. C., & Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. *Advances in Reading/Language Research, 2*, 231–256.

Arribas, M. (2016). Analysing a whole CLIL school: Students' attitudes, motivation, and receptive vocabulary outcomes. *Latin American Journal of Content and Language Integrated Learning, 9(2),* 267–292.

Azizi, M., Sayedi, R., & Asoudeh, F. (2012). Theoretical study of lexical network structure in second language learning. *Procedia - Social and Behavioral Sciences*, *32*, 128–133.

Baba, K. (2002). Test review: Lex30. *Language Testing Update, 32*, 68–71.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bacon, S. M. (1992). Phases of listening to authentic input in Spanish: a descriptive study. *Foreign Language Annals, 25*, 317–334.

Banisaeid, M., & Huang, J. (2014). Self-regulation from Educational Psychology to L2 Pedagogy: an Alternative to Language Learning Strategies. *International Journal of Applied Linguistics and English Literature, 3*(1), 240–244.

Barnett, M. (1989). *More than Meets the Eye*. Englewood Cliff, NJ: Prentice Hall.

Bayazidi, A. (2017). Assessing Reliability of Two Versions of Vocabulary Levels Tests in Iranian Context. *Advances in Language and Literary Studies, 8*(1), 30–43.

Bernaus, M., & Gardner, R. C. (2008). Teacher motivation strategies, student perceptions, student motivation, and English achievement. *Modern Language Journal*, 92(3), 387–401. https://doi.org/10.1111/j.1540-4781.2008.00753.x

Bernstein, B. (1999). Vertical and Horizontal Discourse: an essay. *British Journal of Sociology of Education, 20*, 157–173.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28*, 69–83.

Bialystok, E. (1981). The role of conscious strategies in second language proficiency. *Modern Language Journal, 65*, 24–35.

Bloom, B. S. (1984). *Taxonomy of educational objectives*. Boston: Allyn and Bacon.

Bloomfield, L. (1926). A set of Postulates for the Science of Language. *Language, 2*(3), 153–164.

Bloomfield, L. (1933). *Language*. Henry Holt.

Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). Academic Press.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective in assessment and intervention. *Applied Psychology: An International Review, 54*(2), 199–231.

Boers, F., & Lindstromberg, S. (2006). Cognitive linguistics application in second or foreign language instruction: Rationale, proposals, and evaluation. In G. Kristiansen, et al. (Eds.), *Cognitive Linguistics: Current Applications and Future Perspectives* (pp. 305-358). New York: Mouton de Gruyter.

Boers, F., & Lindstromberg, S. (2008). *Optimizing a Lexical Approach to Instructed Second Language Acquisition*. Basingstoke: Palgrave Macmillan.

Bowen, J.D., Madsen, H., & Hilferty, A. (1985). *TESOL techniques and procedures.* Cambridge: Newbury House.

Breen, M. P. (1987). Learner Contributions to Task Design. In C. Murphy & D. Candlin (Eds.), *Language Learning Tasks* (pp. 23–46). N.J: Prentice-Hall: Englewood Cliffs.

Brown, H. D. (2000). *Principles of Language Learning and Teaching* (4th ed.). Longman.

Bruton, A. (2011a). Are the differences between CLIL and non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal and Moore (2010). *Applied Linguistics, 32*(2), 236–241.

Bruton, A. (2011b). Is CLIL so beneficial, or just selective? Re-evaluating some of the Research. *System, 39*, 523–532.

Bruton, A. (2013). CLIL: some of the reasons why... and why not. *System, 41,* 587–597.

Brumfit, C. J., & Johnson, K. (1979). *The Communicative Approach to Language Teaching*. Oxford University Press.

Bulté, B., Housen, A., Pierrard, M., and Van Daele, S. (2008). Investigating lexical proficiency development over time - The case of Dutch-speaking learners of French in Brussels. *Journal of French Language Studies*, *18*(03), 277 – 298.

Cadierno, T., Hansen, M., Lauridsen, J., Eskildsen, S. W., Fenyvesi, K., Jensen, S. H., aus der Wieschen, M. V. (2020). Does younger mean better? Age of onset, learning rate and short-term L2 proficiency in young Danish learners of English. *VIAL- Vigo International Journal of Applied Linguistics 17*, 57-86.

Calderón, A. M. (2013). The effects of L2 learner proficiency on depth of processing, levels of awareness, and intake. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor in Richard Schmidt* (pp. 103–121). University of Hawai'i, National Foreign Language Resource Center.

Cameron, L. (2001). *Teaching languages to young learners*. Cambridge: Cambridge University Press.

Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, *1*, 1–47.

Canga Alonso, A. (2013a). Receptive vocabulary size of secondary Spanish EFL learners. *Revista de Lingüísticas y Lenguas Aplicadas, 8*, 66–75.

Canga Alonso, A. (2013b). The receptive vocabulary of 6th-grade primary-school students in CLIL instruction: A preliminary study. *Latin American Journal of Content and Language Integrated Learning, 6* (2), 22–41.

Canga Alonso, A. (2015a). Receptive Vocabulary of CLIL and Non-CLIL Primary and Secondary School Learners. *Complutense Journal of English Studies, 23*, 59–77.

Canga Alonso, A. (2015b). The receptive vocabulary size of Spanish 5th grade Primary School students in CLIL and Non-CLIL instruction. *ES: Revista de Filología Inglesa, 36*, 63–85.

Canga Alonso, A., & Arribas García, M. (2014). The benefits of CLIL instruction in Spanish students' productive vocabulary knowledge. *Encuentro Journal, 24*, 15–31.

Caro, K., & Rosado-Mendinueta, N. (2017). Lexis, Lexical Competence and Lexical Knowledge: A Review. *Journal of Language Teaching and Research*, *8*, 205–213.

Carter, R. (1998). *Vocabulary: Applied Linguistic Perspectives*. London & New York: Routledge.

Castellano-Risco, I. (2018). Receptive vocabulary and learning strategies in secondary school CLIL and non-CLIL learners. *Onomázein, 40*(2), 28–48.

Castellano-Risco, I., Piquer-Píriz, A. M. & Alejo González, R. (2020). The development of receptive vocabulary in CLIL learners: Is it all a matter of exposure? *System*, 91. DOI: 10.1016/j.system.2020.102263

Celaya, M.L., & Ruiz de Zarobe, Y. (2008). First languages and age in CLIL and non-CLIL contexts. *International CLIL Research Journal, 3*(1), 60-66.

Celce-Murcia, M. (2008). Rethinking the Role of Communicative Competence in Language Teaching. In E. Alcón Soler & M. P. Safont Jordá (Eds.), *Intercultural Language use and language Learning* (pp. 41–57). Springer.

Celce-Murcia, M., Dörnyei, Z., & Thurrel, S. (1995). Communicative Competence: A Pedagogically Motivated Model with Content Specifications. *Issues in Applied Linguistics*, *6*(2), 5–35.

Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code mixing. In M. P. García Mayo, & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language*. UK: Multilingual Matters.

Cenoz, J., Genesee, D., & Gorter, D. (2013). Critical analysis of CLIL: taking stock and looking forward. *Applied Linguistics, 35*, 1–21.

Chamot, A. U. (1987). The learning strategies of ESL students. In A. Wenden & J. Rubin (Eds.), *Learner strategy in language learning*. Englewood Cliffs: Prentice Hall.

Chamot, A. U. (2004). Issues in language learning strategy research and teaching. *Electronic Journal of Foreign Language Teaching, 1*(1), 14–26.

Chamot, A. U., O'Malley, J., Küpper, L., & Impink-Hernandez, M. V. (1987). *A Study of Learning Strategies in Foreign Language Instruction: First Year Report*. Rosslyn, VA: InterAmerica Research Associates.

Chapelle, C. A. (1998). Construct definition and validity inquiry in SLA research. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 32-70). Cambridge: Cambridge University Press.

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics, 31*, 693–713.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The MIT Press.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 3–23). Norwood, N.J.: Ablex.

Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Teaching & Technology*, *11*(3), 38–63.

Cohen, A.D. (1996). *Second language learning and use strategies: clarifying the issues*. Center for Advanced Research on Language Acquisition University of Minnesota, Minneapolis.

Cohen, A.D. (1998). Strategies and processes in test-taking and SLA. In L. F. Bachman & A. D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 90-111). Cambridge University Press.

Cohen, A. D. (2014). *Strategies in Learning and Using a Second Language.* London and New York: Routledge.

Cohen, A. D., & Macaro, E. (2007). *Language Learner Strategies: 30 years of Research and Practice*. Oxford University Press.

Cohen, A. D., Weaver, S. J., & Li, T.-Y. (1998). The impact of strategies-based instruction on speaking a foreign language. In A. D. Cohen (Ed.), *Strategies in Learning and Using a Second Language* (pp. 107–156). London and New York: Longman.

Cohen, L., & Manion, L. (1994*). Research methods in education*. London: Croom Helm.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*, 240–247.

Cook, V. (2001). *Second language learning and language teaching*. London: Hooder Education.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge University Press.

Council of Europe. (2017). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion volume with new descriptors.*

Council of the European Union. 89/489/EEC: Council Decision of 28 July 1989 establishing an action programme to promote foreign language competence in the European Community (Lingua), 89/489/EEC § (1989). Retrieved from https://publications.europa.eu/en/publication-detail/-/publication/339492e9-b8e3-4f42-9531-81886a6369c5/language-en

Coxhead, A. (2000). A new Academic World List. *Tesol Quartely, 34*(2), 213–239.

Coxhead, A., Nation, P. & Sim, D. (2015). Measuring the Vocabulary Size of Native Speakers of English in New Zealand Secondary Schools. *New Zealand Journal of Educational Studies, 50*(1), 121-135.

Coyle, D. (2007). Content and Language Integrated Learning: Towards a connected research agenda for CLIL pedagogies. *International Journal of Bilingual Education and Bilingualism, 10*, 543–562.

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and Language Integrated Learning*. Cambridge: Cambridge University Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671–684.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology, 104*(3), 268–294.

Cronbach, L. J. (1942). Measuring knowledge of precise word meaning. *The Journal of Educational Research, 36*(7), 528–534.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the age question and some other matters. Working Papers on Bilingualism, 19, 121–129.

Cummins, J. (1998). Immersion education for the millennium: What have we learned from 30 years of research on second language immersion? In M. R. Childs & R. M. Bostwick (Eds.), *Learning through two languages: Research and practice. Second Katoh Gakuen International Symposium on Immersion and Bilingual Education* (pp. 34–47). Japan.

