An integrated approach to incidental collocation learning: learner, item, and context effects

Inés de la Viña

Submitted in fulfilment of the degree of

Doctor of Philosophy by Research in

Linguistics

School of Cultures and Languages, Division of Arts & Humanities

University of Kent

April 2023

Abstract

Collocations, pairs of words that tend to occur together in discourse, are an important aspect of second language (L2) proficiency (Schmitt, 2000), but they stand out as being particularly troublesome for L2 learners (Laufer & Waldman, 2011). Given the limited classroom time available for explicitly teaching collocations, learning through meaning-focused input, such as reading, plays an important role in developing L2 learners' collocational knowledge. The majority of the studies that have investigated incidental collocation learning through reading-(while-listening) are limited in that they only focused on specific linguistic, contextual or individual factors, or failed to consider the impact of relevant confounding variables in L2 vocabulary learning (e.g., Pellicer-Sánchez, 2017; Sonbul & Schmitt, 2013; Szudarski & Carter, 2016; Toomer & Elgort, 2019; Webb & Chang, 2020). Additionally, they produced inconsistent findings, partly due to differences in treatment but also because they did not examine how various factors can contribute independently to learning, or even interact with each other. This dissertation adopted a multifactorial approach to researching collocation learning so as to glean a more comprehensive picture of the nature of collocations, while addressing some research gaps that emerge from existing studies. Three experiments were designed with these goals in mind.

Experiment 1 investigated the effect of reading mode (reading-only vs reading-while-listening), taking into account a range of factors (prior vocabulary knowledge, congruency, compositionality, corpus frequency) to examine the influence of such variables. Unlike previous studies, which have largely focused on fictional short stories as reading materials, this experiment used a single academic text to assess whether the text-related factors could affect learning. The results showed that reading-while-listening did not confer an advantage over reading-only, which contradicts prior studies that used narrative texts as reading materials. However, prior vocabulary knowledge and congruency - characteristics of individual learners and individual collocations, respectively - improved performance.

Experiment 2 examined the effect of repetition by seeding a series of contextually-supportive academic texts with various encounters with collocations. The results indicated that learning gains improved as the number of repetitions increased. Prior vocabulary knowledge again emerged as a reliable predictor of learning. Repetition interacted with prior vocabulary knowledge, suggesting that the facilitative effect of repetition decreases for learners with higher existing vocabulary knowledge.

Experiment 3 explored the role of contextual informativity on collocation learning. Specifically, it compared the effect of three levels of context informativity (low, mid, high) to evaluate whether the degree of contextual support affects learning. As in Experiment 2, a series of short academic texts were used as stimulus materials. The results showed that contextual informativity increased performance to a certain extent, but not beyond it. In particular, it showed that mid informative contexts led to higher gains than low and high informativity contexts. Prior vocabulary knowledge significantly predicted the learning gains.

Overall, the results of this thesis suggest that a multi-factor approach to collocation research can provide a comprehensive account of the effects that unfold in incidental learning conditions. This thesis extends the line of inquiry into the moderator variables that influence collocation learning. Taken as a whole, the results provide useful insights for the learning and teaching of collocations in a more systematic and efficient manner.

Dedicated to the memory of my parents,

Ana Pilar Rodríguez Rodríguez (November 9, 1954 - December 1, 2012) and

Galo Ricardo Sánchez de la Viña (January 5, 1954 - October 7, 2020)

Acknowledgments[1]

Undertaking this PhD has been a truly life-experience filled with many life (and world) events. I wish to thank the remarkable people who have made this journey a genuinely enriching experience.

First of all, I would like to express my heartfelt gratitude to my supervisors, Dr Gloria Chamorro and Dr Christina Kim. This PhD would not have been possible without their guidance and nurturing. Christina was never short of insightful comments. I am grateful to her for her patience and enthusiasm during our numerous data analysis discussions. Gloria was a constant source of encouragement throughout this journey, and I will be eternally grateful to her for remaining my supervisor even after she left Kent.

I am grateful to the many researchers and teachers who have lent me their time to discuss the ideas that appear in this thesis. I owe thanks in particular to Dr. Ana Pellicer-Sánchez for hosting me during my unconventional research stay at UCL, as well as for her practical advice and inspiration. I am grateful to Dr. Imma Miralpeix for her continued support and belief in my work. I am also very thankful to Dr. Beatriz González-Fernández for her encouragement and for leading me working on exciting projects. I must also thank my former teachers at the University of Oviedo for allowing me to interact with their students.

Especially helpful to me during this time were Irene and Rosa, without whom this journey would have not been the same. Thank you for the joyful and uplifting conversations. Thanks also to my office colleague and friend, Luke, for his positive notes on the whiteboard. To my "lingthusiasts" friends, Britt, Clara, Dani, Fer, Judit, and Julie-Anne, who spared time to help me improve my research and shared the stresses and achievements of this journey. Thanks also to my Canterbury friends, who have made this experience much more enjoyable.

I also wish to thank my cousin Javi, for his invaluable comments about statistics, and my lifelong friend María, for her continued friendship and the therapeutic discussions.

Special thanks to George, who put up with my distractions and brought me tea and snacks while I sat in my chair for hours. I am extremely grateful for your patience and unparalleled support.

Last but not least, I would like to thank my family for their love and patience in making this journey meaningful. To my parents, Galo *y* Kuka, who set me off on the road to this journey a long time ago. To Ana, my big sister, whose belief in me has kept my spirits and motivation high throughout this process. To my little nephews, Eri and Leo, for making sure I took breaks to play hide-and-seek. *Finalmente, también debo dar las gracias a mis segundos padres*, *Álex y Tere, por su apoyo incondicional durante estos años (¡y por el escritorio!).*

Declaration

I declare that the work presented in this thesis is my own and was conducted during my time as a PhD student at the University of Kent. Some parts of this thesis have been submitted to peer-reviewed journals:

Experiment 1, presented in Chapter 4, has been submitted as a paper co-authored with Dr Christina Kim and Dr Gloria Chamorro and is currently under review.

The experiments presented in Chapters 5 and 6 are currently being prepared for submission.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AM	Association measure
BNC	British National Corpus
CA	Contrastive Analysis
CLT	Cognitive Load Theory
CPH	Critical Period Hypothesis
COCA	Corpus of Contemporary American English
CEFR	Common European Framework of Reference
DCT	Dual Coding Theory
EAP	English for Academic Purposes
EFL	English as a foreign language
ESL	English as a Second Language
FL	Foreign language
HEI	Higher Education Institution
ILH	Involvement Load Hypothesis
L1	First language
L2	Second language
LEAP-Q	Language Experience and Proficiency Questionnaire
MI	Mutual information
MWU	Multiword unit
PPVT	Peabody Picture Vocabulary Test
RO	Reading-only
RWL	Reading-while-listening
SCT	Sociocultural Theory
SLA	Second language acquisition
VKS	Vocabulary Knowledge Scale
VLT	Vocabulary Levels Test
UG	Universal grammar
ZPD	Zone of Proximal Development

1. Introduction

Vocabulary is crucial for effective communication and an essential aspect of learning a second (L2)/foreign (FL) language. Learning vocabulary entails not only learning individual words, but also learning them in their recurrent word combinations known as multiword units (MWUs). MWUs are "perhaps the essential element" of language, according to Schmitt (2010:146), because they account for at least one-third to one-half of all discourse (Conklin & Schmitt, 2012). Collocations, which are frequently defined pairs of words that tend to occur together, are one of the most researched types of MWUs.

Collocations have been identified as a component of vocabulary knowledge that poses particular difficulty for L2 learners (Laufer & Waldman, 2011), and researchers offer a variety of explanations for their learning burden (Peters, 2016). One explanation concerns the influence of the L1 of the learners, which has been shown to have a negative effect on the learning process of collocations that lack L1-L2 equivalents (i.e., incongruent collocations). Another reason for learners' difficulties with collocations is the general lack of exposure to them (Durrant & Schmitt, 2010; Pellicer-Sánchez, 2017; Meunier, 2012), particularly in English as a Foreign Language (EFL) settings, where opportunities for interaction with L1 speakers are limited (Webb, 2015), as well as the lack of attention paid to them in L2 textbooks (e.g., Brown, 2011). As a result, it is not surprising that collocation acquisition is a slow process (e.g., Altenberg & Granger, 2001; Bahns & Eldaw, 1993; Henriksen, 2013; Laufer & Waldman, 2011; Li & Schmitt, 2010).

Collocational competence is, on the other hand, a key aspect in the acquisition and development of L2 vocabulary and language proficiency (e.g., González-Fernández & Schmitt, 2015; Hill, 2000; Hsu & Chiu, 2008). According to Hill (2000:53), "collocations are found in up to 70% of everything we say, hear, read, or write". For Nation (2011), collocation knowledge is synonymous with language knowledge. Collocations are therefore critical to effective language use. Research indicates that deliberate learning activities such as using concordance lines assist learners in acquiring collocations (e.g., Macis, 2021). However, teaching the vast majority of collocations solely through intentional learning is impossible, especially in the EFL context, where the limited teaching time puts pressure on teaching priorities, which do not always fall into teaching vocabulary, let alone collocations. Therefore, collocations need to be learnt through incidental approaches.

1

Some researchers contend that collocation learning is enhanced through repeated encounters in context (Webb & Nation, 2017). An effective way to increase learners' repeated exposure to collocations is through reading, in its various modes (e.g., reading only, assisted reading (with audio support) or reading with textual enhancement). However, there are still relatively few studies investigating and/or comparing incidental collocation learning through different reading modes (e.g., Dang, Lu & Webb, 2022; Toomer & Elgort, 2019; Vilkaitė. 2017; Vu & Peters, 2021; Webb & Chang, 2020) and the results from these studies have been mixed. Webb and Chang (2020) found that reading-only (RO) contributed to greater gains in collocation than assisted reading or reading-while-listening (RWL). Vu and Peters (2021) corroborated these findings. In contrast, Dang et al. (2022) did not find an advantage of RWL over the RO mode. While the conflicting findings may be explained by differences in treatment variables between studies, there is still not a good understanding of the primary determinants of incidental collocation learning.

In addition, research on collocations show that learners' knowledge of collocations correlate with a range of factors to varying degrees. These factors include but are not limited to the learners' prior vocabulary knowledge, repetition, congruency, compositionality, corpus frequency, and association strength. While previous studies have addressed some of these issues (e.g., Dang et al., 2022; Vilkaitė, 2017; Vu & Peters, 2021; Webb & Chang, 2020), it still remains unclear the degree to which some of these variables moderate the effects of meaning-focused learning of collocations. Unlike earlier studies, which have largely focused on narrative texts (e.g., Pellicer-Sánchez, 2017; Vu & Peters, 2021, Webb et al., 2013; Webb & Chang, 2020), the reading materials to be utilised in this thesis will be expository texts of an academic kind, whose characteristics may have a differential effect on vocabulary learning. Previous studies indicate that distinct characteristics of the stimuli materials can have implications for collocation learning (e.g., Dang et al., 2022; Vilkaitė, 2017). Using academic texts as reading materials, Vilkaitė (2017) found lower collocation gains compared to similar studies using fictional stories. A recent study by Dang et al. (2022) used an academic lecture and reported lower gains through reading-while-listening than reading-only, which contradicts findings from studies using narrative stories as reading materials (e.g., Webb & Chang, 2020, Vu & Peters, 2022a). The difference in text types between these studies and studies using nonacademic input might have led to conflicting findings. Dang et al. (2022) argue that simultaneous presentation of aural and written input can be more disruptive than facilitative for texts of an academic kind. It is likely that text type as well as its level of difficulty (i.e., both in terms

of the difficulty of the language and its topic) can impact reading itself, whether in a reading-only or a reading-while-listening condition, which also affects how L2 readers address a given reading task. For example, L2 readers are more likely to reread parts of a text when the vocabulary load of a text and its information density is high. Nonfiction materials are likely to show such traits, as they often display a broader range of formal and academic language compared to narrative stories, which expose learners to more colloquial language and informal expressions. Since text-based factors (e.g., language style, word frequency, information density, sentence complexity) differ between narrative and academic texts, it worth examining whether academic texts (and their properties) can prompt collocational learning, and, more in particular, whether such learning can occur organically via activities that L2 university learners are already engaging with as part of their degrees. Further, the role of contextual support, which is regarded as an important factor in general vocabulary acquisition (Nation, 2013), will be investigated, as it has yet to be explored in collocation learning. Awareness of the features and demands of different learning conditions should provide additional insights about the type of input needed for the incidental uptake of collocations.

1.1. Aims of this thesis

The purpose of this thesis is to examine the factors that affect incidental collocation learning, as well as to address the research gap described above by investigating the effect of contextual support on collocation learning. The results of this thesis should help to guide material and course developers, teachers, and learners to facilitate learning by identifying factors that affect collocation learning through meaning-focused tasks.

By using pretest-posttest-delayed posttest mixed designs, the three experiments in the project sought to address the following broad research questions:

1. Can collocations be learnt incidentally from reading academic texts?

2. What factors (learner- and collocation-related factors, and features of the learning context) predict incidental collocation learning?

1.2. Structure of the thesis

This thesis reports on three experiments focusing on factors that are involved in the incidental learning of collocations from reading-only and reading-while-listening. It begins with a literature review in Chapter 2, which provides a general theoretical background for the scope of this thesis, starting from general theories of (second) language learning as well as those concerned with incidental vocabulary learning. Chapter 3 focuses on the phenomenon of multiword units in general and of collocations in particular, including its significance for L2 learning and vocabulary development. It also aims to clarify the concept of collocation and outline some of its features, as previous research has taken different approaches to characterise collocations.