Cummins, J. (2000). *Language, Power, and Pedagogy: Bilingual Children in the Crossfire*. Clevedon: Multilingual Matters.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In N. H. Hornberger (ed.), *Encyclopedia of Language and Education*. Springer.

Daller, H., Milton, J., & Treffers-Daller, J. (2007). *Modelling and assessing vocabulary knowledge*. Cambridge University Press Cambridge.

Dallinger, S., Jonkmann, K., Hollm, J., & Fiege, C. (2016). The effect of content and language integrated learning on students' English and history competences – Killing two birds with one stone? *Learning and Instruction, 41*, 23–31. https://doi.org/10.1016/j.learninstruc.2015.09.003

Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) classrooms*. Amsterdam: John Benjamins.

Davies, M. (2008). *Corpus of Contemporary American English (COCA)*. Retrieved from http://corpus.byu.edu/coca.

Davidson, D., & Raschke, V. R. (2009). Metalinguistic awareness in monolingual and bilingual children and its relationship to receptive vocabulary scores and performance on a reading readiness test. In J. Chandlee, M. Franchini, S. Lord, & M. Rheiner (Eds.), *Boston University Conference on Language Development 33 Online Proceedings Supplement*. Somerville, MA: Cascadilla Press.

Davies, M. (2008). *The corpus of contemporary American English: 450 million words, 1990-present.*

Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing*, *25*(4), 447–464.

De Graaff, R., Koopman, G. J., Anikina, Y., & Westhoff, G. (2007). An Observation Tool for Effective L2 Pedagogy in Content and Language Integrated Learning (CLIL). *The International Journal of Bilingual Education and Bilingualism, 10*, 603–624.

Decarrico, J. S. (2001). Vocabulary learning and teaching. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed., pp. 285–299). Boston: Heinle & Heinle.

Dóczi, B. (2020). An overview of conceptual models and theories of lexical representation in the mental lexicon. In S. Webb (Ed.), *The Routledge Book of Vocabulary Studies*. New York: Routledge.

Dóczi, B., & Kormos, J. (2016). *Longitudinal developments in vocabulary knowledge and lexical organization*. Oxford University Press.

Dörnyei, Z. (2005). *The Psychology of Language Learner. Individual differences in Second Language Acquisition.* Lawrence Erlbaum Associates.

Dörnyei, Z. (2009). *The psychology of second language acquisition.* Oxford: Oxford University Press.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of language learner revisited.* New York: Routledge.

Dreyer, C. and Oxford, R.L. (1996) Learning Strategies and Other Predictors of ESL Proficiency among Afrikaans Speakers in South Africa. In: Oxford, R.L., Ed., *Language Learning Strategies around the World: Cross-Cultural Perspectives*, University of Hawaii, Second Language Teaching & Curriculum Center, Honolulu, 61-74.

Dunn, L. M. (1959). *Peabody Picture Vocabulary Test manual.* Minneapolis: American Guidance Service.

Ehram, M., & Oxford, R. L. (1989). Effects of sex differences, career choice, and psychological type on adult language learning strategies. *The Modern Language Journal, 72*(3), 253–265.

Ehram, M., & Oxford, R. L. (1990). Adult language learning styles and strategies in an intensive training setting. *The Modern Language Journal, 74*(3), 313–327.

Ehrman, M., Leaver, B., & Oxford, R. (2003). A brief overview of individual differences in second language learning. *System, 31,* 313–330.

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences*, *8*(7), 301–306. https://doi.org/10.1016/j.tics.2004.05.003

Ellis, R. (1994). *The Study of Second Language Acquisition.* London: Oxford University Press.

Ellison de Matos, M. E. (2014). *CLIL as a Catalyst for Developing Reflective Practice in Foreign Language Teacher Education* (Unpublished PhD dissertation). Universidade do Porto, Porto.

European Commission. (1995). *White Paper on Education and Training. Teaching and Learning: Towards the Learning Society*. Brussels: European Commission.

European Commission. (2002). Presidency conclusions. Barcelona European Council. 15 and 16 March, 2002.

European Commission. (2003). *Promoting Language Learning and Linguistic Diversity*. Brussels: European Commission.

European Commission. (2006). *A New Framework Strategy for Multilingualism*. Brussels: European Commission.

European Commission. (2008). *Multilingualism: an Asset for Europe and a Shared Commitment*. Brussels: Commission of the European Communities. Retrieved from http://ec.europa.eu/education/languages/pdf/com/2008_0566_en.pdf

European Council. (1985). *European Council in Milan*. Milan: European Council.

European Parliament. (1983). *RESOLUTION on measures in favour of minority languages and cultures* (No. C 68/103).

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods, 4*(3), 272–299.

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *RELC Journal, 31,* 105–119.

Fernández Fontecha, A., & Terrazas Gallego, M. (2012). The role of motivation and age in vocabulary knowledge. *Vigo International Journal of Applied Linguistics, 9,* 39–62.

Fernández Fontecha, A. (2014). Motivation and Vocabulary Breadth in CLIL and EFL Contexts. Different age, Same Time of Exposure. *Complutense Journal of English Studies, 23*, 79–95. http://dx.doi.org/10.5209/rev_EIUC.2012.v20.39993

Fernández-Sanjurjo, J., Fernández-Costales, A., & Arias Blanco, J. M. (2017). *Analysing students' content-learning in science in CLIL vs. non-CLIL programmes: empirical evidence from Spain. International Journal of Bilingual Education and Bilingualism*.

Flege, J. E. & MacKay, I. R. A. (2011). What accounts for age effects on overall degree of foreign accent? In M. Wrembel, M. Kul, & K. Dziubalska-Kołaczyk (Eds.), *Polish studies in English language and literature: Achievements and perspectives in SLA of speech* (pp. 65-82). Frankfurt am Main, Germany: Peter Lang.

Folse, K. S. (2004). Myths about teaching and I second language vocabulary: what recent research says. *TESL Reporter, 37(*2), 1–13.

Fries, C. C. (1945). *Teaching and Learning English as a Foreign Language*. The University of

Michigan Press.

Gallardo del Puerto F. & Gómez Lacabez, E. (2017). Oral production outcomes in CLIL: an attempt to manage amount of exposure. *European Journal of Applied Linguistics, 5*(1), 31-54.

García López, M. (2000). Estrategias de aprendizaje de vocabulario de inglés utilizadas por los estudiantes de secundaria. *Lenguaje y Textos, 15*, 61–70.

García Mayo, M. P., & García Lecumberri, M. L. (Eds.). (2003). *Age and the Acquisition of English as a Foreign Language* (Vol. 4). Clevedon (U.K): Multilingual Matters.

García Mayo, M. P. & Gutiérrez-Mangado, J. (Eds.). (2020). *Studies in Second Language and Teaching. Special issue: English language learning in primary schools: variables at play, 10* (3).

Gobierno de España. *Ley Orgánica 2/2006, de 3 de mayo, de Educación* (2006).

Gobierno de España. Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa, Pub. L. No. 12886, 97858 (2013).

Godfroid, A., Boers, F., Housen, A. (2013). AN EYE FOR WORDS: Gauging the Role of Attention in Incidental L2 Vocabulary Acquisition by Means of Eye-Tracking. *Studies in Second Language Acquisition, 35* (3), 483-517.

Goulden, R., Nation, I. S. P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics, 11*, 341–363.

Gyllstad, H., Vilkaite, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats. Issues with guessing and sampling rates. *International Journal of Applied Linguistics, 166*(2), 276–303.

González-Fernández, B., & Schmitt, N. (2019). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics*.

Gouin, F. (1880). *Essai sur une réforme des méthodes d'enseignement. Exposé d'une nouvelle méthode linguistique. L'art d'enseigner et d'étudier les langues.* Paris: Fischbacher.

Graham, S., Courtney, L., Tonkyn, A., & Marinis, T. (2016). Motivational trajectories for early language learning across the primary-secondary school transition. *British Educational Research Journal, 42*, 682-702.

Gavriilidou, Z. (2004). Language learning strategy use in learning Greek as a second language. *Proceedings of the 6th International Conference in Greek Linguistics*. Retrieved from http://www.philology.uoc.gr/conferences/6thICGL/ebook%5Cdefault.htm

Green, J.M., & Oxford, R. (1995). A Closer Look at Learning Strategies, L2 Proficiency, and Gender. *TESOL Quarterly, 29*(2), 261-297.

Griffiths, C. (2003). Patterns of language learning strategy use. *System, 31*, 367–383.

Griffiths, C. (2004). *Language Learning Strategies: Theory and Research.*

Griffiths, C. (2008). Strategies and good language learners. In Carol Griffiths (Ed.), *Lessons from good language learners* (pp. 83–98). Cambridge: Cambridge University Press Cambridge.

Griffiths, C. (2013). *The strategy factor in successful language learning.* Bristol: Multilingual Matters.

Gu, Y. (2002). Gender, Academic Major, and Vocabulary Learning Strategies of Chinese EFL Learners. *RELC Journal, 33*(1), 35- 54.

Gu, Y. (2012). Learning Strategies: Prototypical Core and Dimensions of Variation. *Studies in Self-Access Learning Journal, 3*(4), 330–356.

Gu, Y., & Johnson, R. K. (1996). Vocabulary Learning Strategies and Language Learning Outcome. *Language Learning, 46*, 643–679.

Gyllstad, H., Vilkaite, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats. Issues with guessing and sampling rates. *International Journal of Applied Linguistics, 166*, 276–303.

Halliday, M. A. K. (1990). On Grammar and Grammatics. In R. Hasan, C. Cloran & D. G. Butt (eds.), *Functional Descriptions, Theory in practice* (pp. 1-39). Amsterdam/ Philadelphia: John Benjamins.

Hamzah, M. S.G., Kafipur, R., & Abdullah, S. K. (2009). Vocabulary Learning Strategies of Iranian Undergraduate EFL Students and its Relation to their Vocabulary Size. *European Journal of Social Sciences, 11*(1), 39-50.

Hatami, S., & Tavakoli, M. (2012). The Role of Depth versus Breadth of Vocabulary Knowledge in Success and Ease in L2 Lexical Inferencing. *TESL Canada Journal, 1,* 1–21.

Hayati, A. M. (2005). A comparative study of using bilingual and monolingual dictionaries in reading comprehension of intermediate EFL students. *The Reading Matrix, 5*(2), 61-66.

Hedge, T. (2000). *Teaching and learning in the language classroom*. New York, NY: Oxford University Press.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition, 21*(2), 303–317. Retrieved from https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/article/three-dimensions-of-vocabulary-development/757C3EE5BBA3DDD5110C77259E4E6E40

Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research, 19*(1), 70–88.

Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, *8*, 689–696.

Hong-Nam, K., & Leaven, A.G. (2006). Language learning strategy use of ESL students in an intensive English learning context. *System, 34*(3), 399-415.

Hornby, A. S. (1950). The situational approach in language teaching. *English Language Teaching*, 4, 98-104, 121-8, 150-6.

Housen, A., Bulté, B., Pierrard, M., & Van Daele, S. (2008). Investigating lexical proficiency development over time – the case of Dutch-speaking learners of French in Brussels. *French Language Studies, 18,* 1–22.

Howatt, P. R., & Widdowson, H. G. (2004). *A History of English Language Teaching*. Oxford University Press.

Howitt, D., & Cramer, D. (2000). *An introduction to statistics in psychology: A complete guide for students* (2nd ed.). London: Prentice Hall.

Hsieh, H. C., Moreno, N., & Leow, R. P. (2015). A comparison of level of awareness and depth of processing in two types of instructional media (C-FTF vs. CAI): Revisiting Hsieh (2008). In R. P. Leow, L. Cerezo, & M. Baralt (Eds.), *A psycholinguistic approach to technology and langauge learning*. De Gruyter.

Hsiao, T., & Oxford, R. (2002). Comparing theories of language learning strategies: A confirmatory factor analysis. *Modern Language Journal, 86*, 368–383.

Hsuch-chao, M., & Nation, P. (2000). Unknown Vocabulary Density and Reading Comprehension. *Reading in a Foreign Language*, *13*, 403–430.

Hulstijn, J. H. (2015). *Language Proficiency in Native and Non-native Speakers: Theory and research*. Amsterdam/ Philadelphia: John Benjamins.

Hüttner, J., & Rieder-Bünemann, A. (2010). A cross-sectional analysis of oral narratives by children with CLIL and non-CLIL instruction. In C. Dalton-Puffer, T. Nikula & U. Smit (eds.), *Language Use and Language Learning in CLIL classroom.* Amsterdam/ Philadelphia: John Benjamins.

Hymes, D. (1972). On Communicative Competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics*. Penguin Books.

Ibarretxe-Antuñano, I. (2004). What is cognitive linguistics? A new framework for the study of Basque. *Cahiers*, *10*(2), 3–31.

Intaraprasert, C. (2000). *Language learning strategies employed by engineering students learning English at tertiary level in Thailand*. University of Leeds, U.K.

Intaraprasert, C. (2004). *EST Students and Vocabulary learning strategies: A preliminary investigation.* Muang: Suranaree University of Technology, School of English.

Jaekel, N., Schurig, M., Florian, M., & Ritter, M. (2017). From Early Starters to Late Finishers? A Longitudinal Study of Early Foreign Language Learning in School. *Language Learning, 67*, 631–664.

Jexenflicker, S. (2009). *A comparative study on differences in language output between mainstream and CLIL students at two Austrian colleges of engineering, crafts and arts.* (Unpublished Thesis dissertation).

Jiménez Catalán, R. M. (2002). El concepto de competencia léxica en los estudios de aprendizaje y enseñanza de segundas lenguas. *Atlantis, 24*(2), 149–162.

Jiménez Catalán, R. M. (2003). Sex differences in L2 vocabulary learning strategies. *International Journal of Applied Linguistics, 12*, 54–77.

Jiménez-Catalán, R. M., & Terrazas Gallego, M. (2005). The receptive vocabulary of English Foreign Language young learners. *Journal of English Studies, 5,* 173–192.

Jiménez Catalán, R. M., & Ruiz de Zarobe, Y. (2009). The receptive vocabulary of EFL learners in two instructional contexts: CLIL versus non-CLIL learners. In Y. Ruiz de Zarobe & R. M. Jiménez Catalán (Eds.), *Content and Language Integrated Learning: Evidence from research in Europe* (pp. 81–93). UK: Multilingual Matters.

Jiménez Catalán, R.M., & Agustín-Llach, M.P. (2017). CLIL or time? Lexical profiles of CLIL and non-CLIL EFL learners. *System, 66*, 87-99.

Johnstone, R. (2009). An early start: What are the key conditions for generalized success? In J. Enever, J. Moon, & U. Raman (Eds.), *Young learner English language policy and implementation: International perspectives* (pp. 31-41). Reading, UK: Garnet.

Junta de Extremadura. ORDEN de 30 de agosto de 2000, por la que se establece y regula la impartición de lengua extranjera en el primer ciclo de la Educación Primaria, en el ámbito de la Comunidad Autónoma de Extremadura (2000).

Junta de Extremadura. ORDEN de 10 de agosto de 2001, por la que se establece y regula la implantación progresiva de lengua extranjera en el segundo ciclo de Educación Infantil (2001).

Junta de Extremadura. ORDEN de 27 de mayo de 2004, por la que se convoca y regula la impartición, con carácter experimental, de una segunda lengua extranjera en el tercer ciclo de educación primaria (2004).

Junta de Extremadura. DECRETO 82/2007, de 24 de abril, por el que se establece el Currículo de Educación Primaria para la Comunidad Autónoma de Extremadura. (2007).

Junta de Extremadura. DECRETO 83/2007, de 24 de abril, por el que se establece el Currículo de Educación Secundaria Obligatoria para la Comunidad Autónoma de Extremadura. (2007).

Junta de Extremadura. ORDEN de 1 de septiembre de 2008 por la que se regula la convocatoria para la implementación, con carácter experimental, del Portfolio Europeo de Lenguas en los centros educativos públicos de Extremadura, 2008050317 § (2008).

Junta de Extremadura. Instrucción 24/2013 de 5 de septiembre, de la Secretaría General de Educación, por la que se regula el funcionamiento del Programa de Lengua y Cultura Portuguesa en los centros educativos de la Comunidad Autónoma de Extremadura, 24/2014 § (2013).

Junta de Extremadura. DECRETO 103/2014, de 10 de junio, por el que se establece el currículo de Educación Primaria para la Comunidad Autónoma de Extremadura. (2014). Retrieved from http://doe.juntaex.es/pdfs/doe/2014/1140o/14040122.pdf

Junta de Extremadura. DECRETO 127/2015, de 26 de mayo, por el que se establece el currículo de Educación Secundaria Obligatoria y de Bachillerato para la Comunidad Autónoma de Extremadura, 2015040139 § (2015a).

Junta de Extremadura. DECRETO 98/2016, de 5 de julio, por el que se establecen la ordenación y el currículo de la Educación Secundaria Obligatoria y del Bachillerato para la Comunidad Autónoma de Extremadura, Pub. L. No. 20160401111, 17347 (2016).

Kaylani, C. (1996). The influence of gender and motivation on EFL learning strategy use in Jordan. In R. L. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives* (pp. 75–88). Honolulu: University of Hawai'i.

Ketabi, S., & Shahraki, S. H. (2011). Vocabulary in the Approaches to Language Teaching: From the Twentieth Century to the Twenty-First. *Journal of Language Teaching and Research, 2*(3), 726-731.

Konishi, M. (2003). Strategies for reading hypertext by Japanese ESL learners. *The Reading Matrix, 3*(3), 97–119.

Krashen, S. (1985). *The Input Hypothesis: Issues and implications*. New York: Longman.

Kremmel, B., & Schmitt, N. (2018). Vocabulary Levels Test. In *TESOL Encyclopedia of English Language Teaching*. New Jersey: Wiley-Blackwell.

Kučera, H., & Francis, W. N. (1967). *A Computational Analysis of Present Day American English*. Providence, Rhode Island: Brown University Press.

Lan, R., & Oxford, R. L. (2003). Language learning strategy profiles of elementary school students in Taiwan. *International Review of Applied Linguistics, 41*, 339–379.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar, Vol. I, Theoretical Prerequisites*. Stanford University Press.

Lasagabaster, D. (2008). Foreign Language Competence in Content and Language Integrated Courses. *The Open Applied Linguistics Journal, 1*(1). Retrieved from https://benthamopen.com/ABSTRACT/TOALJ-1-30

Lasagabaster, D. (2011). English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Learning and Teaching, 5*, 3–18.

Lasagabaster, D., & Doiz, A. (Eds.). (2016). *CLIL experiences in Secondary and Tertiary Education. In search of Good Practices*. Bern: Peter Lang.

Lasagabaster, D., & Doiz, A. (2017). A Longitudinal Study on the Impact of CLIL on Affective Factors. Applied Linguistics, 38, 688–712. https://doi.org/10.1093/applin/amv059

Lasagabaster, D., & López Beloqui, R. (2015). The Impact of Type of Approach (CLIL Versus EFL) and Methodology (Book-Based Versus Project Work) on Motivation. *Porta Linguarum, 23*, 41–57.

Laufer, B. (1986). Possible changes in attitudes towards vocabulary acquisition research. *IRAL, 24,* 69–75.

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Clevedon: Multilingual Matters.

Laufer, B. (1990). Why are some words more difficult than others? Some intralexical factors that affect the learning of words. *International Review of Applied Linguistics in Language Teaching, 28*, 293–307.

Laufer, B. (1992). How Much Lexis is Necessary for Reading Comprehension? In H. Bejoint & P. Arnaud (Eds.), *Vocabulary and Applied Linguistics*. Macmillan.

Laufer, B. (1998). The Development of Passive and Active Vocabulary in a Second Language: Same or Different? *Applied Linguistics, 19*, 255–271.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning, 54*, 399–436.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing, 16*, 33–51.

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: effects of Language Learning context. *Language Learning, 48*, 365–391.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*, 15–30.

Lawson, M. J., & Hogben, D. (1996). The Vocabulary-Learning Strategies of Foreign-Language Students. *Language Learning, 46*, 101-135.

Leńko-Szymańska, A. (2020). *Defining and assessing lexical proficiency.* New York & Oxon: Routledge.

Leow, R. P. (1998). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics*, *19*, 133–159.

Leow, R. P. (2001). Attention, awareness and foreign language behavior. *Language Learning*, *51*, 113–155.

Leow, R. P. (2012). Explicit and implicit learning in the L2 classroom: What does the research suggest? *The European Journal of Applied Linguistics and TEFL, 2*, 117–129.