The remaining chapters (Chapters 4, 5, 6) report on the three self-standing experiments. Each of the experiments is concerned with the effects of various factors on incidental collocation learning, which vary according to the specific goals of the experiment. Experiment 1 (Chapter 4) examines the effect of two reading modes (reading-while-listening vs reading-only) on incidental collocation learning through the reading of a single text. Experiment 2 (Chapter 5) seeks to determine the effect of increased repetition through the reading of a series of short academic texts. Experiment 3 (Chapter 6) examines the effect of contextual informativity on gaining knowledge of collocations through reading various short texts. Specifically, it compares the effect of three levels of contextual informativity, i.e., low, medium, and high. Table 1 summarises the three experiments.

Table 1. Summary of the three experiments.

	Experiment 1	Experiment 2	Experiment 3
Focus	Reading mode	Repetition	Contextual support
Input mode	RO vs RWL	Reading-only	
Additional factors included in analysis	▪ Vocabulary knowledge ▪ Congruency ▪ Compositionality ▪ Corpus frequency	▪ Vocabulary knowledge ▪ Congruency ▪ Compositionality ▪ Corpus frequency ▪ Association strength	▪ Vocabulary knowledge ▪ Congruency ▪ Corpus frequency ▪ Association strength
Research design	Pretest-posttest-delayed posttest mixed design	Pretest-posttest-delayed posttest, Latin-square design	
EFL participants	68 L1-Spanish upper-intermediate/advanced learners	84 L1-Spanish upper-intermediate/advanced learners	94 L1-Spanish upper-intermediate/advanced learners
Instruments	▪ Vocabulary (pre)tests ▪ Form recall ▪ Form recognition ▪ Meaning recognition ▪ Reading comprehension ▪ Retrospective questionnaire	▪ Vocabulary (pre)tests ▪ Form recall ▪ Form recognition ▪ Reading comprehension ▪ Retrospective questionnaire	▪ Vocabulary (pre)tests ▪ Form recall ▪ Meaning recall ▪ Reading comprehension ▪ Retrospective questionnaire
Data analysis	Mixed-effects regression models		

Finally, chapter 7 summarises the main results of each experiment and provides a general discussion of the findings. This final chapter also includes broader pedagogical implications of incidental collocation learning and recommendations for future research.

2. Understanding vocabulary

This chapter's goal is to lay the groundwork for the current PhD thesis. It begins with a brief historical overview of theories of first (L1) and second (L2) language acquisition, followed by sections devoted to understanding vocabulary knowledge and how it is conceptualised in the field of second language acquisition (SLA). Special emphasis is placed on how vocabulary is learnt under incidental learning conditions. The chapter concludes with an overview of the conceptual frameworks for incidental vocabulary learning, which are of particular interest to this thesis.

2.1. (Second) language learning theories

How language is acquired remains one of the central questions in human learning. Several L1 acquisition theories have been proposed and revisited over the years (Lightbown & Spada, 2013). This section presents an overview of these theories, which have influenced theories of SLA.

Early L1 acquisition research followed mainstream linguistic and psychological theories (Gass, Behney, & Plonsky, 2013). In the 1950s and 1960s, the behaviourist theory explained how people learn to speak by stressing the environment and the learner's observable experience with it. Language learning was simply the acquisition of a new behaviour (Skinner, 1957). Children develop language competence by imitation, practice (e.g., responding to parental or carer speech), and receiving feedback on success (de Bot, Lowie & Verspoor 2005). Skinner (1957) described language acquisition as the process of developing good language habits through conditioning (Mitchell, Myles, & Marsden, 2013). The explanations offered by behaviourism influenced SLA researchers, who proposed that L2 learners learn in a similar way, with the exception that L2 learners already hold habits formed in acquiring the L1, which interact with L2 learning. Such interaction or transfer can be positive or negative (i.e., interference). Behaviourism was associated with language transfer research and the contrastive analysis (CA) hypothesis, which postulates that L2 errors can be traced back to the learners' L1.

With new findings from child language acquisition, beliefs began to switch in the 1960s and 1970s (VanPatten & Williams, 2015). Children produce utterances that are not present in the input, indicating that their knowledge of language extends beyond what they learn solely through exposure (Gass et al., 2013). For the first time, language learning was depicted as distinctive from other types of learning. CA

6

studies did not predict all errors in L2 research, so behaviourism began to lose ground as a result of its overemphasis on environmental factors and lack of empirical evidence (VanPatten & Williams, 2015). Chomsky's early review of Skinner's work (1959) concluded that behaviourism was either irrelevant or meaningless in human language acquisition, effectively ending behaviourism in mainstream acquisition research as he developed his Poverty of Stimulus arguments throughout the 1960s and 1970s.

The first language-specific model of L2 learning, the Monitor Model, was developed in the 1970s, and the field of SLA began to rely on its hypotheses and research about the model (Krashen, 1982). Krashen's model is made up of five hypotheses: The Acquisition-Learning Hypothesis, the Monitor Hypothesis, the Natural Order Hypothesis, the Input Hypothesis, and the Affective Filter Hypothesis. The distinction between acquisition and learning is the most influential of the five (Gass et al., 2013). Krashen defines acquisition and learning as two distinct ways of developing linguistic skills in an L2. Acquisition is defined as a natural, unconscious learning process, whereas learning is defined as explicit and conscious. Such distinction sparked the explicit/implicit learning debate. According to Krashen, the source of spontaneous communication is implicit knowledge (i.e., acquisition), while explicit knowledge is of little use as it cannot be converted into implicit knowledge.

In the 1980s, nativist approaches, based on arguments developed by Chomsky since the 1950s (Chomsky, 1959), started to gain popularity among SLA researchers, and the Universal Grammar (UG) theory became increasingly influential over the years (Mitchell et al., 2010). UG postulates that children are born with an innate ability to learn language. According to Chomsky, the language faculty contains innate knowledge of linguistic rules, constraints, and principles, which constitutes the language faculty's initial state (Gass et al., 2013). Therefore, Chomskyan linguistics contends that Skinner's conditioning cannot account for the ultimate language attainment because children acquire knowledge for which no evidence exists in the input (VanPatten & Williams, 2015). This is known as the Poverty of Stimulus argument. Other traditional UG claims include the idea that there is a sensitive (critical) period for language learning (Lennenberg, 1967), the existence of language universals, i.e., all natural languages possess certain properties (Chomsky, 1986), and the fact that most children acquire language: deaf children learn to sign, and children with limited cognitive abilities build complex language systems (VanPatten & Williams, 2015). The Critical Period Hypothesis (CPH) (Lenneberg, 1967) contends that there are maturational constraints on the time an L1 or an L2 can be fully attained due to age-related

decline in neural plasticity, and has been used to support UG-based explanations of increasing difficulty in L2 learning with age (Penfield & Roberts, 1959).

In SLA, UG is a property theory which tries to explain the nature and acquisition of the evolving language or the "interlanguage competence" (Selinker, 1972) of L2 learners. The term interlanguage is used to define the interim linguistic system of L2 learners, which is believed to be a separate, unique system produced by L2 learners in their process of learning. There are three different views on the availability of UG principles in the L2 context: no access, partial access, and full access to UG (Mitchell et al., 2010). Some SLA research has shown that the interlanguage competence of L2 learners goes beyond L2 input, suggesting that UG is also operating in L2 acquisition (VanPatten & Williams, 2015). However, there are a few objections to innatism in SLA, because the learning mechanisms underlying L2 acquisition are different from those underlying L1 acquisition. First of all, there is no possible language transfer in learning an L1 but L2 is subjected to transfer, be it positive or negative, i.e., L2 learners have already acquired a language (Richards & Schmidt, 2013). L1 acquisition is a natural, unconscious learning process whereas adult L2 learning is conscious and effortful. Input also differs with regards to type, i.e., L1 input is mostly oral (at the beginning), quantity, i.e., children acquiring their L1 receive more input than L2 learners, and quality, i.e., children acquiring their L1 receive input mostly from other native speakers, while L2 learners may only receive input from other L2 learners. Further, the cognitive abilities of children cannot be compared to those of adult learners. Overall, there is evidence for and against UG principles in SLA (Gass et al., 2013), although the UG approach is appealing to SLA researchers because it provides a detailed descriptive framework which enables researchers to formulate hypotheses about the task facing the learner, and to analyse learners' language in a more focused manner (Mitchell et al., 2013).

Usage-based approaches or cognitive perspectives to language acquisition were proposed in the 1990s and 2000s (e.g., Goldberg, 1995; Tomasello, 2000; Ellis, 2005). Usage-based approaches are closely aligned with Cognitive-Functional linguistics (e.g., Bybee, 1995; Croft, 2000; Langacker, 1987), and with frequency-based (Ellis, 2002), connectionist (Elman, 2005), and emergentist theories (MacWhinney, 1999; O'Grady, 2005, 2008). The fundamental assumption of usage-based approaches is that language acquisition emerges from language use (Ellis, 2002). Two working hypotheses reconcile these approaches: 1) language acquisition is based on learners' exposure to L2 input in use, and 2) learners infer L2 rules by employing general cognitive processes that mediate any kind of

learning, including language learning (Ellis & Wulff, 2015). From a usage-based perspective, L1 and L2 acquisition draw on the same domain-general processes of perception, memory, categorisation, and generalisation. In essence, language acquisition depends on the child's general learning abilities and the contributions of the environment. While usage-based approaches reject the idea of domain-specific language acquisition mechanisms like UG, they do not propose an alternative property theory[2]. They also assume the existence of linguistic features in the lexicon but they do not explain where these come from. The focus is thus on the input and the processing of the input, and how domain-general learning mechanisms shape L2 learning. Those mechanisms, in theory, operate without awareness (implicitly), and they do not necessitate explicit knowledge of the language (Mitchell et al., 2013).

Other usage-based theories to language acquisiton include statistical techniques and models to analyse and process large amounts of linguistic data. From this standpoint, statistical learning accounts of language acquisition. Evidence from the statistical learning literature suggests that chid utterances are stored and integrated in chunks (Lany & Saffran, 2010). Children repeat regularly occurring words that are non-idiomatic more accurately than when the same words form lower frequency linguistic units (Bannard & Matthews, 2008). Statistical approaches to language learning are of great relevance in identifying patterns and regularities in linguistic data through statistical analysis, but they are not without pitfalls as they prioritise descriptive accuracy over explanatory power. Further, some statistical models heavily rely on frequency-based information.

The sociocultural theory (SCT) of human development is another influential theory that can be traced back to 18th and 19th century German philosophy (Lantolf, Thorne, & Poehner, 2015). Based on Vygotsky's theory of cognitive development (1986), the sociocultural orientation to learning takes an emergentist, usage-based view of how language is constructed through meaningful interaction (Lantolf & Thorne, 2006). Learning is a socially mediated process that involves oral interaction, regulation, and internalisation (Lantolf et al., 2015). Thus, language is a symbolic tool that is first social, then individual (Lightbown & Spada, 2003). Learning occurs ideally in the Zone of Proximal Development (ZPD), which is achieved by guiding children/L2 learners through successive steps (scaffolding) and engaging the

---

[2] In theory construction, a property theory is a theory that addresses the nature of L2 acquisition (Gregg, 2001). Property theories attempt to uncover the underlying mechanisms and principles that govern the process of acquiring a language.

child/learner in meaningful activities that are beyond their current level of competence (Donato, 1994). Sociocultural theories are concerned with the process of learning the L2 system rather than how learners construct it. Educators like this approach because it focuses on the individual learner. A major flaw of the SCT is that it lacks a property theory and it fails to integrate social and psychological perspectives (VanPatten & Williams, 2015).

This section has examined L1 and L2 language acquisition theories. The following section delves deeper into the conceptualisation and acquisition of vocabulary knowledge in the field of SLA, which is the focus of this thesis.

2.2. What is vocabulary knowledge?

Language emerges as words first, both in terms of how we learn to speak our mother tongue and any second or additional languages. Vocabulary knowledge plays thus a critical role in L2 learning and teaching. Even learners are aware of its relevance as they "carry around dictionaries and not grammar books" (Schmitt, 2010:4). In the last two decades, the importance of vocabulary knowledge has moved from being a peripheral aspect of language learning into a central position (Webb & Nation, 2017). Prior to investigating how vocabulary is acquired, we must look first at how it is operationalised. This section presents the main constructs in the vocabulary literature and how they have been applied to vocabulary research.

2.2.1. What's in a word?

Discussions of vocabulary acquisition often begin by referring to it as one of the most "intriguing puzzles" in SLA (Schmitt, 1998). Vocabulary knowledge is a multidimensional and complex construct that is acquired through a variety of learning processes (Read, 2000). To describe such learning processes, linguists, SLA researchers, psychologists, and neurobiologists must work together (Chacon-Beltran, Abello-Contesse, & Torreblanca-Lopez, 2010). To date, there is no widely accepted theory of vocabulary acquisition, though different theories for specific lexical processes have been proposed (e.g., Carey, 1978; de Bot, Paribakht, & Wesche, 1997).