Leow, R. (2015). *Explicit Learning in the L2 classroom: a student-centered approach.* New York: Routledge.

Lewis, M. (1993). *The lexical approach.* Hove: Language teaching publications.

Lewis, M. (1997). *Implementing the Lexical Approach: Putting Theory into Practice*. Hove: Language teaching publications.

Lewis-Beck, M. S., Bryman, A., & Futing Liao, T. (2004). *The SAGE Encyclopedia of Social Science Research Method*. Sage Publications.

Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: a web-based interface for bilingual research. *Language History Questionnaire: A Web-Based Interface for Bilingual Research, 38*, 202–210.

Lightbown, P. M., & Spada, N. (1999). *How languages are learned*. Oxford: Oxford University Press.

Lightbown, P. M., & Spada, N. (2014). *How languages are learned (4ᵗʰ ed.)*. Oxford: Oxford University Press.

López-Mezquita, M. T. (2005). La evaluación de la dimensión léxica a nivel receptivo: un marco para el diseño de instrumentos de medida. In M. L. Carrió Pasto (Ed*.), Perspectivas interdisciplinares de la lingüística aplicada* (pp. 381–390). España: Asociación Española de Lingüística Española.

Lorenzo, F., Casal, S., & Moore, P. (2009). The Effects of Content and Language Integrated Learning in European Education. Key Findings from the Andalusian Bilingual Sections Evaluation Project. *Applied Linguistics, 31*(3), 418-442.

Lorenzo, F., & Rodríguez, L. (2014). Onset and expansion of L2 cognitive academic language proficiency in bilingual settings: CALP in CLIL. *System, 47,* 64–73.

Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *The Modern Language Journal, 9,* 320–337.

Madsen, H. (1983). *Techniques in testing.* Oxford: Oxford University Press.

Manchón, R. M. (2001). Trends in the conceptualisations of second language composing strategies: A critical analysis. *International Journal of English Language Studies, 1*(2), 47–70.

Manzano Vázquez, B. (2014). Lexical transfer in the written production of a CLIL group and a non-CLIL group. *International Journal of English Studies, 14*(2), 57–76.

Marqués-Aguado, T., & Solís-Becerra, J. (2013). An Overview of Translation in language teaching methods: implications for EFL in secondary education in the region of Murcia. *Revista de Lingüística y Lenguas Aplicadas, 8,* 38-48.

Martínez Agudo, J.D. (2020). The impact of CLIL on English language competence in a monolingual context: a longitudinal perspective. *The Language Learning Journal,* 48, 36-47.

Marsh, D. (2002). *Content and Language Integrated Learning: The European Dimension-Actions, Trends and Foresight Potential.* Brussels: European Commission.

Matthews, J., & Cheng, J. (2015). Recognition of high frequency words from speech as a predictor of L2 listening comprehension. *System, 52,* 1-13.

McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.

Mclaughlin, B. (1987). *Theories of second-language learning*. Edward Arnold.

Meara, P. (1980). Vocabulary acquisition: a neglected aspect of language learning. *Language Teaching and Linguistics: Abstracts, 13*(4), 221–246.

Meara, P. (1982). Word associations in a foreign language. *Nottingham Linguistic Circular*, *11*(2), 29–38.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper, & P. R. Howatt (Eds.), *Interlanguage* (pp. 225–235). Edinburgh University Press.

Meara, P. (1992). *EFL vocabulary tests* (1st ed.). Wales University: Swansea Centre for Applied Language Studies.

Meara, P. (1996a). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and Competence in Second Language Acquisition* (pp. 35–53). Cambridge: Cambridge University Press.

Meara, P. (1996b). *The vocabulary knowledge framework*. Vocabulary Acquisition Research Group Virtual Library.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 109–121). Cambridge: Cambridge University Press.

Meara, P. (2005). Lexical frequency profiles: A Monte Carlo analysis. *Applied Linguistics, 26*, 32–47.

Meara, P. (2007a). Simulating Word Associations in an L2: Approaches to Lexical Organisation. *International Journal of English Studies*, *7*(2), 1–20.

Meara, P. (2007b). Simulating word associations in an L2: the effects of structural complexity. *Language Forum, 33*(2), 13–31.

Meara, P. (2010). *EFL vocabulary tests* (2nd ed.). Wales University: Swansea Centre for Applied Language Studies.

Meara, P., & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect, 16*(3), 5–19.

Meara, P., & Fitzpatrick, T. (2000). Lex 30: An improved method of assessing productive vocabulary in an L2. *System, 28,* 19–30.

Meara, P., & Milton, J. (2003). *X_Lex, The Swansea Levels Test*. Newbury: Express.

Meara, P., & Wolter, B. (2004). V_Linkis: Beyond vocabulary depth. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Angles on the English-speaking world* (pp. 85–96). Copenhagen: Museum Tusculanum Press.

Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL: Content Language Integrated Learning and Multilingual Education*. Oxford: Macmillan Education.

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition, and Pedagogy* (pp. 84–102). Cambridge: Cambridge University Press.

Merikivi, R., & Pietilä, P. (2014). Vocabulary in CLIL and in Mainstream Education. Journal of *Language Teaching and Research, 5*(3), 487–497.

Merino, J. A., & Lasagabaster, D. (2017). The effect of content and language integrated learning programmes' intensity on English proficiency: A longitudinal study. *International Journal of Applied Linguistics, 28*(1), 18-30.

Mihaljević Djigunović, J., & Lopriore, L. (2011). The learner: Do individual differences matter? In J. Enever (Ed.), *Early language learning in Europe* (pp. 43-60). London: British Council.

Milla, R., & Gutierrez-Mangado, M. J. (2019). Language learning strategy reported choice by bilingual children in CLIL: The effect of age, proficiency and gender in L3 learners of English. *System*, *87*.

Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. UK: Multilingual Matters.

Milton, J. (2010). The development of vocabulary breadth across the CEFR levels. In I. Vedder, I. Bartning, & M. Martin (Eds.), *Communicative proficiency and linguistic development: intersections between SLA and language testing research* (pp. 211–232). European Second Language Association.

Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindqvist, & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use*. European Second Language Association.

Milton, J., & Fitzpatrick, T. (2014). *Dimensions of vocabulary knowledge*. UK: Palgrave Macmillan.

Miralpeix, I. (2007). Lexical Knowledge in instructed Language Learning: The effects of age and exposure. *International Journal of English Studies*, 7(2), 61-83.

Mitchell, R., Myles, F. & Marsden, E. (2004). *Second Language Learning Theories*. New York & London: Routledge.

Monaikul, N. (2015). *Towards an Integrated Model of the Mental Lexicon*. University of Illinois at Urbana-Champaign.

Moreno-Espinosa, S. (2003). Vocabulary: Reviewing trends in EFL/ESL instruction and testing. *Odisea*, *4*, 97–112.

Mullins, P. (1992). *Successful Language Learning Strategies of Students enrolled at the Faculty of Arts, Chulalongkorn University, Thailand*. Doctoral Dissertation, United States International University, San Diego, California.

Muñoz, C. (2010). On how age affects foreign language learning. In A. Psaltou-Joycey & M. Mattheoudakis (Eds.), *Advances in research on language acquisition and teaching: Selected papers* (pp. 39–49). Thessaloniki, Greece: Greek Applied Linguistics Association.

Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics*, 35, 463e482.

Naiman, N., Frohlich, M., Stern, H., & Todesco, A. (1978). *The good language learner*. Toronto:

Nation, I. S. P. (1974). Techniques for teaching vocabulary. *English Teaching Forum, 12*(3), 18–21.

Nation, I. S. P. (1975). Teaching vocabulary in difficult circumstances. *English Language Teaching Journal, 29,* 115–120.

Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines, 5*(1), 12–25.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* New York City: Newbury.

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language.* Cambridge: Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63*, 59–82.

Nation, I. S. P. (2012a). *The BNC/COCA word family lists.*

Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9–13.

Nation, I. S. P. (2013). *Learning Vocabulary in Another Language* [2nd edition]. Cambridge: Cambridge University Press.

Nation, P., & Coady, J. (1988). Vocabulary and Reading. In Ronald Carter & M. McCarthy (eds.), *Vocabulary and Language Teaching* (pp. 97-110). London and New York: Routledge.

Nightingale, R., & Safont, P. (2019). Conversational Style and Early Academic Language Skills in CLIL and Non-CLIL Settings: A Multilingual Sociopragmatic Perspective. *English Language Teaching, 12* (2), 37-56.

Nikula, T. (2007). Speaking English in Finnish content-based classrooms. *World Englishes, 26* 206-223.

Nunan, D. (1989). *Designing Tasks for the Communicative classroom.* Cambridge: Cambridge University Press.

Nunan, D. (2004). *Task-Based Language Teaching.* Cambridge University Press.

Nyikos, M. (1987). *The effect of color and imagery as mnemonic strategies on learning and retention of lexical items in German* (Unpublished Doctoral Dissertation). Purdue University, West Lafayette.

O'Dell, F. (1997). Incorporating vocabulary into the syllabus. In Norbert Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 258–278). Cambridge: Cambridge University Press.

Olsson, E. (2015). Progress in English academic vocabulary use in writing among CLIL and non-CLIL students in Sweden. *Moderna språk,* 105, 51-74.

Olsson, E., & Sylvén, L. K. (2019). English Productive Proficiency. In Liss K. Sylvén (ed.)., *Investigating Content and Language Integrated Learning: Insights from Swedish High Schools* (117-135). Bristol: Multilingual Matters.

O'Malley, J., & Chamot, A. U. (1990). *Learning Strategies in Second Language Acquisition.* Cambridge: Cambridge University Press.

Ouazizi, K. (2016). The Effects of CLIL Education on the Subject Matter (Mathematics) and the Target Language (English). *Latin American Journal of Content and Language Integrated Learning, 9(1)*, 110–137.

Oxford, R. (1989). Use of Language Learning Strategies: a synthesis of studies with implications for strategy training. *System, 17* (2), 235–247.

Oxford, R. (1990). *Language Learning Strategies*. New York: Newbury House.

Oxford, R. (2011). *Teaching and Researching Language Learning Strategies*. Great Britain: Longman.

Oxford, R. (2017). *Teaching and Researching Language Learning Strategies: self-regulation in Context* (2nd ed.). New York: Routledge.

Oxford, R., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *The Modern Language Journal, 83*(3), 291–300.