There is converging evidence in the literature that vocabulary acquisition is incremental in nature (Schmitt, 2010). The gradual process of word learning is based on the idea that different aspects of a

word are mastered at different rates (Schmitt, 2008). Word acquisition is defined in the literature as knowledge of the form-meaning link both in and out of context. A word is acquired when learners can identify its meaning and use it appropriately and naturally in oral/written communication (Nation, 2001; Richards, 1976; Schmitt, 2000). As a result, different aspects of word knowledge must be considered. Word knowledge is classified into three categories: form, meaning, and use. Nation's (2001) framework of word knowledge aspects is the nearest reference frame we have to describe what it means to know a word. His framework characterises word knowledge in terms of nine word knowledge components: form (spoken and written forms, word parts), meaning (form and meaning, concepts and references, associations), and use (grammatical functions and collocations), each of which is divided into receptive [R] and productive [P] mastery (see Table 2). Learning a word implies thus knowing a range of word knowledge aspects, which are thought to be acquired on a continuum and at varying rates (Nation, 2001; Schmitt, 2010). For learners to master a word, it must be encountered several times in various modes, contexts, and registers (Nagy, Anderson, & Herman, 1987; Nagy & Scott, 2000; Nation, 2001; Schmitt, 1998, 2010). As a result, as learners gain more vocabulary exposure, they may encounter more aspects of word knowledge (Schmitt, 2010). To put it another way, the more types of word knowledge a learner has, the more likely he or she will be able to use it appropriately in the right contexts. This demonstrates that lexical items are not isolated units in the language but rather belong to a variety of related systems and levels (Nation, 2001).

Table 2. Nation's (2013: 49) framework of the components involved in knowing a word.

FORM	Spoken	[R]	What does the word sound like?
		[P]	How is the word pronounced?
	Written	[R]	What does the word look like?
		[P]	How is the word written and spelled?
	Word parts	[R]	What parts are recognisable in this word?
		[P]	What word parts are needed to express the meaning?
MEANING	Form and meaning	[R]	What meaning does this word form signal?
		[P]	What word form can be used to express this meaning?
	Concept and referents	[R]	What is included in the concept?
		[P]	What items can the concept refer to?
	Associations	[R]	What other words does this make us think of?
		[P]	What other words could we use instead of this one?
USE	Grammatical functions	[R]	In what patterns does the word occur?
		[P]	In what patterns must we use this word?
	Collocations	[R]	What words or types of words occur with this one?
		[P]	What words or types of words must we use with this one?
	Constraints on use	[R]	Where, when and how often would we expect to meet this word?
		[P]	Where, when, and how often can we use this word?

Within this framework, the most researched component of vocabulary knowledge is the form-meaning link, which will be examined in this thesis. Some authors argue that the relationship between a word's form and its meaning is the first of the word knowledge aspects to be learnt, and the most important for communication (Agustín Llach & Moreno Espinosa, 2014; Laufer & Goldstein, 2004). As a result, the majority of vocabulary study measures assess productive and/or receptive knowledge of the form-meaning link (Laufer & Goldstein, 2004). Laufer and Goldstein (2004) classify the form-meaning aspect into four levels of difficulty (or strength of knowledge): active recall (the most difficult), passive recall, active recognition, and passive recognition (the easiest) (see Section 2.2.4). These degrees of knowledge are important for comprehending the various levels of declarative vocabulary knowledge as well as making sense of the learning gains reported in vocabulary tests (Read, 2000).

2.2.3. The learning burden of a word

Another relevant construct from L2 vocabulary literature is the concept of a word's "learning burden" (Nation, 1990). The learning burden of a word is the amount of effort required to learn and remember it. This "burden" is determined by three factors: 1) the learners' prior knowledge of the target language and their mother tongue, 2) how the word is learnt and taught, and 3) the inherent difficulty of

the word. This essentially means that intrinsic and extrinsic lexical features, as well as non-linguistic factors, influence the development of L2 vocabulary.

A number of studies have shown that the relative similarity/dissimilarity between the learners' L1 and the target L2 also influence vocabulary learning (e.g., Jiang, 2002; Lotto & De Groot, 1998; Schmitt, 2010). The phenomenon of cognate words is an example of L1 lexical influence (Schmitt, 2010). In terms of how a lexical item is learnt or taught, Nation (2001) contends that a number of factors can interfere with the learning process. The concept of "unteaching" (Nation, 2001) is a good example of the impact of how a word is taught. Unteaching (unhelpful teaching) is the practise of not organising the way vocabulary knowledge is taught, thereby increasing the learning burden. An example of unteaching would be to teach words with similar forms at the same time (e.g., affect-effect) or exceptions before a rule has been fully understood. As Papagno, Valentine, and Baddeley (1991) and Papagno and Vallar (1992) demonstrate, intrinsic properties of words such as phonological distinctiveness, word length, part of speech, and pronunciation difficulties also influence the acquisition of new L2 words. Further, De Groot and Keijzer (2000) suggest that form-meaning mappings are easier to learn and less prone to forgetting if they are concrete (vs abstract) words, cognate (vs non-cognate) words, loanwords (for unrelated languages), and high frequency (vs low frequency) words (weak effect). Idiomaticity is also seen as a difficulty-inducing factor (Schmitt, 2010).

2.2.3. Breadth/depth of knowledge

Another influential construct in L2 vocabulary knowledge is that of size/breadth vs depth/quality of knowledge. Anderson and Freebody (1981: 93) defined vocabulary breadth as the number of words a language user knows in terms of relevant aspects of meaning whereas vocabulary depth is "the quality or depth of understanding". The term vocabulary size is also used (and preferred) instead of breadth because it refers to quantitative aspects, i.e., how many words are known (Nation, 1990). Depth, on the other hand, is concerned with qualitative aspects. Suffixes, word associations, collocations, inflectional affixes, spelling, grammatical information, meaning knowledge beyond the most frequent, dictionary-based meaning, and synonyms are examples of depth of word knowledge. In simple terms, depth refers to how well a language user understands a word or the knowledge of a given word at various levels (Nation, 2001).

Breadth has been defined and operationalised as knowledge of the form-meaning link of multiple words (for example, word recognition in Yes/No tests or meaning-matching of a given L2 form or vice versa). Breadth has been extensively researched, and a variety of standardised tests have been used to estimate vocabulary size. The Vocabulary Size Test (VST) (Nation & Beglar, 2007) and the Peabody Picture Vocabulary Test (PPVT) are two widely used receptive vocabulary size tests (Dunn & Dunn, 2007). The former is a multiple-choice test divided into frequency bands in which learners must recognise the meaning of a given word form, and the latter presents various pictures for learners to choose the one that represents the word they hear. Aside from these two tests, which were specifically designed to assess learners' vocabulary size, researchers contend that the majority of traditional vocabulary tests currently available (e.g., L1-L2 translation tests, recognition of word meaning, or recall of word form) also assess vocabulary size because they typically assess form-meaning link knowledge.

In contrast to breadth, a precise definition of vocabulary depth has proven elusive (Schmitt, 2014) and so it remains a construct difficult to operationalise (Nation, 1990). As a result, depth has received less attention than breadth in the L2 vocabulary acquisition literature (Henriksen, 1999; Nation & Webb, 2011; Schmitt, 2010). Read (1993: 357) defines depth as "the quality of the learner's vocabulary knowledge", which translates into knowledge of synonyms and collocations. Others define it as the learner's understanding of the relationships between individual words, more specifically knowledge of multiple semantic associations with other lexical items in the mental lexicon (Haastrup & Henriksen, 2000; Meara, 1996). The various definitions point to a richer understanding of individual words and their connections (Anderson & Freebody, 1981; Beck, McKeown, & Kucan, 2013; Qian & Schedl, 2004). Testing depth entails a thorough examination of individual words and their various aspects of word knowledge. Although the Vocabulary Knowledge Scale (VKS) has been used in some studies to assess depth, researchers agree that administering a battery of tests is preferable to measure different aspects of word knowledge (e.g., Schmitt & Meara,1997; Pigada & Schmitt, 2006; Webb, 2009). As a result, multiple tests are required for an accurate and reliable depth assessment.

In sum, the breadth/depth dichotomy is a useful descriptive framework to describe vocabulary knowledge and its dimensions. The concept of depth may yield helpful pedagogical insights by emphasising that word knowledge is more than knowledge of the form-meaning link (Schmitt & Schmitt, 2020). In terms of acquisition, it should be noted that breadth and depth do not necessarily develop in parallel (Schmitt, 2014). It is possible to know very little about a large number of words, or to know a lot

about a small amount of words (ibid). An example would be an L2 speaker using their L2 in a very constrained work situation, e.g., a taxi driver.

2.2.4. Receptive/productive distinction

The distinction between receptive and productive knowledge (sometimes referred to as passive and active knowledge) has been widely taken up in the conceptualisation of vocabulary (Meara 1990; qtd in Nation, 2001; Schmitt & Schmitt, 2014). The terms receptive/productive describe what language learners can do with words. Receptive knowledge refers to words that can be recognised and understood while listening or reading, whereas productive knowledge refers to words that learners can recognise, comprehend, recall, and use correctly in oral and written discourse (Webb, 2008). A learner's receptive/productive knowledge of a lexical item is frequently measured in tests using recognition/recall formats. Recognition is a receptive measure in which learners recognise the form or meaning of a target word (Read, 2000). In recall tests, learners produce the target word from memory in response to a stimulus (ibid). Schmitt (2010) developed a framework for defining receptive (recognition) and recall (productive) knowledge based on how it is measured. As shown in Table 3, Schmitt (2010) divides the form-meaning link of a word into four types of knowledge: form recall, form recognition, meaning recall, and meaning recognition (in Laufer and Goldstein's (2004) terms: active recall, active recognition, passive recall, and passive recognition). Schmitt's (2010) terminology will be used in the experiments of this thesis.

Table 3. Framework for defining recognition and recall knowledge (from Schmitt, 2010: 86).

		Word knowledge tested	
		Recall	Recognition
Word knowledge given	Form	Form recall (supply the L2 item)	Form recognition (select the L2 item)
	Meaning	Meaning recall (supply definition/L1 translation)	Meaning recognition (select definition/L1 translation)

The distinction between a learner's receptive and productive vocabulary sizes is usually referred to as receptive/productive gap (e.g., Laufer, 1998; Schmitt & Meara, 1997; Webb, 2008). The transition from no knowledge to receptive and productive vocabulary has been frequently addressed (Schmitt, 2010). Past research has focused on whether receptive knowledge or productive knowledge is acquired first, primarily in terms of vocabulary size (Aitchison, 2003; Laufer, 1998; Laufer & Paribakht, 1998; Melka, 1997). The evidence available shows that receptive and productive knowledge are learnt incrementally, with receptive mastery being typically reached before productive mastery, in part because productive mastery requires knowledge of more word knowledge aspects (Henriksen, 1999; Nation, 2013). Therefore, research into the receptive/productive gap shows that learners' receptive vocabulary knowledge is larger than their productive one (e.g., Laufer, 1998; Mondria & Wiersma, 2004; Webb, 2008; Schmitt, 2010). Though it is widely accepted that passive knowledge becomes active in an incremental fashion (Meara, 1997), there is still considerable uncertainty with regard to the threshold at which receptive knowledge becomes productive (Laufer & Goldstein, 2004; Schmitt, 2010). The receptive/productive gap poses a real challenge for learners, particularly in the EFL context, where the transition from receptive to productive vocabulary knowledge is also difficult to track (Laufer & Nation, 1995). The next section zooms in on three prominent conceptual approaches to understanding vocabulary knowledge.

2.3. Approaches to vocabulary knowledge

Vocabulary research has emphasised the existence of different degrees and types of word knowledge involved in learning vocabulary, leading to the development of various approaches to vocabulary knowledge. There are three main approaches to vocabulary knowledge: the components approach, the developmental approach, and the lexical networks approach.

2.3.1. Components approach

The components approach (or the dimensional approach) divides vocabulary knowledge into its various components and describes the differences between them (Milton & Fitzpatrick, 2014). It seeks to assess the quality of vocabulary knowledge, i.e., how many different word knowledge aspects learners are aware of (Schmitt, 2010).

Cronbach (1942) proposed one of the earliest descriptions of the multidimensional nature of word knowledge. He identified five dimensions of vocabulary knowledge: generalisation (ability to define a word), application (ability to use a word), breadth of meaning (ability to know the various meanings of a word), precision of meaning (ability to use the various meanings appropriately in different contexts), and availability (ability to use a word productively). Richards (1976) developed a more comprehensive list of what it means to know a word, which included eight assumptions dealing with features such as pronunciation, register, and morphological information. Nation (1990) expanded Richards' (1976) work by introducing new aspects as well as receptive and productive knowledge of each aspect. Nation's (2013) model is currently regarded as the most comprehensive approach to vocabulary dimensions (Webb, 2019). The components approach has some benefits, including its comprehensiveness due to measuring multiple aspects of word knowledge (Schmitt, 1998). From the researcher's perspective, this approach aids in breaking down lexical acquisition into more manageable units (Schmitt, 2010). Nation's model, however, has limitations. To begin with, measuring all word knowledge features is impractical, so some researchers have advocated for a simplification of Richards' and Nation's lists (e.g., Chapelle, 1998; Qian, 2002; Henriksen, 1999). According to Schmitt (2010), the overall research trend is to concentrate on a small number of components. Another criticism levelled at the use of Nation's model is the fact that it does not specify the relationships between the components. Vocabulary knowledge is not a feature of single words, but whole lexical networks (Meara, 1997; Meara & Wolter, 2004; Schmitt, 2014).

2.3.2. Developmental approach

The developmental approach reflects the incremental nature of vocabulary acquisition, so it makes use of developmental scales to measure vocabulary knowledge (Schmitt, 2010). According to this approach, word knowledge is acquired in a series of different learning stages that represent each word knowledge aspect, thus it can be organised according to developmental scales ranging from lower degrees of knowledge to higher degrees. This approach characterises aspects of word knowledge by examining different stages of word acquisition.