Oxford, R., Nyikos, M., & Ehram, M. (1988). Vive la Différence? Reflections on Sex Differences in Use of Language Learning Strategies. *Foreign Language Annals, 21*, 321–329.

Palmer, H. E. (1917). *The Scientific Study and Teaching of Languages*. London: Oxford University Press.

Palmer, H. E. (1921). *The Principles of Language-study*. London, Harrap. London: Harrap.

Paran, A. (2013). Content and Language Integrated Learning: Panacea or Policy Borrowing Myth? *Applied Linguistics Review, 4*, 317–342. https://doi.org/10.1515/applirev-2013-0014

Park, G. (1997). Language learning strategies and English proficiency in Korean University students. *Foreign Language Annals, 30*, 211–221.

Pavlenko, A. (2009). Conceptual representation in the bilingual lexicon and second language vocabulary learning. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches*. Multilingual Matters.

Pavlov, I. P. (1897). *The work of the digestive glands*. London: Griffin.

Pavlov, I.P. (1927). *Conditioned Reflexes: An investigation of the physiological activity of the cerebral cortex*. Simply psychology: http://psychclassics.yorku.ca/Pavlov/lecture6.htm.

Peacock, M., & Ho, B. (2003). Student language learning strategies across eight disciplines. *International Journal of Applied Linguistics, 13*, 179–200.

Pérez Cañado, M. L. (2012). CLIL research in Europe: Past, present and future. *International Journal of Bilingual Education and Multilingualism, 1*5, 315–341.

Pérez Cañado, M. L. (2018). The effects of CLIL on L1 and content learning: Updated empirical evidence from monolingual contexts. *Learning and Instruction, 57,* 18–33. https://doi.org/10.1016/j.learninstruc.2017.12.002

Pérez-Cañado, M. L., & Lancaster, N. K. (2017). The effects of CLIL on oral comprehension and production: a longitudinal case study. *Language Culture and Curriculum, 30*, 300-316.

Pfenninger, S. E., & Singleton, D. (2016). Affect trumps age: A person-in-context relational view of age and motivation in SLA. *Second Language Research, 32,* 311-345.

Phakiti, A. (2008). Strategic competence as a fourth-order factor model: A structural equation modelling approach. *Language assessment quarterly, 5*, 20-42.

Philips, V. (1991). A look at learner strategy use and ESL proficiency. *CATESOL Journal, 4*(1), 57–67.

Piquer-Píriz, A. M., & Boers, F. (2019). La lingüística cognitiva y sus aplicaciones a la enseñanza de lenguas extranjeras. In I. Ibarretxe-Antuñano, T. Cadierno López, & A. Castañeda Castro (Eds.), *Lingüística cognitiva y español LE/L2* (pp. 52–70). Routledge.

Politzer, R. (1983). An Exploratory Study of Self-reported Language Learning Behaviors and Their Relation to Achievement. *Studies in Second Language Acquisition, 6*, 54–65.

Psaltou-Joycey, A. (2008). Cross-cultural differences in the use of learning strategies by students of Greek as a second language. *Journal of Multilingual and Multicultural Development, 29*, 310–324.

Psaltou-Joycey, A. (2010). *Language Learning Strategies in the Foreign Language Classroom.* Thessaloniki: University Studio Press.

Psaltou-Joycey, A., & Kantaridou, Z. (2009a). Foreign language learning strategy profiles of university students in Greece. *Journal of Applied Linguistics, 25,* 107–127.

Psaltou-Joycey, A., & Kantaridou, Z. (2009b). Plurilingualism, language learning strategy use and learning style preferences. *International Journal of Multilingualism, 39*, 460–474.

Psaltou-Joycey, A., & Sougari, A.-M. (2010). Greek young learners' perceptions about foreign language learning and teaching. In A. Psaltou-Joycey & M. Matthaioudakis (Eds.), *Advances in research on language acquisition and teaching: Selected papers* (pp. 387–401). Thessaloniki: Greek Applied Linguistics Association.

Psaltou-Joycey, A., Mattheoudakis, M., & Alexiou, T. (2012). *Language learning strategies in CLIL and non-CLIL classes: Which strategies do young learners claim they use?* Paper presented at the Thai TESOL Conference, Bangkok.

Purpura, J. E. (1994). *The role of learner strategies in language learning and tests.*

Qian, D. D. (2002). Investigating the Relationship Between Vocabulary Knowledge and Academic Reading Performance: An Assessment Perspective. *Language Learning, 52*, 513–536.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language: Selection, Acquisition and Testing* (pp. 209–227). Amsterdam & Philadelphia: John Benjamins Publishing.

Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly, 10*, 77–89.

Richards, J. C., Platt, J., & Platt, H. (1992). *Language teaching and applied linguistics (2nd ed.)*. Essex: Longman.

Richards, J. C., & Rodgers, T. S. (1986). *Approaches and Methods in Language Teaching*. Cambridge University Press.

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Pearson.

Rigney, J. (1978). *Learning strategies: a theoretical perspective. In H. F. O'Neil (Ed.), Learning strategies* (pp. 165–205). USA: Academic Press.

Rimmer, W. 2009. A closer look at CLIL. *English Teaching professional*, *64*, 4-6.

Rothman, J., & Guijarro-Fuentes, P. (2010). Input Quality Matters: Some Comments on Input Type and Age-Effects in Adult SLA. *Applied Linguistics, 31*, 301-306.

Rubin, D. (1981). Study of cognitive processes in second language learning. *Applied Linguistics, 11*, 117–131.

Rubin, J. (1975). What the 'Good Language Learner' can teach us. *TESOL Quarterly, 9* (1), 41–51.

Rubin, J. (1987). *Learner strategies: theoretical assumptions, research history and typology. In Learner Strategies in Language Learning*. New York: Prentice Hall.

Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque Country. *International CLIL Research Journal, 1*, 60–73.

Ruiz Gómez, D. A. & Nieto García, J. M. (2009). Las secciones bilingües en Secundaria y Bachillerato. Marco organizativo. Dificultades y propuestas. In A. Bueno González, J. M. Nieto García, & D. Cobo López (Eds.), *Atención a la diversidad en la enseñanza plurilingüe. I, II y III Jornadas Regionales de Formación del Profesorado* (CD-ROM). Jaén: Delegación Provincial de Educación de Jaén y Universidad de Jaén.

Sahbazian, S. (2004). *Perceived vocabulary learning strategies of Turkish university students* (Unpublished Ph.D. Dissertation). Oklahoma State University, Oklahoma, USA.

Salkind, N. J. (2010). *Encyclopedia of Research Design*. Sage Publications.

Sanjuan Álvarez, M. (1991). Qué significa 'conocer' una palabra: la complejidad de la competencia léxica. *Cuadernos de Filología Inglesa, 17,* 89–101.

Saussure, F. (1916). *Course in General Linguistics*.

Scarcella, R., & Oxford, R. (1992). *TAPESTRY OF LANGUAGE LEARNING: The Individual in the Communicative Classroom*. Heinle & Heinle.

Schachter, D. L. (1989). On the relation between memory and consciousness: Dissociable interactions and conscious experience. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving*. LEA.

Schmid, M. S. (2002). *First language attrition use and maintenance: The case of German Jews in anglophone countries*. John Benjamins.

Schmidt, R. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics*, *11*, 129–158.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness. In R. Schmidt (Ed.), *Attention and awareness in foreign language teaching and learning (Technical Report No. 9)* (pp. 1-64). Honolulu: University of Hawai'i at Manoa.

Schmitt, N. (1995). A fresh approach to vocabulary using a word knowledge framework. *RELC Journal 26*(1), 86-94.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy, *Vocabulary: Description, acquisition and pedagogy* (pp. 199-227). Cambridge: Cambridge University Press.

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning, 48*(2), 281–317.

Schmitt, N. (2000). *Vocabulary in language teaching.* Cambridge: Cambridge University Press.

Schmitt, N. (2010). *Researching vocabulary: A Vocabulary Research Manual*. Palgrave Press.

Schmitt, N. (2019). Understanding vocabulary acquisition, instruction, and assessment: a research agenda. *Language Teaching, 52*(2), 261-274.

Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2017). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching*, *50*, 212–226.

Schmitt, N., Nation, P., & Kremmel, B. (2019). Moving the field of vocabulary assessment forward: The need for more rigorous test development and validation. *Language Teaching*, 52(4).

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47, 484-503.

Schmitt, N., Schmitt, D., & Claphman, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Tests. *Language Testing, 18(*1), 55–88.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The Percentage of Words Known in a Text and Reading Comprehension. *Modern Language Journal*, *95*(1), 26–43.

Scot, D., & Beadle, S. (2014). *Improving the effectiveness of language learning: CLIL and computer assisted language learning* Brussels: European Commission.

Seliger, H. W., & Shohamy, E. (1989). *Second language research methods*. Oxford: Oxford University Press.

Sims, V. M. (1929). The reliability and validity of four types of vocabulary tests. *Journal of Educational Research, 20*, 91–96.

Singleton, D. (2007). How integrated is the integrated mental lexicon? In Z. Lengyel & J. Navracsics (Eds.), *Second language lexical processes: Applied linguistics and psycholinguistic perspectives* (pp. 3–16). Multilingual Matters.

Smit, U. 2007. Introduction. Vienna English Working Papers, 16(3): 3-5.

Smith, J. L. (1983). The Gouin Method. *American Annals of the Deaf, 38*(3), 177-186.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, *1*, 214–241.

Snow, R. E., Corno, L., & Jackson, D. N. (1996). Individual differences in affective and conative functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 243–310). New York: Macmillan.

Spolsky, B. (1995). *Measured Words.* Oxford: Oxford University Press.

Stern, H. (1992). *Issues and options in language teaching.* (P. Allen & B. Harley, Eds.). Oxford: Oxford University Press.

Stieglitz, G. J. (1955). The Berlitz Method. *The Modern Language Journal*, *39*(6), 300–310.

Stöffer, I. (1995). *University foreign language students' choice of vocabulary learning strategies as related to individual difference variable* (Unpublished Doctoral Dissertation). University of Alabama, U.S.A.

Surmont, J., Struy, E., Van Den Noort, M., & Van De Craen, P. (2016). The effects of CLIL on mathematical content learning: A longitudinal study. *Studies in Second Language Learning and Teaching, 6*, 319–337.

Sylvén, L. K. (2010). *Teaching in English or English teaching? On the effects of content and language integrated learning on Swedish learners' incidental vocabulary acquisition.* Göteborg: University of Gothenburg: Acta Universitatis Gothoburgensis.

Sylvén, L.K. (ed.) (2019). *Investigating Content and Language Integrated Learning. Insights from Swedish High Schools.* Bristol: Multilingual Matters.

Sylvén, L. S., & Thompson, A. S. (2015). Language learning motivation and CLIL: Is there a connection? *Journal of Immersion and Content Based Language Learning, 3*, 28-50.

Takac, V. P. (2008). *Vocabulary Learning Strategies and Foreign Language Acquisition.* Clevedon: Multilingual Matters.

Takala, S. (1984). *Evaluation of Students' Knowledge of English Vocabulary in the Finnish Comprehensive School.* (No. 350). Jyväskylä, Finland: Institute of Educational Research.

Tercanlioglu, L. (2004). Gender effect on adult foreign language learning. *Issues in Educational Research, 14*. Retrieved from http://iier.org.au/iier14/tercanlioglu.html

Terrazas Gallego, M., & Agustín-Llach, M. P. (2009). Exploring the Increase of Receptive Vocabulary Knowledge in the Foreign Language: A Longitudinal Study. *International Journal of English Studies, 9*(1), 113–133.

Thoma, D. (2009). *Strategic attention in Language Testing: Metacognition in Yes/No Business English Vocabulary Test*. Frankfurt: Peter.

Thornbury, S. (2002). *How to Teach Vocabulary*. Essex: Longman.

Thorndike, E. L., & Lorge, I. (1944). *The Teacher's word book of 30,000 words*. New York: Teacher's College, Columbia University.

Tilley, H. C. (1936). A Technique for Determining the Relative Difficulty of Word Meanings Among Elementary School Children. *The Journal of Experimental Education, 5*(1), 61–64.

Trask, R. (1995*). Language: The basics*. London: Routledge.

Trenchs, M. (1996). Writing strategies in a second language: Three case studies of learners using electronic mail. *Canadian Modern Language Review, 52*, 464–497.

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*, 78–102.

Tyacke, M., & Mendelsohn, D. (1986). Student needs: cognitive as well as communicative. *TESL Canada Journal, 1*, 171–183.

Vallbona González, A. (2014). *L2 Competence of Young Language Learners in Science and Arts CLIL and EFL Instruction Contexts. A Longitudinal Study*. Autonomous University of Barcelona. Unpublished Doctoral Dissertation.

van den Branden, K. (Ed.). (2006). *Task-based Language Education: From Theory to Practice*. Cambridge: Cambridge University Press.

van EK, J. A. (1975). *The Threshold Level*. Council of Europe.

Van Patten, P., & Benati, A. G. (2010). *Key Terms in Second Language Acquisition*. U.K.: Continuum International Publishing Group.

Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, *34*, 457–479.

Vandergrift, L. (1997). The strategies of second language (French) listeners: A descriptive study. *Foreign Language Annals, 30*(3), 387-409.

Velicer, W. F., & Jackson, D. N. (1990). Component Analysis versus Common Factor Analysis: Some issues in Selecting an Appropriate Procedure. *Multivariate Behavioral Research, 25*(1), 1–28.

Vilkaite-Lozdiene, L., & Schmitt, N. (2020). Frequency as a Guide for Vocabulary Usefulness: High-, Mid-, and Low-Frequency Words. In S. Webb (Ed.), *The Routledge Handbook of Vocabulary Studies* (pp. 81–96). Routledge.

Vlcková, K., Berger, J., & Völkle, M. (2013). Classification theories of foreign language learning strategies: an exploratory analysis. *Studia Paedagogica, 18*(4), 93-113.

Vrettou, A. (2009). Language learning strategy employment of EFL Greek-speaking learners in junior high school. *Journal of Applied Linguistics, 25,* 85–106.

Vrettou, A. (2011). *Patterns of language learning strategy use by Greek-speaking young Learners of English*. Aristotle University of Thessaloniki.

Vygostky, L. (1934). *Thought and Language*.

Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language Learning, 50*(2), 203–244.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition, 27*, 33–52.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language, 20*(2), 232–245.

Webb, S. (ed.) (2019). *The Routledge Handbook of Vocabulary Studies* (Routledge Handbooks in Linguistics). Oxon: Taylor and Francis.

Webb, S., & Nation, P. (2017). *How vocabulary is learned.* Oxford: Oxford University Press.

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *International Journal of Applied Linguistics*, *168*, 33–69.

Weinstein, C. E., Husman, J., & Dierking, D. R. (2000). Self-regulation interventions with a focus on learning strategies. In M. Boekaerts, P. R. Printrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 727–747). San Diego: Academic Press.

Wesche, M., & Paribakht, T. S. (1996). Assessing Second Language Vocabulary Knowledge: Depth Versus Breadth. *Canadian Modern Language Review, 53*, 13–40.

West, M. (1930). Speaking-vocabulary in a foreign language. One Thousand Words. *The Modern Language Journal*, *14*(7), 509–521.

West, M. (1953). *A General Service List of English Words*. London: Logman, Green & Co.

Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language Learning, 50*, 203–244.

Whong, M. (2011). *Language Teaching: Linguistic Theory in Practice.* Edimburgh: Edimburgu University Press

Widdowson, H. G. (1978). *Teaching Language as Communication*. Oxford: Oxford University Press.

Wilkins, D. A. (1972). *Linguistics and language teaching*. London: Edward Arnold.

Wilks, C., & Meara, P. (2002). Untangling word webs: Graph theory and the notion of density in second language word association networks. *Second Language Research*, *18*(4), 303–324.

Willis, J. (2004). Perspectives on task-based instruction: Understanding our practices, acknowledging different practitioners. In B. Leaver & J. Willis (Eds.), *Task-based instruction in foreign language education: Practices and Programs* (pp. 3–44). Washington D.C.: Georgetown University Press.

Wolff, D. (2005). Approaching CLIL. In Marsh, D. et al,. (Eds), *Project D3 - CLIL Matrix. The CLIL quality matrix*. Central Workshop Report 6/2005, 10-25.

Xanthou, M. (2010). Current trends in L2 vocabulary learning and instruction: Is CLIL the right approach? In *Advances in Research on Language Acquisition and Teaching: Selected Paper*s, 459–471.

Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching: Practice and Critique, 10*(4), 116–126.

Xhaferi, B., & Xhaferi, G. (2008). *Vocabulary learning strategies used by students at SEEU in terms of gender and teachers' attitudes toward teaching vocabulary*. Tetovo: South East European University.

Xue, G., & Nation, N. (1984). A university word list. *Language Learning and Communication, 3*(2), 215–229.

Yaremko, R. M., Harari, H., Harrison, R. C., & Lynn, E. (1986). *Handbook of Research and Quantitative Methods in Psychology: For Students and Professional*s. Hillsdale, NJ: Lawrence Erlbaum Associates.

Yen Dang, T.N. (2019). High-frequency words in academic spoken English: corpora and learners *ELT journal*, *74*(2), 146-155.

Yen Dang, T.N., Webb, S., & Coxhead, A. (2020). Evaluating lists of high-frequency words: Teachers' and learners' perspectives. *Language Teaching Research*, 1-25.

Yi Lo, Y., & Jeong, H. (2018). Impact of genre-based pedagogy on students' academic literacy development in Content and Language Integrated Learning (CLIL). *Linguistics and Education, 47*, 36-46.

Ying-Chun, L. (2009). Language Learning Strategy Use and English Proficiency of University Freshmen in Taiwan. *TESOL Quarterly, 43*(2), 255-280.

Zimmerman, C. (1997). Do Reading and Interactive Vocabulary Instruction Make a Difference? An Empirical Study. *TESOL Quarterly*, *31*, 121–140.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego: Academic Press.

APPENDICES

Appendix A: Vocabulary Learning Strategies classifications

A.1. Stoffer's classification (1995)

Stoffer's classification of Vocabulary Learning Strategies
• Strategies involving authentic language use
• strategies used to create mental linkages
• strategies involving physical action
• strategies used for self-motivation
• memory strategies
• strategies used to overcome anxiety
• strategies used to organize words
• strategies involving creating activities
• visual/ auditory strategies

A.2. Gu & Johnson (1996)

Gu & Johson's classification of Vocabulary Learning Strategies
• Metacognition regulation
• Dictionary strategies
• Note-taking strategies
• Memory strategies rehearsal
• Encoding
• Activation strategies
• Guessing strategies

A.3. Schmitt's taxonomy (1997)

Groups	Subgroup	Strategies
Discovery strategies	Determination	(1) Analysing the part of speech, (2) analysing affixes and roots, (3) Checking for L1 cognates, (4) analysing any available pictures or gestures, (4) guessing from textual context, (5) bilingual dictionary, (6) monolingual dictionary, (7) word lists and (8) flash cards.
	Social	(1) Asking teacher for an L1 translation, (2) asking teacher for paraphrasing or synonym of new word, (3) asking teacher for a sentence including the new word, (4) asking classmates for meaning and (5) discovering new meaning through group work activity.
Consolidation strategies	Social	(1) Studying and practicing meaning in a group, (2) teacher checking students' flash cards or word lists for accuracy and (3) interacting with native speakers.
	Memory	(1) Studying word with a pictorial representation of its meaning, (2) imagining word's meaning, (3) connecting word to a personal experience, (4) associating the word with its coordinates, (5) connecting the word to its synonyms and antonyms, (6) using semantic maps, (7) using 'scales' for gradable adjectives, (8) Peg Method, (9) Loci Method, (10) grouping words together to study them, (11) grouping words together spatially on a page, (12) using new word in sentences, (13) grouping words together within a storyline, (14) studying the spelling of the word, (15) studying the sound of a word, (16) saying new word aloud when studying,(17) imaging word form, (18) underlining initial letter of the word, (19) configuration, (20) using Keyword Method, (21) affixes and roots (remembering), (22) part of speech (remembering), (23) paraphrasing the words meaning, (24) using cognates in study, (25) learning the words of an idiom together, (26) using physical action when learning a word, and (28) using semantic feature grids
	Cognitive	(1) Verbal repetition, (2) written repetition, (3) word lists, (4) flash cards (5) taking notes in class, (6) using the vocabulary section in your textbook, (7) listening to tape of word lists, (8) putting English labels on physical objects and (9) keeping a vocabulary notebook.
	Metacognitive	(1) Using English-language media, (2) testing oneself with word tests, (3) using spaced word practice, (4) skipping or passing new word and (5) continuing to study word over time.