Scales have been widely used in the field of L2 vocabulary research and are based on the assumption that the development of certain aspects of vocabulary knowledge is gradual (Read, 2000;

Schmitt, 2010). One influential elicitation tool is the VKS[3] (Wesche & Paribakht, 1996), which describes word mastery as following a continuum from not knowing anything about a word to full mastery, characterised by the ability to correctly use the word across contexts. This scale has a number of limitations[4]. Because previous research has not reached a consensus on a principled developmental scale, the operationalisation of the developmental process remains unfeasible (Schmitt, 2010). The main limitation of the scales is that they rely on self-reported data or because testees do not to provide an accurate self-assessment. Another shortcoming of this approach is that there is still little theoretical guidance as to how many stages an acquisition scale should include (Schmitt, 2010).

2.3.3. Lexical networks approach

The third general approach to conceptualising vocabulary knowledge is the lexical networks approach. It envisages vocabulary knowledge as the lexical network that makes the L2 learners' mental lexicon. The different models within this approach are based on the observation that words do not exist in isolation but instead entangle and interconnect with other words to form a network (e.g., Crossley, Salsbury, & McNamara, 2010; Meara, 2009). Within the lexical approach we can find Aitchison's (2003) "web" of words or Meara and Wolter's (2004) interrelated network model, which represents depth of knowledge and the connections between words as a web in which greater linkage is associated with better connectedness and organisation (see

Figure *1*). These models posit that words will only be mastered productively if they develop appropriate links and networks in the mental lexicon (Meara, 1997). Overall, the lexical network framework is considered more theoretical than practical for L2 vocabulary research, although previous studies on this approach set out to produce measurement tools (see Meara & Wolter, 2004; Wilks & Meara, 2007).

---

[3] Wesche and Paribakht (1996) adopted and expanded Dale's (1965) L1 developmental model to L2 acquisition.

[4] Read (2000) provides a critique of the VKS and its limitations.
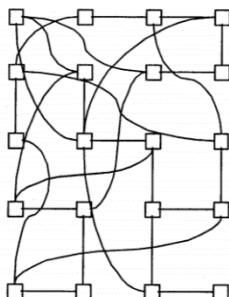
Figure 1. Vocabulary size and organisation, a network model of vocabulary (Meara & Wolter, 2004: 89).

An important point to mention is that the approaches described in this section have not been empirically tested, and thus, supported. As a result, the actual nature of vocabulary knowledge and how it develops is not yet clear, and needs to be further examined (Li & Kirby, 2015; Milton & Fitzpatrick, 2014).

The experiments reported in this thesis use incidental learning conditions, in which learners acquire lexical items without a conscious intention to learn those specific items while engaging in meaning-focused activities such as reading. The remainder of this chapter focuses on the learning conditions that can lead to vocabulary learning. It begins by looking at the concept of incidental vocabulary learning. Then, it presents the conceptual frameworks that can be used to support it.

2.4. Vocabulary learning conditions

The literature on vocabulary learning has consistently followed the tradition of two learning approaches: incidental learning and intentional learning (also known as implicit/unattended or explicit/attended learning) (Hulstijn, 2003). In applied linguistics, incidental learning is often defined as the acquisition of a linguistic feature while the primary focus is on another (Schmidt, 1994). Hulstijn (2001) describes it as the by-product of any meaning-focused activity such as reading a novel, listening to the radio, or watching a movie. In contrast, intentional learning refers to any activity that requires a conscious effort to deliberately learn new words. In an intentional learning condition, learners are instructed to engage with specific features in the input (e.g., spelling, form-meaning connection) and pay deliberate attention to them. The distinction between incidental and intentional learning may be also based on the use of explicit instructions that forewarn learners about an upcoming vocabulary test (Mäntylä, 2001; Craik, 1972). Although they are often conceptualised as opposites, incidental and

intentional learning complement each other (Schmitt, 2010), and are beneficial for L2 lexical growth (Webb, 2020). Some researchers contend that, though useful, narrow definitions of incidental and intentional learning do not provide a clear-cut distinction (Hulstijn, 2001), "particularly if we accept that most learning involves conscious attention" (Nation, 2001: 349). Therefore, a more practical approach is to describe the conditions for learning, thus framing the distinction between incidental vs intentional learning conditions (e.g., Nation & Webb, 2011; Pellicer-Sánchez & Boers, 2019).

2.4.1. Incidental vocabulary learning

Most L2 vocabulary, except for the first few thousand words, is learnt incidentally (Rieder, 2003). Meaning-focused tasks such as reading a novel or listening to the radio can promote incidental word learning. Research shows that the gains made under incidental learning conditions are relatively small compared to the amount of learning reported in intentional learning studies (Hulstjin, 1992; Webb & Nation, 2017). The relatively low learning rates should be a benefit rather than a cause of concern. Incidental learning occurs in small increments as learners gain knowledge of new vocabulary through repeated and meaningful encounters (Webb, 2020). As previously seen in this chapter, the degree to which a word is known includes different aspects of word knowledge: written form, form-meaning link, collocation, etc. Incidental learning is thus likely to provide new instantiations of a new or a partially known word, which helps developing vocabulary depth (Schmitt, 2008). Some lexical and semantic information may be better learnt through incidental learning conditions (Nation, 2001). Knowledge of collocations and register, for example, is likely to be reinforced by the varied encounters embedded within different contexts (Nation, 2001; Schmitt, 2010).

The value of incidental learning in word acquisition is well established. There is plenty of evidence that L2 vocabulary can be learnt through (extensive) reading (e.g., Pellicer-Sánchez & Schmitt, 2010; Pigada & Schmitt, 2006; Vu & Peters, 2022a; Waring & Takaki, 2003; Webb, 2007), listening (e.g., Jin & Webb, 2020; Van Zeeland & Schmitt, 2013), and watching TV (e.g., Peters & Webb, 2018; Rodgers & Webb, 2020). So, how does incidental vocabulary learning take place? The next section presents various theoretical frameworks for incidental vocabulary learning.

2.5. Conceptual frameworks for incidental vocabulary learning

The mechanisms underlying incidental vocabulary learning have been explained using a variety of theoretical frameworks. In the 1970s, Craik and Lockhart (Craik & Lockhart, 1972; Craik & Tulving, 1975) introduced the Depth of Processing Hypothesis. According to this theory, the amount of learning is dependent on the relative amount of cognitive effort (or processing levels) which can range from shallow to deep, at which information is processed. In the revised version, Lockhart and Craik (1990) highlight that effective learning is triggered by two stages: an initial sensory analysis of the properties of the target word (e.g., orthographic and phonological features) and its subsequent retrieval and consolidation of its semantic associations and conceptual referents, which requires a deeper level of analysis (Eckerth & Tavakoli, 2012). Put simply, the deeper the level of processing, the more powerful the memory traces and the easier it is to recall the input. One criticism levelled at this theory is that it does not provide a method for determining and evaluating different levels of processing (see Hulstijn, 2001).

One of the most influential theories in the field of SLA is Krashen's Input Hypothesis (1982, 1989). This theory posits that input exposure to the L2 is necessary and sufficient for SLA. Krashen claims that extensive and sustained reading alone can trigger the acquisition of spelling and vocabulary knowledge (Cobb, 2007). The L2 input must be slightly beyond the current level of proficiency of the learners (i+1). What he calls "comprehensible input" is the driving force behind acquisition (VanPatten & Williams, 2015). Some researchers disagree with Krashen and maintain that wide reading alone is not sufficient for vocabulary learning (e.g., Nation, 2001; Sternberg, 1987). The benefit of reading for incidental lexical learning is, however, empirically proven (e.g., Nation & Waring, 2020; Macis, 2018; Pellicer-Sánchez & Schmitt, 2010).

Building on various psychological learning theories and his own empirical studies, Schmidt (1990, 2001) developed the Noticing Hypothesis on the basis that nothing is learnt unless is has been "noticed". This theory claims that acquisition begins when L2 learners become aware of a particular language feature in the input. Originally, Schmidt (1990:29) proposed that incidental learning occurs only when a task triggers learners' attention, i.e., noticing is "necessary and sufficient" for input to become intake. Schmidt (2001) later proved that noticing strongly affects second language learning, although he concedes that unattended learning may be also possible. The Noticing Hypothesis gained the support

of numerous SLA scholars (e.g., Ellis, 1994, 1997; Gass; 1988; Long, 1991; Skehan, 1998), but it also raised questions regarding the importance of awareness and attention.

Laufer and Hulstijn (2001) applied the conceptual framework proposed by Craik and Lockhart (1972) to incidental vocabulary learning in the Involvement Load Hypothesis (ILH). This theory proposes that the level of involvement load induced by a task determines its effectiveness in the process of vocabulary acquisition. The ILH includes a motivational-cognitive construct with one motivational aspect (*need*) and two cognitive information processing components (*search* and *evaluation*). According to Laufer and Hulstijn (2001), the motivation of the learner to determine the meaning of an unfamiliar word triggers the subsequent processes of searching for and evaluating the word's meaning. In other words, the more involvement, the more likely vocabulary will be retained. This validates Schmitt's (2008:329) claim that "the overriding principle for maximizing vocabulary learning is to increase the amount of engagement learners have with lexical items". However, the theory fails to account for factors such as time on task (e.g., Keating, 2008), frequency exposure (e.g., Folse, 2006), and student-related variables such as individual learning strategies (e.g., Schmitt, 2008).

Yanagisawa and Webb (2021a) recently revised the ILH, emphasising factors such as frequency and test format. They created a broader framework for incidental vocabulary learning by employing formulas that help calculate the relative effectiveness of a learning activity: the Involvement Load Hypothesis Plus. The formulas include seven components: need, search, evaluation, sentence-level varied use, composition-level varied use, frequency, and mode. However, the authors point out that this framework still does not address many other factors that can predict vocabulary gains, such as learner characteristics (e.g., proficiency), task-related features (e.g., time on task), or target word characteristics (Yanagisawa & Webb, 2021b).

All of the theories discussed in this section help us understand the various processes that occur during incidental vocabulary learning. They take into account various processes (e.g., noticing, involvement) and perspectives as a whole (e.g., cognitive, motivational). They have each contributed valuable insights and have been useful in L2 vocabulary research. Nevertheless, none of these theories fully capture the complexities of incidental vocabulary learning. Providing a single framework that encompasses all of the observed phenomena may prove challenging, given the range of factors that underpin the dynamic nature of vocabulary learning. A substantial amount of research shows that various interlexical factors (e.g., cognate status, L2 proficiency) and intralexical factors (e.g., spelling,

word length, morphological complexity, conceptual difficulty) affect vocabulary learning (Laufer, 1997; Nagy, McClure, & Mir, 1995). Other variables, such as learner characteristics (e.g., educational level, prior vocabulary knowledge, engagement), materials (text type, length), and learning activities (input mode), have only recently been included in incidental learning studies (see also Webb, Uchihara, & Yanagisawa, 2023). Collectively, these studies show that vocabulary learning is a complex and adaptive process that comprises the interactions of many different factors (Larsen-Freeman & Cameron, 2008). Other theories are likely to emerge, advancing our understanding of the processes underlying incidental vocabulary acquisition, but until then, some unanswered questions about the many factors that mediate incidental vocabulary learning call for more empirical research.

2.6. Summary

This chapter has established the framework for the current thesis. It has presented language acquisition theories that are relevant to our understanding of how vocabulary knowledge is acquired, as well as the various approaches to vocabulary knowledge that are available. It has also shown that vocabulary knowledge extends far beyond knowing a word's form and meaning and presented the ways in which vocabulary learning occur. Sources of meaning-focused input (such as reading or listening) are likely to contribute to incidental vocabulary learning. The following chapter will focus on vocabulary beyond single words, that is, multiword combinations, and in particular on one category: collocations, or words that tend to cluster together.

3. Incidental collocation learning

This chapter focuses on the incidental learning of multiword units (MWUs), specifically collocations. It begins with a discussion of the importance of MWUs and collocations in L2 learning. I then review the two main approaches to the identification of MWUs, with a special focus on collocations. Then, I discuss the reasons for encouraging incidental collocational learning and the challenges that adult L2 learners may face when learning collocations. Finally, I discuss a range of factors that influence collocation learning, some of which I examine in my experiments.

3.1. Defining Multiword Units

Evidence from psycholinguistics (Hoey, 2005; Wray, 2002), sociolinguistics (Coulmas, 1981), corpus linguistics (Moon, 1998a), and L1 and L2 acquisition (Pawley & Syder, 1983) shows that language users rely on recurring word combinations and have "preferred formulations" for expressing messages when communicating (Wray, 2006: 591). This is manifested in the widespread use of collocations ("heavy rain"), binomials ("black and white"), idioms ("break a leg"), and phrasal verbs ("drop off"), to name a few examples (Ellis, Simpson-Vlach, & Maynard, 2008). Various terms exist to denote such patterned use of language, the most common of which are lexical phrases (Nattinger & DeCarrico, 1992), lexical bundles (Biber & Conrad, 1999), multiword items (Moon, 1998b), phrasal lexemes (Moon, 1998b), formulaic sequences (Schmitt & Carter, 2004; Wray, 2002), or recurrent phrases (Hoover, 2002; Stubbs & Barth, 2003). Although conceptually related, the terms focus on distinctive features associated with specific word strings (Wray, 2002), and they differ depending on the field from which they are approached. As a result, the term formulaic language is used as the umbrella term to describe these recurring multiword sequences (Wray, 2002). One limitation of the term is that it is frequently associated with claims of mental storage, that is, the idea that formulaic sequences are more easily processed because they are stored holistically as prefabricated chunks (see Wray, 2002). Since I make no claims about the holistic storage of formulaic sequences in this thesis, I will use the term MWUs, which include collocations (Grant & Bauer, 2004; Nation, 2013; Nation & Webb, 2011).