A.4. Nation (2001)

Group	Strategies
Planning	Choosing words
	Choosing aspects of word knowledge
	Choosing strategies and planning repetition
Source	Analysing the word
	Using word parts
	Learning from word cards
	Using context
	Using a dictionary
	Consulting a reference source in L1 and L2
	Using parallels in L1 and L2
Processing	Noticing
	Retrieving
	Generating

A.5. Tseng, Dornyei & Schmitt (2006)

Group	Strategies
Commitment control	Once the novelty of learning vocabulary is gone, I easily become impatient with it.
	When I feel stressed about vocabulary learning, I know how to reduce this stress.
	When I am studying vocabulary and the learning environment becomes unsuitable, I try to sort out the problem.
	When learning vocabulary, I have special techniques to achieve my learning goals.
Metacognitive control	When learning vocabulary, I have special techniques to keep my concentration focused.
	I feel satisfied with the methods I use to reduce the stress of vocabulary learning.
	When learning vocabulary, I believe I can achieve my goals more quickly than expected.
	During the process of learning vocabulary, I feel satisfied with the ways I eliminate boredom.
Satiation control	When learning vocabulary, I think my methods of controlling my concentration are effective.

	When learning vocabulary, I persist until I reach the goals that I make for myself.
	When it comes to learning vocabulary, I have my special techniques to prevent procrastination.
	When I feel stressed about vocabulary learning, I simply want to give up.
Emotion control	I believe I can overcome all the difficulties related to achieving my vocabulary learning goals.
	When learning vocabulary, I know how to arrange the environment to make learning more efficient
	When I feel stressed about my vocabulary learning, I cope with this problem immediately.
	When it comes to learning vocabulary, I think my methods of controlling procrastination are effective.
Environment control	When learning vocabulary, I am aware that the learning environment matters
	During the process of learning vocabulary, I am confident that I can overcome any sense of boredom.
	When feeling bored with learning vocabulary, I know how to regulate my mood in order to invigorate the learning process.
	When I study vocabulary, I look for a good learning environment.

A.6. Mayuree 2007

Group	Strategies
Strategies to Discover the Meaning of New Vocabulary Items:	(1) Guess the meaning from a single vocabulary item, (2) guess the meaning from contexts, (3) guess the meaning from word classes, such as nouns, verbs, adjectives, adverbs, (4) guess the meaning from grammatical structure of a sentence, (5) guess the meaning by analysing the structure of words (prefixes, roots, and suffixes) to discover the meaning of new vocabulary items, (6) guess the meaning from aural features, such as stress, intonation, pronunciation, (7) guess he meaning from real situations, (8) guess the meaning from gestures, (9) use an English-English dictionary, (10) use an English-Thai dictionary, (11) use a Thai-English dictionary, (12) ask classmates or friends, (13) ask teachers of English, (14) ask other people, such as members of one's family, native speakers of English.
Strategies to Retain the	(1) Say a single vocabulary item with its meanings repeatedly, (2) say vocabulary items in sentences repeatedly, (3) say vocabulary items with their lexical sets

Knowledge of Newly-Learned Vocabulary Items:	repeatedly, (4) say vocabulary items in rhymes repeatedly, (5) listen an English conversation of other people (classmates, friends, teachers, native speakers of English), (6) use vocabulary items to converse with classmates or friends, (7) use vocabulary items to converse with teachers of English to retain the knowledge of newly-learned vocabulary items (8) sing English songs, (9) review previous English lessons to retain the knowledge of newly-learned vocabulary items, (10) look at words' affixes (prefixes and suffixes), (11) make a vocabulary list with meanings and examples in one's notebook, (12) write vocabulary items with meanings on papers and stick them in one's bedroom, (13) group vocabulary items according to the synonyms and antonyms, (14) group vocabulary items according to the similarity of meaning, pronunciation and spelling, (15) do English exercises after class, (16) use newly-learned vocabulary items to practise writing in sentences, (17) associate pictures to vocabulary items, (18) look at real objects and associate them with vocabulary items, (19) associate newly-learned vocabulary items with previously-learned ones, (20) connect newly-learned vocabulary items to one's previous learning experience, (21) use semantic maps.
Strategies to Expand the Knowledge of New Vocabulary Items:	(1) Practise listening to English lectures, presentation, or cassettes of conversation, (2) listen to English songs, (3) listen to English radio programmes, (4) converse in English with classmates and friends in English, (5) converse with teachers of English in English, (6) converse with foreigners in English, (7) converse with foreigners in English through the Internet to expand the knowledge of vocabulary, (8) read English articles from different sources, such as texts, newspaper, brochures, leaflets, to expand the knowledge of vocabulary, (9) read a book of English-Thai conversation in various situations, (10) study vocabulary items from advertisements, public relations notices, traffic signs, etc.,(11) watch English programme channels on TV, (12) watch an English-speaking films with subtitles to expand the knowledge of vocabulary, (13) search for English information through the Internet to expand the knowledge of vocabulary, (14) practise using a dictionary regularly to expand the knowledge of vocabulary, (15) practise translating articles from English to Thai, or from Thai to English to expand the knowledge of vocabulary, (16) do extra English exercises from other sources, such as texts, newspapers, Internets, to expand the knowledge of vocabulary, (17) build a word-network to expand the knowledge of Vocabulary, (18) play English games, such as scrabble, crossword puzzles, to expand the knowledge of vocabulary, and (19) take an extra job at tour offices, hotels, etc. to expand the knowledge of vocabulary

Appendix B: Language history questionnaire

CUESTIONARIO PERFIL GENERAL DEL ALUMNADO

Estimados alumnos: mediante este cuestionario se pretende conocer algunos aspectos relativos a vuestra trayectoria en el aprendizaje de lenguas extranjeras, que nos serán muy útiles para explicar qué aspectos pueden influir en el aprendizaje de lenguas extranjeras.

Estas respuestas serán usadas exclusivamente con fines científicos y en ningún caso se usará nombre alguno.

Muchas gracias por vuestra colaboración.

1. DATOS DEL ALUMNO

1. Inicial de tu nombre: _____
2. Dos primeras letras de tu primer apellido: _____
3. Edad: _____
4. Sexo: ☐ Masculino   ☐ Femenino
5. Curso: ☐ 4º E.S.O bilingüe   ☐ 4º E.S.O sección no bilingüe
6. Lengua materna: _____

   Si la lengua materna de tus padres es diferente del español, indica cuál es a continuación:

   - Lengua materna del padre: _____

   - Lengua materna de la madre: _____

   Si la lengua de comunicación normal en tu casa no es el español, por favor, indica cuál es la lengua utilizada: _____

2. LENGUAS APRENDIDAS EN LA ESCUELA

1. En tu centro de Primaria, ¿ tuviste una enseñanza bilingüe (se enseñaban en otro idioma algunas asignaturas) ?   ☐ SÍ   ☐ NO

   -¿En qué idiomas estudiabas las otras asignaturas?

   Inglés ☐   Francés ☐   Portugués ☐

   - ¿En qué cursos has seguido una enseñanza bilingüe? (Marca con una X los cursos)

1º E.P.	2º E.P.	3º E.P.	4º E.P.	5º E.P.	6º E.P.	

-Marca con una X las asignaturas que estudiabas en Primaria en otro idioma:

Matemáticas	
Conocimiento del Medio	
Educación Artística	
Educación Física	
Educación para la ciudadanía	
Otra:	

2. ¿Qué idiomas has aprendido en Primaria? (marca con una X los cuadros que correspondan)

Lengua	1º Primaria	2º Primaria	3º Primaria	4º Primaria	5º Primaria	6º Primaria
Inglés						
Francés						
Portugués						

3. EXPERIENCIAS EXTRAESCOLARES CON LENGUAS

1. ¿Has estudiado o estudias alguna lengua extranjera fuera del colegio/instituto? (Por ejemplo en una academia) Sí ☐   No ☐

En caso afirmativo, indique qué idioma, dónde y cuánto tiempo

Idioma	Lugar (Academia, clases extraescolares en el colegio, campamentos..)	Número de meses/ años que has estudiado

Muchas gracias por tu colaboración

Appendix C: 2K and academic bands of the Vocabulary Levels Test administered at T1 and T2 (Schmitt et al. 2001)[6]

This is a vocabulary test.   You must choose the right word to go with each meaning. Write the number of that word next to its meaning.   Here is an example.

1   business
2   clock   _____ part of a house
3   horse   _____ animal with four legs
4   pencil   _____ something used for writing
5   shoe
6   wall

You answer it in the following way.

1   business
2   clock   __6__ part of a house
3   horse   __3__ animal with four legs
4   pencil   __4__ something used for writing
5   shoe
6   wall

Some words are in the test to make it more difficult.   You do not have to find a meaning for these words.   In the example above, these words are business, clock, and shoe.

If you have no idea about the meaning of a word, do not guess.   But if you think you might know the meaning, then you should try to find the answer.

---

[6] I would like to thank Professor Diane Schmitt for allowing me to reproduce this material.