It is likely that MWUs are a universal feature of languages (Buerki, 2020). Estimates of the number of MWUs used in productive language vary, but they constitute a significant part of the linguistic

repertoire of L1 and L2 users (Nattinger & DeCarrico, 1992). Conklin and Schmitt (2012) claim that MWUs account for at least one-third to one-half of all language, making them an important aspect of vocabulary learning and teaching. Erman and Warren (2000) calculated that a variety of prefabs (as they called them) accounted for 58.6% of the spoken English discourse and 52.3% of the written discourse they examined. Foster (2001) asked raters to look for MWUs in transcripts of unplanned oral native English, and the raters determined that MWUs made up 32.3% of the speech. This variance around the numbers can be most likely explained by differences in how researchers defined what constitutes a MWU. Comparisons of written and spoken corpora show that MWUs are more frequent in oral discourse (Biber, Johansson, Leech, Conrad, & Finegan, 1999), but they are also used in written language (ibid), where they provide cohesion and structure (Wood, 2020). Previous studies on MWUs in English academic writing show that L1 and L2 speakers use MWUs differently. Pérez-Llantada (2014) discovered, for example, that while proficient L2 speakers use a wide range of English MWUs, their usage is not fully native-like.

MWUs play a substantial role in promoting both receptive and productive fluency in an L2. MWUs help to ease the wording of complex ideas and maintain the flow of oral discourse (Wray, 2004; Wood, 2010; Nattinger & DeCarrico, 2000). A good command of MWUs is thus beneficial to perceptions of fluency and range of expression (Ellis, 2008), and is an indicator of native-like competence (Wray, 2002). Knowledge of MWUs is especially important for L2 learners living in the L2 environment because it aids in their social integration and assimilation in language communities (Foster, 2001). For example, when communicating in culturally specific situations, the use of situation-bound utterances (e.g., "bless you") is required (Kecskes, 2007). Listening skills are also influenced by knowledge of MWUs (Kremmel, Brunfaut, & Alderson, 2017). According to Bradford (2010), L2 learners must actively comprehend MWUs in order to follow various media such as films, television, radio, or a conversation with peers.

Taken together, previous research shows that learning MWUs is an important aspect of vocabulary knowledge for using an L2 fluently and accurately (e.g., Schmitt, 2004; Wray, 2002). The next section presents the two main approaches to the identification of MWUs.

3.2. Identification criteria

As stated in the preceding section, there is a wide range of MWUs. Many attempts have been made to differentiate between various types of MWUs and provide distinct classifications, which tend

to overlap and change over time. Taxonomies have been developed from a variety of perspectives, including linguistic (Granger & Paquot, 2008; Moon, 1998b; Cowie, 1981), pedagogical (Biber, 2009; Howarth, 1998; Nattinger & DeCarrico, 1992), and natural language processing (NLP) (Baldwin, 2006). There are two main approaches to identifying MWUs: form-based and meaning-based approaches (Nation, 2013).

3.2.1. Form-based approaches

Form-based approaches to the identification of MWUs are bottom-up, corpus-driven methods. They are also known as the frequency-based approach or distributional approach (e.g., Evert, 2004; Firth, 1962; Nesselhauf, 2004; Sinclair, 1987). This approach of MWU (and collocation) identification is statistically grounded as it identifies MWUs on the tendency for recurrent word strings to appear with high frequency in corpus searches (Henriksen, 2013). From this perspective, it is possible to extract and quantify MWUs using methods such as clusters, n-grams, collocations, or co-occurrence analysis (Stubbs, 2009). Researchers employing a form-based approach use strength-of-association measures, i.e., statistical measures that indicate the extent to which observed regular co-occurrence of words differs from what we would expect by chance (Webb et al., 2013). Section 3.5.4.4. will introduce such measures.

In this tradition, MWUs are viewed as combinations of linguistic items (including words, parts of speech, and semantic fields), which have a high probability of co-occurrence. Examples include high frequency collocations (e.g., "fast food"), colligations (e.g., the verb "to depend" colligates with the preposition "on"), semantic preferences (e.g., words related to express + feelings), and semantic prosody (e.g., negative concept + happen; positive concept + provide) (Durrant & Mathews-Aydınlı, 2011). It should be noted that frequency-data does not directly indicate how a MWU is stored in the mental lexicon. Instead, an assumption is made that certain distributional patterns imply storage as a whole unit, because a corpus is believed to capture the usage patterns of speakers in a way analogous to behavioural experiments. Therefore, corpus-based frequency evidence allows for the identification of, for example, salient word combinations that would go overlooked otherwise (Biber et al., 2004). One major disadvantage with this approach is that "it lumps together units that are linguistically quite different" (Granger, 2019: 231). Moreover, frequency-based approaches do not consider semantic

factors that affect the learnability of MWUs, such as semantic transparency of meaning (Moon, 1998c; Nesselhauf, 2003).

3.2.2. Meaning-based approaches

Meaning- or phraseological-based approaches are top-down methods that categorise MWUs according to semantic criteria (Benson et al., 1997; Cowie, 1998b; Gyllstad & Wolter, 2016; Henriksen, 2013; Howarth, 1998a). One criterion is compositionality (or transparency), or whether the meaning of a MWU can be predicted from the meaning of its parts. The other is substitutability or restriction, which refers to the commutability of the constituents, or whether the relations between words are restricted or variable to varying degrees. In this tradition, MWUs are placed on an opacity/transparency continuum with fixed and semantically opaque MWUs at one end (e.g., "spill the beans") and high transparency, variable MWUs at the other (e.g., "a big house") (Howarth, 1998b). Collocations, sometimes called restricted collocations (e.g., "perform a task"), are in the middle of the continuum. Different categories of MWUs have been proposed within the phraseological tradition (e.g., Cowie, 1981, 1988a, 1988b, 1994; Howarth, 1998). While the phraseological approach aids in the identification of MWUs that are psychologically salient and pedagogically valuable (Henriksen, 2013; Simpson-Vlach & Ellis, 2010), it is also limited in various ways. Firstly, transparency and restriction are not easily operationalised as they can be defined and/or quantified in multiple ways (Granger, 2019). Secondly, this approach can be rather subjective because judgement of the semantic relationship between words varies from person to person (e.g., Boers & Webb, 2018; Schmitt, 2010).

The form- and meaning-based approaches are used to identify, operationalise, and discuss qualitative distinctions between different types of MWUs. In practice, however, some types of MWUs are not so easily operationalised in one single category as they can partly overlap with others. For example, some collocations can have an idiomatic meaning (Wulff, 2008). A growing number of studies have started to integrate aspects of both form- and meaning-based traditions (e.g., Boone, De Wilde, & Eyckmans, 2023; Fioravanti, Senaldi, Lenci, & Siyanova-Chanturia, 2021; Szudarski & Carter, 2016). For example, Boone et al. (2023) investigated word collocability by means of both phraseology- and corpus-based criteria. By adopting a hybrid definition of collocations, they found that collocation learning is mediated by a range of corpus-based and word-related factors such as frequency, collocational association strength, and imageability. Fioravanti et al. (2021) used behavioural data and computational

indices to measure semantic compositionality of word combinations. They found evidence that L1 and L2 speakers perceive lexical fixedness and compositionality of collocations differently. In conclusion, a hybrid approach minimises the frequency- and phraseology-based shortcomings while providing a more thorough identification of MWUs.

The experiments presented in this thesis will thus employ a mixed-method approach. Form-based measures will be employed to identify target collocations, while the meaning-based approach will help identifying semantic factors that affect collocational learning.

3.3. Collocations

MWUs have been classified using a range of criteria, including structural, semantic, or syntactic properties (collocations, idioms, metaphors, proverbs, compounds, phrasal verbs), pragmatic function (lexical phrases, pragmatic formulas) and distribution in language corpora (lexical bundles, variable expressions) (e.g., Granger & Paquot, 2008; Grant & Bauer, 2004; Martinez & Schmitt, 2012; Nattinger & DeCarrico, 1992; Simpson-Vlach & Ellis, 2010; Wood, 2020). Collocations are one of the most researched subclasses of MWUs in L2 vocabulary learning studies, if not the most (González-Fernández & Schmitt, 2015). They have aroused a great deal of attention for various reasons. Firstly, they are linguistically interesting. A word's typical collocates provide information about the semantic cohesion of its elements (Durrant, 2008) which enables a word to be used in a certain lexical context. This link between collocation and meaning was first introduced by Firth (1951) and his oft-quoted definition of collocation as "the company words keep". Halliday (1966) extended and refined this idea by pointing out that a word's conventional collocates offer a profile that can distinguish it semantically from other words with analogous meanings. The near synonyms "powerful" and "strong" illustrate this point clearly as they can be differentiated by virtue of their collocations: strong/*powerful tea; *strong/powerful engine. Collocations are of special interest to L2 vocabulary acquisition researchers because it is an aspect of lexical knowledge which L2 learners struggle with, even at advanced levels of proficiency (Granger, 1998; Howarth, 1998; Nesselhauf, 2003, 2005).

Widely varying definitions of collocation have emerged since Firth (1951, 1957) developed the concept. Collocations have been defined as combinations of two or more lexemes in a grammatical construction (Cowie, 1981), or as Nesselhauf (2005:1) puts it, as "arbitrarily restricted lexeme combination". The definition of collocation has evolved, and the term can be broadly defined as the

repeated co-occurrence of words. Boers (2020) defines collocation as strong word associations, i.e., recurrent co-occurrences of (content) lexical items. In meaning-based approaches to MWUs, a collocation is understood as a word combination displaying various degrees of fixedness, whereas the form-based approach treats collocations as lexical units for which the co-occurrence probability of the component parts is quantifiable in statistical terms (Nesselhauf 2005). In addition, some researchers use the term in its narrow linguistic sense to refer to lexical units which hold a syntagmatic relationship among the words which co-occur (Brown, 2014, Wood, 2020), while others conceptualise collocational knowledge as an aspect of depth of vocabulary knowledge itself, because knowing a collocation involves knowledge of the single words' linguistic characteristics which constitute it (Nation, 2001; Qian, 1999, 2002). For the purposes of this thesis, collocations are characterised as the co-occurrence of two words displaying a high statistical strength-of-association (using association measures) which are also valued along semantic dimensions such as congruency and compositionality. It is important to note that there may be cases where an intervening article is needed for grammatical reasons or to modify the collocation's meaning (e.g., "make a call"). This definition combines key aspects from both the form- and meaning-based approaches to collocations.

This thesis will focus on two well-studied types of lexical collocations: adjective+noun and noun+noun (e.g., Dang et al., 2022; Durrant & Schmitt, 2010; Kasahara, 2011; Macis, Sonbul, & Alharbi, 2021; Peters, 2016; Siyanova & Schmitt, 2008; Webb & Kagimoto, 2009; Webb & Chang, 2020; Vu & Peters, 2022a). Lexical collocations refer to word combinations which usually consist of two content words (full lexical items) and possibly a grammatical word (e.g., "spicy food", "subliminal message") (Lewis, 2000). These can be distinguished from grammatical collocations (also called colligations, see Bartsch, 2004), which include at least one content word and at least one preposition or grammatical structure, such as "to + infinitive"" or "that-clause" (e.g., "under attack", "immune to"). Both grammatical and lexical collocations are a source of difficulty for L2 learners, although the challenges they present vary due to their properties. The variable and less predictable usage patterns of lexical collocations may be particularly challenging even for advanced learners, compared to grammatical collocations, whose usage patterns are syntactically restricted and therefore easier to master (Howarth, 1998b; Ackerman & Chen, 2013).

The remainder of this chapter will focus on aspects of incidental collocation learning. The following section considers the advantages of using incidental learning conditions to promote collocations/MWUs development, with a focus on the acquisition of collocations by adult L2 learners.

3.4. Incidental learning of collocations

The notions of incidental and intentional vocabulary learning were introduced in Chapter 2. Collocations have been traditionally taught through focus-on-form (intentional) activities such as translation exercises, contrastive analysis, dictionary lookups, and word cards (e.g., Boers, Lindstromberg, & Eyckmans, 2014; Chan & Liou, 2005; Laufer & Girasi, 2008). Perhaps the most typical activity used to encourage collocation retention is to provide L2 learners with two lists of lexical items and instruct them to match up the items of the lists together to form collocations (Brown, 1974). One limitation of these tasks is that collocations appear decontextualised, which may affect the degree to which they are learnt and retained (e.g., Boers et al., 2014; Zhang, 2017). Previous research indicates that collocations are not always addressed appropriately in L2 textbooks (e.g., Boers, Demecheleer, Deconinck, Stengers, & Eyckmans, 2017; Brown, 2011; Vu & Michel, 2021) and that MWUs in general are not representative of native speaker usage (Northbrook & Conklin, 2019). The vast number and variety of collocations are also unlikely to be learnt solely under intentional learning conditions. For these reasons, collocations must be taught using meaning-focused input.

Research shows that collocations can be learnt from incidental exposure (e.g., Dang et al., 2021; Pellicer-Sánchez, 2017; Vilkaitė, 2017; Vu & Peters, 2022a, Webb et al., 2013). One potential facilitating effect of incidental learning is that learners may already have some knowledge of high- and mid-frequency collocations, i.e., they may be familiar with one of the components that make up the collocations. Such familiarity could promote learning through meaning-focused input, especially if the collocation's immediate context is informative of its meaning (e.g., Toomer & Elgort, 2019; Webb & Chang, 2020). Most importantly, research shows that a substantial amount of vocabulary knowledge is acquired incidentally from context (Nation, 2001). L2 learners use lexical inferencing for guessing the meaning of MWUs, at least for idioms (Cooper, 1999, qtd in Schmitt & Schmitt, 2020). Thus, there can be ways to improve incidental collocation learning through context. An additional factor that has a positive learning effect is sound patterning. Collocation learning and recall may be aided by features

such as rhyme (e.g., "brain drain"), alliteration (e.g., "wonderful world"), and assonance (e.g., "special effects") (Pellicer-Sánchez & Boers, 2019).

Incidental learning typically results in small collocation gains (Pellicer-Sánchez, 2020). However, these small gains are valuable as they reflect the gradual/incremental nature of collocation acquisition through meaning-focused input (Webb & Chang, 2020). Learners can not only develop their knowledge of collocations through repeated exposure, but they can also establish knowledge of previously known words/expressions, gain awareness of how specific vocabulary is used, gain knowledge of different word components (e.g., multiple meanings, derivatives), and learn new concepts (Webb & Nation, 2017). Incidental vocabulary learning helps to develop other skills like reading, grammatical knowledge, and oral fluency (Nation, 2001). The next section discusses the difficulties that adult L2 learners have with collocational learning.

3.4.1. Adult learners' difficulties with collocations

The literature on the acquisition of collocations by adult learners contends that collocations present special difficulties for learners, even for advanced ones (e.g., Bahns & Eldaw, 1993; Farghal & Obeidat, 1995; Granger, 1998; Nesselhauf, 2003, 2005). Bahns and Eldaw (1993:108) claim that EFL learners' collocational knowledge lags behind other language skills and that their general vocabulary knowledge "far outstrips their knowledge of collocations". Some researchers believe that adult L2 learners fail to notice and remember collocations in the input because, unlike L1 speakers, they focus on single words, rather than on meaningful multiword items (Wray, 2002). For example, on encountering the collocation "boot camp", an adult learner will break it down into its parts "boot" and "camp", therefore failing to notice it as a whole meaningful unit. This non-holistic/analytical approach to collocations may make it difficult for adult learners to learn collocations, even if they have a strong command of general vocabulary knowledge (e.g., Peters, 2006; Wray, 2002). Nonetheless, they have been shown to make relatively rapid progress in some other areas of vocabulary knowledge, particularly written, as a result of this analytical approach in their learning (see Muñoz, 2006; Schmitt, Sonbul, Vilkaitė-Lozdiene, & Macis, 2019). Wray (2002) argues that the difficulties that adult learners have in attaining knowledge of collocations (and MWUs in general) stem from other sources, including social and cognitive factors. Adult L2 learners, especially those in EFL settings, underuse (pragmatic) MWUs/collocations (e.g.,

"sounds lovely", "great joy") because their attention is not focused on communication, or they are unfamiliar with the items which express what they want to communicate (Hill, 2000).

Language awareness can help to explain the difficulties with collocations. Adult learners find it difficult to identify and master collocations used by L1 speakers (Pawley & Syder, 1983; Wray, 2002; Schmitt, 2010), especially if the collocations are not salient, which partly explains their slow acquisition (Boers et al., 2017). According to Wray (2002), learners' general lack of awareness of MWUs impedes the processes of noticing and decoding collocations in the input, both of which are required for L2 input to become intake (Schmidt, 1990, 1995). Previous research has nonetheless cast doubt on Wray's (2002) theory. Durrant and Schmitt (2010), for example, found that adult L2 learners can retain knowledge of collocational associations between words from exposure, although this study did not measure learning. Durrant and Schmitt (2010) conclude that poor collocational development is caused by a lack of input exposure rather than the learners' approach to learning collocations. More research is thus needed to assess the effect of incidental learning approaches using "adequate" input on adult learners' acquisition of collocations.

Statistical learning approaches give a different explanation for the acquisition of collocations (and MWUs). It argues that collocations can be derived directly from statistical properties of natural language (McCauley & Christiansen, 2019a), allowing for the detection of collocations, and generalisation across their patterns (Isbilen, McCauley, & Christiansen, 2022). According to this perspective, adult learners pick up and recognise collocations using statistical patterns and regularities present in the L2 input, i.e., they process frequencies and associations between words to discern collocations and other linguistic patterns. This learning process, however, requires substantial exposure to L2 input on a regular basis, which is arguably problematic for adult learners, particularly in the EFL context. In such environments, learners will experience difficulties in acquiring low-frequency items (e.g., "awful mistake") compared to high-frequency collocations (e.g., "fatal mistake"). For example, an eye-tracking study by Sonbul (2015) provided evidence that adult learners are sensitive to collocation frequency. Despite the benefits of data-driven approaches to collocation profiling and grouping words into latent word classes, statistical learning does not integrate relevant factors such as semantics or language variability, i.e., collocations can vary between different dialects, registers, and contexts. For example, a collocation with a literal meaning (e.g., "rainbow colours") will most probably be learnt via statistical learning, whereas a collocation with a figurative meaning (e.g., "red tape") will require context-rich language exposure to be

learnt, which, again, is problematic for L2 mature learners. Crucially, what this shows that focusing solely on statistical data and adhere to linguistically blind categorisations of collocations without taking into account their particular properties may not be the best approach (Szudarski, 2023).

Many adult learners study English in a classroom setting. The characteristics of the EFL context may pose a problem in incorporating collocational knowledge into classroom pedagogy. Some issues are related to limited contact with the L2, insufficient output or limited opportunities to practise English outside of the classroom, incorrect input manipulation (e.g., using non-realistic language), or limited time dedicated to (vocabulary) learning (Siyanova-Chanturia & Webb, 2016). One frequently mentioned factor influencing collocational learning is the traditional single-word-centered approach to vocabulary teaching, which has prevailed over the otherwise prominent role of MWUs (e.g., Henriksen, 2013; Boers et al., 2014). Collocations/MWUs can be omitted from speech addressed to learners in instructional settings (Irujo, 1986). Earlier research also indicates that there is frequently a lack of authentic input in the L2 classroom, as well as insufficient exposure in EFL settings (Muñoz, 2008a), where classroom time may be limited to only a few hours per week, which is insufficient for vocabulary growth, let alone learning collocations (Siyanova-Chanturia & Webb, 2016). Further, standard classroom teaching materials may not present adequate practice with collocations (Lancker-Sidtis & Rallon, 2004). The chances of acquiring collocations will also inevitably depend on the learner's intrinsic learning processes and strategies and the different scenarios for collocation uptake as well as the activities through which a learner engages with the language (Webb, 2020).

3.5. Variables affecting collocation learning

This final section provides a detailed description of a variety of factors relevant to the current PhD thesis. Some of these variables have received considerable attention (for example, repetition), while others have yet to be thoroughly investigated (e.g., context informativity). Examining these factors will provide a more fine-grained picture of the factors that contribute to incidental collocational learning.

3.5.1. Features of the learning context

3.5.1.1. Input mode

Input plays a central role in learning a FL/L2 (Ellis & Shintani, 2013; Gass, 2017). It is crucial for learners to receive substantial amounts of input because "if [they] do not receive exposure to the target

language they cannot acquire it" (Ellis, 2005:38). Input in its various modalities (images, written text, audio) is used in vocabulary learning. There is a large body of evidence indicating that input mode affects incidental vocabulary learning (Webb, 2020). Words can be incidentally learnt through reading (e.g., Pigada & Schmitt, 2006; Waring & Takaki, 2003), listening (e.g., Pavia et al., 2019; Van Zeeland & Schmitt, 2013), reading while listening (e.g., Brown et al., 2008; Webb & Chang, 2012), and viewing (e.g., Peters & Webb, 2018; Rodgers & Webb, 2020). The differential effects of input mode on vocabulary learning can be explained using the theoretical foundations of Paivio's (1986) Dual Coding Theory (DCT) and the Cognitive Load Theory (CLT) (Sweller & Chandler, 1994; Sweller, 2010).

The DCT describes how the brain processes new information or comprehensible input. According to the DCT, cognition is made up of two distinct coding systems, one verbal and one nonverbal (Sadoski & Paivio, 2001). The verbal code is responsible for dealing with language, whereas the nonverbal code is responsible for dealing with non-linguistic information (e.g., objects, events). The DCT's central claim is that presenting information in both modes can improve recall, implying that exposing learners to two different modes improves word learnability because learners can make associations between verbal and nonverbal information (Sadoski, 2005).

The CLT is concerned with the distribution of cognitive resources during learning (Sweller, 1988; 1994). The amount of information that our working memory (WM) can hold at any given time is referred to as "cognitive load" (Cooper, 1998). There are three types of cognitive load in the CLT framework: intrinsic, extraneous, and germane. Intrinsic load is concerned with the inherent characteristics of the material to be learnt, extraneous load is caused by the instructional design used to present the material, and germane load (effortful learning) is imposed by learning processes (de Jon, 2010). According to Paivio (1986), the CLT predicts that learning from a mixed input mode (or bimodal input) is more effective than learning with a single input mode (unimodal). This is known as the modality effect.

Ginn's meta-analysis (2005) provides empirical support for the modality effect. The use of CLT in the context of L2 vocabulary learning/teaching has the potential to shed some light on the procedural aspects of vocabulary learning. A higher cognitive load can aid lexical learning because high demands on the cognitive system necessitate greater depth of processing, which has been shown to have more long-term effects than superficial processing (Craik & Tulving, 1975; Sandford & Sturt, 2002). However, it can also hamper learning when the learners' cognitive abilities are overburdened by the task at hand (Sweller, Van Merrienboer & Paas, 1998). Some learners may find the task of attending to multiple input

modes to be too demanding. As a result, teachers must consider cognitive load when employing various modes to expose learners to the L2. It is worth noting that different modes may lead to different results depending on the abilities of the learners and the tasks used (Vu & Peters, 2022a). There is also evidence that learners have input mode preferences, particularly when it comes to aural input (ibid).

The remainder of this section reviews the relevant studies that have looked and/or compared the impact of reading-only (RO) and reading-while-listening (RWL), which are the modes investigated in this thesis.

Input mode is frequently classified as unimodal, bimodal, and multimodal input. Unimodal input refers to language presentation in the form of written or aural input only. L2 vocabulary research on reading has attracted the most attention, partially because of the solid two-way relationship between reading and vocabulary growth (Nation, 2001). Reading can be further classified as extensive or intensive. The former refers to the reading of large quantities of texts, often chosen by the learners themselves, with 5% or less unknown running words (Webb & Nation, 2017), whereas the latter refers to close reading, which involves learners reading shorter texts in great detail in order to gain a complete understanding of every part of the texts. For lexical growth, extensive reading must have around 95% of vocabulary coverage. This means that learners should be familiar with approximately 95% of the words in the text, allowing them to understand the overall meaning without having to look up too many unknown words. A higher vocabulary coverage, ideally around 97% or more, is recommended for intensive reading (Nation, 2001). This ensures that learners have a good grasp of the language used in the text, allowing for a more in-depth analysis of the content.

A number of studies have been conducted to investigate the incidental learning of collocations through (extensive) RO (e.g., Pellicer-Sánchez, 2017; Szudarski, 2012; Vu & Peters, 2022a; Vilkaitė. 2017; Webb & Chang, 2020). Findings from these studies suggest that collocations can be learnt incidentally through reading, though some contradictory findings have been reported, particularly regarding the degree of learning of different collocation knowledge aspects. Pellicer-Sánchez (2017) explored the incidental learning of collocations in a classroom setting. Forty-one L2 learners read a text containing six adjective-pseudoword collocations. One week after the treatment, the learners were tested at the recall and recognition levels (i.e., form and meaning) of the target items. Results showed increased collocational knowledge at form recognition and recall of the target items. Vilkaitė (2017) also found that reading (academic texts) led to the incidental learning of transparent verb-noun collocations

(e.g., "generate debate") at the level of form recognition, but no effect was found at the level of form recall. On the other hand, Szudarski (2012) did not find a significant effect of RO on incidental learning of verb-noun collocations at the levels of form recall and form recognition. The lack of a significant effect in this study may be due to the type of collocations under investigation, i.e., verb-noun combinations with delexical verbs (e.g., "take a step"), as these items tend to be particularly challenging for L2 learners (Altenberg & Granger, 2001).

Given the importance of L2 reading, SLA researchers and teachers are concerned with finding ways to support the reading process and maximise the learning potential of reading (Krashen, 2004). Combining written text with other input sources, such as auditory input, is one way to support reading development. This is referred to as bimodal input or assisted reading. This type of incidental situation involves students reading and listening to the text at the same time. It is believed that assisted reading can have a scaffolding effect, particularly for lower-level learners who may find auditory input more difficult (Çekiç & Demirezen, 2020). This may be especially useful for learning collocations because they are thought to carry sound patterns. Prosody (e.g., stress, intonation) of multiword combinations highlights semantic unit boundaries (Pellicer-Sánchez & Boers, 2019). Lin (2010a, b) was the first to claim that MWUs exhibit phonological coherence, i.e., they form a single intonation unit that can then assist learners in decoding texts in RWL conditions. From a psycholinguistic standpoint, auditory input may aid learners in processing text in larger meaningful chunks (Conklin, Alotaibi, Pellicer-Sánchez, & Vilkaitė-Lozdiene, 2020). Another possible explanation for the added value of aural input is that MWUs are more common in oral discourse than in written discourse (Shin & Nation, 2008), so listening can be an effective way to develop both receptive and productive knowledge of MWUs. However, it is important to note that the facilitative effect of assisted reading has been primarily reported in studies using fictional stories as stimuli materials (e.g., Webb et al., 2013; Vu & Peters, 2022a). This does not necessarily mean that RWL is also more effective than RO when learners are exposed to other types of input, as increasing complexity of a text (plus listening to it) can result in cognitive overload (e.g., Dang et al., 2022).

A study by Webb et al. (2013) was the first to reveal that collocations can be learnt incidentally through RWL. In a classroom experiment, learners read while listening to one of four versions of a modified graded reader containing 18 target collocations repeated 1, 5, 10, or 15 times. Results revealed greater learning gains at the level of form and meaning recognition when compared to form

and meaning recall, although productive knowledge of the form and meaning of the collocations was not pretested. This study used immediate posttests only, so little is known about the retention of the learning gains. Repetition had a positive effect on collocation learning.