Version 1   The 2,000 word level

1 birth
2 dust   _____ game
3 operation   _____ winning
4 row   _____ being born
5 sport
6 victory

1 choice
2 crop   _____ heat
3 flesh   _____ meat
4 salary   _____ money paid regularly for
5 secret   doing a job
6 temperature

1 cap
2 education   _____ teaching and learning
3 journey   _____ numbers to measure with
4 parent   _____ going to a far place
5 scale
6 trick

1 attack
2 charm   _____ gold and silver
3 lack   _____ pleasing quality
4 pen   _____ not having something
5 shadow
6 treasure

1 cream
2 factory   _____ part of milk
3 nail   _____ a lot of money
4 pupil   _____ person who is studying
5 sacrifice
6 wealth

1 adopt
2 climb   _____ go up
3 examine   _____ look at closely
4 pour   _____ be on every side
5 satisfy
6 surround

1 bake
2 connect   _____ join together
3 inquire   _____ walk without purpose
4 limit   _____ keep within a certain size
5 recognize
6 wander

1 burst
2 concern   _____ break open
3 deliver   _____ make better
4 fold   _____ take something to someone
5 improve
6 urge

1 original
2 private   _____ first
3 royal   _____ not public
4 slow   _____ all added together
5 sorry
6 total

1 brave
2 electric   _____ commonly done
3 firm   _____ wanting food
4 hungry   _____ having no fear
5 local
6 usual

Version 1   Academic Vocabulary

1 benefit
2 labor   _____ work
3 percent   _____ part of 100
4 principle   _____ general idea used to
5 source   guide one's actions
6 survey

1 element   _____ money for a special
2 fund   purpose
3 layer   _____ skilled way of doing
4 philosophy   something
5 proportion   _____ study of the meaning
6 technique   of life

1 consent
2 enforcement   _____ total
3 investigation   _____ agreement or permission
4 parameter   _____ trying to find information
5 sum   about something
6 trend

1 decade
2 fee   _____ 10 years
3 file   _____ subject of a discussion
4 incidence   _____ money paid for services
5 perspective
6 topic

1 colleague
2 erosion   _____ action against the law
3 format   _____ wearing away gradually
4 inclination   _____ shape or size of something
5 panel
6 violation

1 achieve
2 conceive   _____ change
3 grant   _____ connect together
4 link   _____ finish successfully
5 modify
6 offset

1 convert
2 design   _____ keep out
3 exclude   _____ stay alive
4 facilitate   _____ change from one thing
5 indicate   into another
6 survive

1 anticipate
2 compile   _____ control something skillfully
3 convince   _____ expect something will
4 denote   happen
5 manipulate   _____ produce books and
6 publish   newspapers

1 equivalent
2 financial   _____ most important
3 forthcoming   _____ concerning sight
4 primary   _____ concerning money
5 random
6 visual

1 alternative
2 ambiguous   _____ last or most important
3 empirical   _____ something different that
4 ethnic   can be chosen
5 mutual   _____ concerning people from
6 ultimate   a certain nation

Version 2   The 2,000 word level

1 copy
2 event   _____ end or highest point
3 motor   _____ this moves a car
4 pity   _____ thing made to be like
5 profit   another
6 tip

1 accident
2 debt   _____ loud deep sound
3 fortune   _____ something you must pay
4 pride   _____ having a high opinion of
5 roar   yourself
6 thread

1 coffee
2 disease   _____ money for work
3 justice   _____ a piece of clothing
4 skirt   _____ using the law in the right
5 stage   way
6 wage

1 clerk
2 frame   _____ a drink
3 noise   _____ office worker
4 respect   _____ unwanted sound
5 theater
6 wine

1 dozen
2 empire   _____ chance
3 gift   _____ twelve
4 opportunity _____ money paid to the
5 relief   government
6 tax

1 admire
2 complain   _____ make wider or longer
3 fix   _____ bring in for the first time
4 hire   _____ have a high opinion of
5 introduce   someone
6 stretch

1 arrange
2 develop   _____ grow
3 lean   _____ put in order
4 owe   _____ like more than something
5 prefer   else
6 seize

1 blame
2 elect   _____ make
3 jump   _____ choose by voting
4 manufacture _____ become like water
5 melt
6 threaten

1 ancient
2 curious   _____ not easy
3 difficult   _____ very old
4 entire   _____ related to God
5 holy
6 social

1 bitter
2 independent _____ beautiful
3 lovely   _____ small
4 merry   _____ liked by many people
5 popular
6 slight

Version 2 Academic Vocabulary

1 area
2 contract _____ written agreement
3 definition _____ way of doing something
4 evidence _____ reason for believing
5 method   something is or is not true
6 role

1 alter
2 coincide _____ change
3 deny _____ say something is not true
4 devote _____ describe clearly and exactly
5 release
6 specify

1 debate
2 exposure _____ plan
3 integration _____ choice
4 option _____ joining something into a
5 scheme   whole
6 stability

1 correspond
2 diminish _____ keep
3 emerge _____ match or be in agreement
4 highlight   with
5 invoke _____ give special attention
6 retain   to something

1 access
2 gender _____ male or female
3 implementation _____ study of the mind
4 license _____ entrance or way in
5 orientation
6 psychology

1 bond
2 channel _____ make smaller
3 estimate _____ guess the number or size
4 identify   of something
5 mediate _____ recognizing and naming
6 minimize   a person or thing

1 accumulation
2 edition _____ collecting things over time
3 guarantee _____ promise to repair a broken
4 media   product
5 motivation _____ feeling a strong reason or
6 phenomenon   need to do something

1 explicit
2 final _____ last
3 negative _____ stiff
4 professional _____ meaning `no' or `not'
5 rigid
6 sole

1 adult
2 exploitation _____ end
3 infrastructure   machine used to move
4 schedule   people or goods
5 termination _____ list of things to do at
6 vehicle   certain times

1 abstract
2 adjacent _____ next to
3 controversial _____ added to
4 global _____ concerning the whole world
5 neutral
6 supplementary

Appendix D: 2K and academic bands of Productive Levels Test (Nation & Laufer, 1999) administered at T2[7]

PRODUCTIVE VOCABULARY LEVEL TESTS (Nation & Laufer, 1999)

Complete the underlined words. The example has been done for you.

He was ri_____ a bicycle.--> He was riding a bike.

The 2000-word level

1. I'm glad we had this opp_____ to talk.
2. There are a doz_____ eggs in the basket.
3. Every working person must pay income t _____.
4. The pirates buried the trea_____ on a desert island.
5. Her beauty and cha_____ had a powerful effect on men.
6. La_____ of rain led to a shortage of water in the city.
7. He takes cr_____ and sugar in his coffee.
8. The rich man died and left all his we_____ to his son.
9. Pup must hand in their papers by the end of the week.
10. This sweater is too tight. It needs to be stret_____.
11. Ann intro_____ her boyfriend to her mother.
12. Teenagers often adm_____ and worship pop singers.
13. If you blow up that balloon any more it will bur_____.
14. In order to be accepted into the university, he had to impr_____ his grades.
15. The telegram was deli_____ two hours after it had been sent.
16. The differences were so sl_____ that they went unnoticed.
17. The dress you're wearing is lov_____.
18. He wasn't very popu_____ when he was a teenager, but he has many friends now.

---

[7] I would like to thank professor Nation for allowing me to reproduce this material.

The University Word List level

1. There has been a recent tr_____ among prosperous families towards a smaller number of children.

2. The ar_____ of his office is 25 square meters.

3. Phil_____ examines the meaning of life.

4. According to the communist doc_____, workers should rule the world.

5. Spending many years together deepened their inti_____.

6. He usually read the sport sec_____ of the newspaper first.

7. Because of the doctors' strike the cli_____ is closed today.

8. There are several misprints on each page of this te_____.

9. The suspect had both opportunity and mot to commit the murder.

10. They insp_____ all products before sending them out to stores.

11. A considerable amount of evidence was accum_____ during the investigation.

12. The victim's shirt was satu_____ with blood.

13. He is irresponsible. You cannot re_____ on him for help.

14. It's impossible to eva_____ these results without knowing about the research methods that were used.

15. He finally att_____ a position of power in the company.

16. The story tells us about a crime and subs_____ punishment.

17. In a hom_____ class all students are of a similar proficiency.

18. The urge to survive is inh_____ in all creatures.

Appendix E: Vocabulary Learning Strategies questionnaire

Cuestionario sobre estrategias de aprendizaje de vocabulario

Por favor, puntúa en las siguientes tablas, donde 1 significa nunca y 5 se corresponde con siempre, el uso que le das a las estas estrategias para aprender vocabulario.

1. Perfil del alumnado
    1. Sexo:
    2. Edad:
    3. ¿Estás en una sección bilingüe?
    4. En caso de que estés en una sección bilingüe, ¿cuánto tiempo llevas en ella?

	Cuando encuentro palabras nuevas, para averiguar lo que significa....	Nunca 1	a veces 2	muy a menudo 3	Siempre 4
1	Analizo el tipo de palabra que es (nombre, adjetivo, verbo..), por ejemplo en la frase "I go home", aunque no sepa lo que significa 'go', puedo averiguar que es un verbo, porque aparece el sujeto ('I') y sé que la otra palabra no es un verbo.				
2	Analizo sus prefijos o sufijos, por ejemplo,en la palabra 'unhappy', me fijo en el prefijo 'un-', que significa no, para averiguar que significa infeliz.				
3	Analizo si se parece a alguna palabra en castellano, por ejemplo, la palabra 'produce' en inglés se parece a producir.				
4	Uso las imágenes de los libros para intentar averiguar qué significa				
5	Consulto el significado en un diccionario bilingüe (español-inglés/ inglés-español)				
6	Pido al profesor que me dé el significado de la palabra				
7	Pido al profesor que me diga un sinónimo de esa palabra en la lengua extranjera, por ejemplo, si aparece la palabra 'big' y no sé lo que significa, el profesor podría decirme 'it's the opposite of small'				
8	Averiguo el significado preguntando a mis compañeros				

Appendix E: Vocabulary Learning Strategies questionnaire

	Una vez sé lo que significa la palabra, para memorizar su significado...	Nunca	a menudo	muy a menudo	Siempre
		1	2	3	4
9	Practico esa palabra en actividades en grupo, por ejemplo, si estoy aprendiendo verbos sobre la rutina diaria ('lay the table', 'do my homework') trato de usarlas en las actividades en grupo				
10	Relaciono la palabra con algún dibujo o imagen				
11	Conecto la palabra con una experiencia personal, por ejemplo, si tengo que memorizar la expresión "it's raining cats and dogs" puedo asociarlo a un día que no pude salir de casa porque llovía mucho.				
12	Asocio el significado de la palabra con un sinónimo o antónimo, por ejemplo si tengo que aprender la palabra 'happy' lo asocio a su antónimo 'sad'				
13	Agrupo las palabras que tienen relación para estudiarlas juntas, por ejemplo, si estoy estudiando el vocabulario de deportes y de rutinas, agrupo por un lado, todas las palabras que tengan que ver con el deporte ('football, basketball..'), y por otro lado las rutinas.				
14	Uso la palabra en una frase, por ejemplo, para recordar que 'many' se usa con nombres contables, recuerdo la frase "How many students are there in class?"				
15	Realizo físicamente la acción que indica la palabra, por ejemplo, si tengo que aprender lo que significa 'run', hago como si estuviera corriendo.				
16	Repito en voz alta muchas veces la palabra hasta que me la aprendo				
17	Escribo varias veces las palabras para recordar cómo se escribe				
18	Hago listas con el vocabulario que tengo que estudiar				
19	Escucho canciones y películas en inglés				
20	Ignoro una palabra que no sé que significa cuando comprendo en general lo que quiere decir el texto				
21	Repaso las palabras cada cierto tiempo para no olvidarlas				

Si utilizas algún truco más, por favor escríbelo en el siguiente cuadro:
Muchas gracias por tu colaboración