The few studies that have compared RWL to RO have produced mixed results (Dang et al., 2022; Webb & Chang, 2020; Vu & Peters, 2022a). Webb and Chang (2020) compared the effects of RO, listening, and RWL on the incidental acquisition of collocations from a grader reader. Their EFL learners were asked to read only, listen to, or read while listening to a graded reader containing seventeen collocations in six sessions over the course of three weeks. They found that RWL had an advantage over both RO and listening conditions. In a similar vein, Vu and Peters (2021) compared the effects of three reading modes: RO, RWL, and reading with textual input enhancement. Their findings corroborated Webb et al. (2020), as they also found that RWL to a grader reader resulted in more collocation gains than RO, although reading with the collocations typographically enhanced was the most beneficial input mode for collocation learning. By contrast, Dang et al. (2022) did not find an advantage of RWL over RO. Unlike Webb and Chang (2020) and Vu and Peters (2021), this study used an academic lecture as source of input, which may have differing effects compared to using fictional stories. This contradictory evidence about the benefit of RWL point to an effect of the type of reading materials employed in these studies (fiction vs nonfiction; academic vs non-academic). Crucially, more research is needed about the impact of using different text types in L2 vocabulary studies.

3.5.1.2. Text type

Text type is an input-related variable that has yet to be thoroughly investigated. Previous studies have largely focused on narrative texts (i.e., non-academic input), particularly on fictional stories (Pellicer-Sánchez, 2017; Szudarski, 2012; Szudarski & Carter, 2016; Vu & Peters, 2022a; Webb & Chang, 2020; Webb et al., 2013). Expository texts and academic input have not been examined to the same extent (e.g., Dang et al., 2022, Vilkaitė, 2017). While both academic texts and fiction offer valuable vocabulary learning opportunities, there are some key differences when it comes to learning collocations. Academic input is formal and adhere to established conventions, while fiction often allows for more creative and imaginative language use. As for content specificity, fiction covers a wider range of topics with a more general language, while academic texts concentrate on specialised subjects and technical language, leading to more domain-specific collocations. Lexical differences between text

types may affect the degree to which vocabulary (and what type of vocabulary) is learnt. For example, in a corpus analysis of narrative and expository texts for children, Gardner (2004) found that narrative texts were superior for high frequency vocabulary, whereas expository texts were better for academic vocabulary and unique word acquisition. Academic texts are likely to have a higher level of complexity as they tend to use a rich and diverse vocabulary (Vilkaitė, 2017). This may pose a greater challenge for learners than nonacademic texts, particularly if different input modalities are combined in a given treatment (see Dang et al., 2022). The language register also reflects this. Academic materials have a formal and technical language register, while fiction can encompass a wide range of registers, from informal to poetic. Further, academic texts have a formal and precise style of writing, and learning collocations from such texts can help learners adopt a more appropriate language register for academic writing and communication.

The demands of different types of input (and their features) affect the amount of learning (and the learning experience itself) (e.g., Dang et al., 2022 vs Webb & Chang, 2020). Text-related features (e.g., more/less specialised vocabulary, information density, language register, degree of abstraction) may thus make distinctive contributions to the incidental learning of collocations. Investigating whether learners can effectively learn collocations from reading academic texts is worthwhile because it provides insights into the effectiveness of using authentic, domain-specific materials for vocabulary learning. While learning collocations from fiction has its benefits, academic texts offer unique opportunities for learners to acquire collocations specific to their fields of study and to develop language skills necessary for academic and professional success. Importantly, academic texts are an essential part of the learning journey for EFL learners who are pursuing higher education or academic studies. Understanding collocations in academic contexts enhances their ability to read, comprehend, and produce academic materials effectively. For this reason, this PhD thesis focuses on expository texts (of an academic kind) rather than narrative short stories. If learning of collocations from expository texts can be demonstrated, this would open up a large potential source of incidental learning for advanced/university-level learners, who are already reading and engaging with texts of this type as part of their studies.

3.5.1.3. Repetition

A central factor in incidental vocabulary learning is repetition (e.g., Ellis, 2002; Nation, 2013). Repetition refers to the quantity of encounters with a lexical item. The more repetitions with a word, the

more likely is to be learnt. Frequency of encounters of new and partially known lexical items is a significant factor in both attaining word knowledge and improving vocabulary retention (e.g., Nation & Wang, 1999; Waring & Takaki, 2003; Webb, 2007; Pellicer-Sánchez, 2020). There is however no commonly accepted threshold number for incidental learning to occur, and estimates vary, from six and eight to more than ten for individual words (Pellicer-Sánchez & Boers, 2019). There is robust evidence attesting to the value of repetition (e.g., Chen & Truscott, 2010; Horst et al.,1998; Pellicer-Sánchez & Schmitt, 2010; Pigada & Schmitt, 2006; Waring & Takaki, 2003; Webb, 2007). A meta-analysis conducted by Uchihara et al. (2019) found a medium effect correlation between repetitions and learning, which translated into 12% of the variance in incidental vocabulary learning (Schmitt & Schmitt, 2020).

Different repetitions technique are used to increase repetition. Repetitions can be either spaced or massed together. The practise of leaving intervals between encounters with a lexical item is referred to as spaced repetition. It is a common technique in intentional vocabulary learning in which teachers take advantage of repeating vocabulary over time (Webb & Nation, 2017). The process of seeding the input with several encounters with a word to increase its salience is known as massed repetition[5]. This is typically done over a short space of time and is thus perceived as a shortcut way of enriching the input (Durrant & Schmitt, 2010). Repetitions can be further divided into verbatim or varied. The former refers to various encounters with a target word in the same context, while the latter refers to varied encounters with a word in different contexts. These typical repetition techniques can have an impact on the strength of vocabulary learning (e.g., Durrant & Schmitt, 2010; Sonbul & Schmitt, 2013; Toomer & Elgort, 2019).

So far, the message is clear: repetition affects incidental vocabulary learning. According to Zipf's law, however, the likelihood of learners encountering the same collocations repeatedly in a text or in language textbooks is not very high (Boers et al., 2014, Pellicer-Sánchez & Boers, 2019). Due to the relative rarity of occurrence of collocations (compared to that of single words) (e.g., Durrant & Schmitt, 2010; Martinez & Schmitt, 2012; Shin & Nation, 2008), incidental collocation learning relies heavily on the learning/teaching (reading) materials and how these are designed to include collocations at higher rates of recurrence than in naturally occurring language. This is the rationale behind the promotion of extensive reading programmes for collocational development (Webb & Nation, 2017). Repeated

---

[5] Massed repetition is known as "input flood" technique in the focus-on-form literature.

encounters with collocations during reading is beneficial because learners acquire the knowledge of the form and meaning whilst improving the fluency with which the collocation is processed in context (Pellicer-Sánchez, 2020). Most learners, however, do not read extensively in the L2 (Peters, 2018; Lindgren & Muñoz, 2013). They prefer other out-of-class activities such as TV viewing (Kuppens, 2010). One weakness of previous research is that it fails to consider the learners' preferences as well as the type of input they are exposed to. Other concerns relate to the use of long, narrative texts as reading materials (Vilkaitė, 2017). These considerations point to the need for further research examining collocation learning from short texts. Short texts may be an effective way to stimulate learning if the input quality is adequate and repetitions occur in a short period of time. For such learning to occur, collocations need to be seeded into them. Massed repetition may be thus particularly beneficial as the intervals left in spacing repetitions may weaken its effect (Boers, Demecheleer, Coxhead, & Webb, 2014). In sum, the acquisition of collocations not only relies on the quantity of repetitions, but also the quality and aspects related to the type and amount of input the learners are exposed to (e.g., Hoey, 2005).

Research into the incidental learning of collocations has shown varying effects of repetition, with most studies demonstrating a facilitative effect of repeated exposure (Durrant & Schmitt, 2010; Toomer & Elgort, 2019; Vilkaitė, 2017; Webb et al., 2013; Webb & Chang, 2020). Webb et al. (2013) looked at the effect of repetition (1, 5, 10, 15) on the acquisition of verb-noun collocations through reading-while-listening to graded readers. Unsurprisingly, results showed that 15 encounters was the most advantageous figure for collocation intake, although 5 repetitions also led to significant gains. This study only included immediate posttests, so it is unclear how durable these learning gains were. In a similar study, Webb and Chang (2020) found a positive relationship between repetition and learners' uptake of collocations from reading-while-listening. A frequency effect, however, was not found in the reading-only condition, which suggests that assisted reading could have inflated the learning gains in Webb et al. (2013).

Focusing on retention, Durrant and Schmitt (2010) also found a positive effect of repetition. In this study, they examined adult learners' retention of adjective-noun collocations in three conditions: single presentation (one repetition in a single context), verbatim repetition (two repetitions in the same sentence context), and varied repetition (two repetitions in two different sentences). Both varied and verbatim repetition improved recall over the single exposure condition, with verbatim being slightly more

effective. This result suggests that repeated exposure to the same contexts has a greater influence on learners' collocation recall than repeated exposure to the same collocations in different contexts. The authors speculate that varied repetition increases the learners' cognitive burden, whereas the nature of the fluency-based instruction does not. It should be noted that this study used sentence-level reading rather than full-text reading.

Vilkaitė (2017), Sonbul and Schmitt (2013) and Toomer and Elgort (2019) focused on academic texts. Vilkaitė (2017) examined the incidental learning of verb-noun collocations. Her participants read two texts (~2,000 words) containing 4 repetitions with 15 target collocations in one reading session. Results showed that 4 encounters can lead to incidental acquisition of collocation at both form recognition and form recall levels. Sonbul and Schmitt (2013) compared the effects of enriched (incidental), enhanced (collocations presented in red font and bolded), and decontextualized input (collocations taught in isolation) on the acquisition of collocations on the acquisition of medical collocations. The incidental treatment (enriched condition) included three repetitions of each target item. They found that the decontextualised treatment was not significantly better for long-term retention than the enriched condition which included three contextual exposures to the collocations in the reading text. However, the collocations in the incidental treatment were embedded in meaningful contexts. In a conceptual replication and an extension of Sonbul and Schmitt (2013), Toomer and Elgort (2019) looked at the effects of spaced repetition on the acquisition of medical collocations. Their participants read nine texts (~500 words) containing 9 occurrences with the target collocations over two consecutive days. Results showed that repetition resulted in significant gains in both the immediate and delayed form-recognition posttests.

On the other hand, a few previous studies did not find an effect of repetition on incidental learning of collocation (Macis, 2018; Pellicer-Sánchez, 2017; Szudarski & Carter, 2016). This lack of an effect is likely due to salient features that come into play when researching collocations. Szudarski and Carter (2016) explored the effects of input flood and input flood plus enhancement on L1-Polish EFL learners. The reading materials consisted of six short stories which contained 6 and 12 encounters with the target collocations. They found that input flood alone did not lead to gains while input flood plus enhancement did. Interestingly, the results showed significant gains for form recall and form recognition when learners encountered the collocations 6 times, while 12 exposures improved meaning recall only. The authors concluded that more repetitions do not necessarily lead to greater learning gains at all levels of

collocational knowledge. Pellicer-Sánchez (2017) investigated the acquisition of adjective-pseudowords collocations. Her intermediate English language learners read two versions of a short story (~2000 words) with 6 target items embedded in it 4 and 8 times. One week after the treatment, learners took a combination of paper and pencil and interview tests on the form, meaning, and collocation of the target items at recognition and recall levels. She found that repetition did not have a significant effect on incidental learning of form and meaning of collocations. Unlike this prior research, Macis (2018) conducted three case studies to examine whether repetition promoted meaning recall of collocations. Her participants read a modified version of a novel with 38 target verb-noun and adjective-noun collocations embedded in it. Results showed that, while repetition was not always statistically significant, it was consistently positive.

These studies show that repeated exposure is a central factor in incidental vocabulary learning. Importantly, learning may not be solely dependent on repetition alone as other factors are likely to interact with repletion, which influence how collocations are acquired and retained by learners. Some of these factors concern the quality of the encounter (Laufer, 2005), the collocation's unique properties (e.g., novelty, semantic transparency, formal salience), its relevance for comprehension (particularly for reading), the availability of an L1 equivalent (congruency), and the characteristics of the learners (e.g., level of proficiency) (Pellicer-Sánchez & Boers, 2019).

3.5.1.4. Context informativity

Context[6] plays a significant role in vocabulary learning from two obvious reasons. To begin with, most L1 and L2 vocabulary growth occurs through inferring word meanings from context (Nagy, 1995; Paribakht, & Wesche, 1999; Rott, 1999). Second, word meanings are context sensitive, changing from one context to another (Labov,1973). Though it is widely acknowledged that context affects word learning, the role and extent to which context influences vocabulary learning has been debated. Some L1 researchers have voiced scepticism about the importance of context (Beck, McKeown, & McCaslin, 1983; Schatz & Baldwin, 1986). For example, Beck et al. (1983) stated that many natural texts do not provide supportive contexts as they are often ambiguous or even misleading for learning word

---

[6] The term context can be used more broadly to refer to various aspects of learning contexts; here, context in vocabulary learning refers to contextual support or context informativity.

meanings. In an article subtitled "All contexts are not created equal" they proposed a model of contextual support that displays context as a very unreliable source of information. By contrast, Nagy and colleagues (Nagy, Herman, & Anderson, 1985) claim that context aids L1 vocabulary acquisition.

The role of contextual support also remains unclear in L2 vocabulary research, partly because of the various ways in which context has been conceptualised and operationalised, but also because the types of contexts that facilitate the learning of unfamiliar words remain uncertain (e.g., Zahar et al., 2003). In general, research shows that the chances of a word being learnt from context is small (Nation, 2001). These small, gradual gains, however, should not discount the value of context for L2 vocabulary development (Nagy et al., 1985). Small gains can lead to large ones if learners do large quantities of reading and/or pay some deliberate attention to learning vocabulary (Nation, 2001). Inferencing is also the most commonly used strategy by L2 learners when reading (Schmitt, 2010). Some consider the ability to infer the meaning of new vocabulary using context clues to be a prerequisite for incidental learning while reading (and listening) (Van Zeeland, 2014; Webb & Nation, 2017). Context is thus an important factor that should not be overlooked. The studies that have examined the role of context have largely focused on single-word vocabulary (e.g., Joe, 2010; Laufer & Shmueli,1997; Teng, 2016; Webb, 2008; Zahar et al., 2001). There are good reasons to believe that context may aid incidental collocation (and MWUs in general) learning by facilitating meaning inference from linguistic environments (Nation, 2001).

Guessing from context is the most frequent strategy used by L2 learners when dealing with novel MWUs (Schmitt & Schmitt, 2020). It can be argued that the incidental uptake of collocations, particularly of (partly) incongruent items (e.g., "raise cash"), could be largely dependent on the contextual information that learners could use to infer their meanings. For example, it is likely that a learner will infer the meaning of a new collocation if they have a good lexical coverage of the text. In such scenario, the likelihood of a collocation being learnt is likely to increase when context cues are readily available and provide an adequate source of information about the collocation. That is, in favourable conditions, guessing from context is likely to occur (Nation, 2013).

Quantifying the "quality" of context can be laborious as it frequently involves the application of various scales and/or models of contextual support. For example, Webb (2007a) developed a context specification scale to rate the level of informativity of a given context. Webb's scale is meant to display how likely a word can be guessed correctly using a scale from 1 (extremely unlikely, no context cues

available/misleading) to 4 (good chance of guessing the meaning correctly). This scale has been used in subsequent studies (e.g., Hu, 2013; Teng, 2016; Webb, 2008). Other researchers (e.g., Joe, 2010; Zahar et al., 2001) have used Beck et al.'s (1983) model of contextual support where contexts can be identified using a four-point nominal scale: misdirective (misleading), nondirective (which do not provide particular clues about the meaning), general (some basic information provided) or directive contexts (word presented in an implicit definition). For example, Zahar et al. (2001) assessed the degree of contextual richness of the immediate context of 30 target words and categorised it using Beck et al.'s model.

Findings from single-word studies serve as a starting point for further research into its effect on collocation learning. Taken together, results from single-word studies indicate a positive effect of context, particularly in gaining knowledge of word meanings (e.g., Teng, 2016; Webb, 2008). Webb (2008) investigated the effect of context (i.e., more informative vs less informative) on the incidental learning of 10 single words. Fifty EFL Japanese learners were asked to read three sets of 10 sentences, each sentence containing 1 of 10 target words. The contexts in which the target words appeared were rated by two native speakers of English according to Webb's context rating scale (adapted from Webb, 2007a). Participants were then administered four tests that measured recall of form, recognition of form, recall of meaning, and recognition of meaning. Results showed that more informative contexts triggered greater incidental acquisition of word meaning, but context had little effect on the acquisition of form. In line with Webb (2008), Teng (2019) also found that contextual richness fosters the acquisition of word meanings, but contrary to Webb (2008), he also found a significant effect for learning the word forms. This difference in the results is likely due to the fact that Teng (2019) included three different frequencies of word exposure (1, 5, and 15 repetitions). Findings from these studies support the claim that learners might overlook words embedded in contexts which do not supply contextual cues for inferring their meanings (Hulstijn, Hollander, & Greidanus 1996). It should be noted, however, that both Webb (2008) and Teng (2019) included two distinctions of "context" (informative/non-informative) only, and the fact that Webb's study used disguised L2 forms for the target words to ensure that participants did not have prior knowledge of the words, e.g., the form "ancon" was used instead of the target item "hospital". If the facilitative role of context for incidental learning collocation can be demonstrated, this would have an important bearing on L2 pedagogy. To fill this gap, Experiment 3 will investigate how different levels

of context informativity (high, mid, low) affect form recall and meaning recall of lexical collocations by adult EFL learners.

3.5.2. Learner-related factors

The most researched learner-related factor is the learners' prior vocabulary knowledge, which has been shown to have positive effects on the incidental learning of collocations (Puimège & Peters, 2019, 2020; Vilkaitė, 2017; Vu & Peters, 2022a). It is widely accepted that individual differences in overall vocabulary knowledge affect all aspects of L2 language development and vocabulary growth (Milton, 2013; Stæhr, 2008). This is the reason learners with larger vocabulary sizes learn more collocations and perform better in L2 collocation studies (e.g., Vu & Peters, 2022a). This is likely due to the Matthew effect[7], or the reciprocal causal relationship between gains in lexical knowledge and previously existing vocabulary knowledge (Stanovich, 1986). It should be also noted that lexical coverage increases exponentially with a learner's vocabulary knowledge, which in turn results in better reading comprehension skills (Webb & Chang, 2015). This appears to enhance processing, so that L2 learners with large vocabularies (by extension, higher proficiency level) can devote more attentional resources to noticing and learning new lexical items (Vu & Peters, 2022a). From this standpoint, the message is clear: prior vocabulary knowledge predicts collocational knowledge.

3.5.3. Item-related factors

The learning burden of a collocation is also influenced by a range of item-related factors. Four item-related factors that are of particular interest to this thesis will be discussed in more detail in the following sections: congruency, compositionality, corpus frequency, and association strength.

---

[7] The Matthew effect, also known as the rich-get-richer phenomenon, is a term in vocabulary research that refers to the idea that L2 learners with larger vocabularies tend to learn new words more easily and quickly compared to those with a smaller vocabularies.

3.5.3.1. Congruency

Congruency refers to the probability of L1-L2 collocational equivalence. Collocations that have L1-L2 lexical similarities are categorised as congruent whereas collocations that have lexical items that are different in the two languages are labelled as incongruent (Nesselhauf, 2003). When the congruency between L1 and L2 is high, it can facilitate the learning of collocations. Conversely, when congruency is low, it can lead to difficulties in acquiring collocations in the target language. For example, the English collocation "fast food" is congruent with its Spanish counterpart "*cómida rápida*". On the other hand, the English collocation "soft drinks" has no literal equivalent in Spanish (i.e., "*refrescos*"). Previous research shows that learners are likely to transfer from the L1, i.e., they use their existing language knowledge to make connections and learn the collocations more effectively. However, if collocations are quite different between the two languages, learners may struggle to find analogous structures, leading to (L1-induced) errors (Peters, 2016). Collocations can be also sensitive to subtle semantic nuances that may not have direct equivalents in the learners' L1. Learners with high congruency between their L1 and L2 may find it easier to grasp these nuances and use collocations appropriately. Collocations can differ between languages with different cultural backgrounds because they are frequently influenced by situational and cultural contexts (Wood, 2020) . If learners have been exposed to similar collocations in their L1, high congruency may ensure that this is the case, making it simpler for them to recognise and pick them up in the target language through increased frequency. These are the main reasons why congruency affects the difficulty learners may experience in learning collocations (e.g., Nesselhauf, 2003; Peters, 2016; Wolter & Yamashita, 2015) and the reason why the learning burden of collocations is lighter when they are similar in the L1 and the L2 (Wolter, 2020). Previous studies show a production and processing advantage for L2 congruent collocations that have L1 equivalents (e.g., Nesselhauf, 2003; Yamashita & Jiang, 2010; Laufer & Waldman, 2011, Wolter & Gyllstad, 2011, 2013). There is also ample evidence showing that L2 learners tend to directly transfer collocational patterns from their L1 to the L2 (e.g., Bahns, 1993; Nesselhauf, 2003, 2005; Peters, 2016; Wolter & Gyllstad, 2013), particularly in L2 writing. For example, in their analysis of learner corpus of written essays, Laufer and Waldman (2011) found that a third of the collocations used by EFL learners were atypical, with L1 being the main source of about half of the errors.

The studies that have looked at the effect of congruency on incidental collocation learning have yielded mixed results. For example, a study by Vu and Peters (2021) showed that congruency was a

significant predictor of incidental collocation recall, which suggests that L2 learners find incongruent collocations more difficult to understand and recall due to the lack of L1 equivalents. These findings are supported by processing studies by Wolter and Gyllstad (2011, 2013) and Yamashita and Jiang (2010) who found that congruency has a positive effect on the processability of collocations. Nguyen and Webb (2017) found that congruency predicted receptive knowledge of collocation, but Puimège and Peters (2020) failed to find such an effect, as they found that congruency was not related to the incidental learning of MWUs (including collocations) from watching audiovisual input. However, the interpretability of these findings might depend on the nature of the intervention, the mode of input, and the type of collocational knowledge tested (i.e., form or/and meaning).

One of the reasons why previous studies find differences between congruent and incongruent collocations is because incongruent collocations are opaque while congruent ones are transparent (Conklin & Carroll, 2018). Collocational congruency between the L1 and the L2 affects collocation learning, particularly concerning transparency and opaque meanings. Although semantic transparency plays a role in the processing (and learning) difficulty of collocations (see Gyllstad & Wolter, 2015), it is a factor that has been often overlooked in favour of congruency for two main reasons. First, due to the availability of large corpora like the COCA and the existence of association measures (see Section 3.5.3.4.), many previous studies have adopted a frequency-based approach to collocations because it is less time consuming and more efficient than the phraseological approach. Second, human judgments, which can be arbitrary and differ from person to person, are necessary to determine the level of semantic transparency of a given collocation. Due to this limitation, previous research did not consider semantics as a factor. However, using measures of probabilistic strength of association as the only criterion is also limited, particularly if we accept that factors other than frequency may be more salient for L2 learners (Wray, 2002). In the current PhD thesis, a mixed-method approach to collocations was followed so as to consider semantic factors. One of these is semantic compositionality.

3.5.3.2. Compositionality

Compositionality is often seen as a factor affecting MWU/collocation difficulty (Martinez & Murphy, 2011). Semantic compositionality refers to the degree to which the meaning of a collocation can be predicted from the meaning of its parts. For example, the meaning of the collocation "office hours" is compositional (transparent) whereas the meaning of the collocation "hot dog" is non-

compositional (opaque) as its meaning cannot be inferred from the meaning of its individual components ("hot" and "dog") (Cruse, 1986). Non-compositional collocations pose a greater challenge for learners because, even with knowledge of the meaning of the constituents making up a collocation, the meaning of that collocation cannot be derived from combining the meaning of the constituents (Martinez & Murphy, 2011).

Collocations do not necessarily fall into a "fully compositional" or "non-compositional" category; on the contrary, they are known to display a continuum of compositionality (e.g., Boers & Webb, 2015; McCarthy, Keller & Carrol, 2003). The fact that collocations are not (non)compositional to the same extent has led some to claim that semi-compositional collocations are the most problematic for learners to acquire (Nesselhauf, 2005). A semi-compositional collocation (e.g., "zebra crossing") will be less noticeable than a fully non-compositional (e.g., "cloud nine") or compositional one (e.g., "swimming pool"), which can mislead learners about the collocation's real meaning. The degree to which collocations may be perceived as more or less compositional has been shown to vary from person to person (Boers & Webb, 2015), which poses problems in our ability to quantify the degree of compositionality in absolute terms. For example, a study by Macis and Schmitt (2017) reported a range of individual judgements on a transparency task they used to assess the degree of compositionality of figurative meanings of collocations. This is possibly one of the reasons why little attention has been paid to examine its role on collocation/MWUs learning. In the cited study, Macis and Schmitt (2017) found that semantic compositionality did not affect the knowledge of the idiomatic meaning of collocations. This result was unexpected because the semantic cohesion of the individual parts of a collocation are thought to be salient, which should be useful for the interpretation of its figurative meaning. The authors conclude that the relationship between the individual constituents of a collocation is too subjective and so this result reflects that learners' intuitions are often inaccurate.

3.5.3.3. Corpus frequency

The frequency in which a lexical item/collocation occurs in a language influences its acquisition, use, and processing (Ellis, 2002). It is regarded as one of the best predictors of usefulness and vocabulary acquisition (Schmitt & Schmitt, 2020). By definition, the most frequent items (i.e., high frequency lexical items) in language are more likely to occur in the input, and thus learners will acquire them before less frequently occurring words (Schmitt, 2010). Frequency is an index of a learner's prior

encounters with a particular linguistic item (Janssen & Barber, 2012). Frequency information is extracted using language corpora (Arnon & Snider, 2010; Bannard & Matthews, 2008; Siyanova-Chanturia et al., 2011; Tomasello, 2005; Wray, 2006). As a result, corpus frequency is used as an indicator of exposure information. Two main measures are frequently employed in extracting the frequency of collocations: raw-frequency of occurrence and log-frequency-scores of the occurrence of specific sequences. Although corpus measures have some limitations, including the inability to accurately reflect the amount (and type) of input received by individual learners (González-Fernández & Schmitt, 2015), they serve to quantify the amount of exposure learners may have had by employing various parameters.

Corpus frequency has been found to explain a large part of the variance in vocabulary acquisition (e.g., González-Fernández & Schmitt, 2015; Redwood & Schmitt, 2011; Wolter & Gyllstad, 2013; Wolter & Yamashita, 2018). With regard to collocation learning, a meta-analysis of nineteen tests of collocational knowledge by Durrant (2014) found that corpus frequency (extracted from the BNC and COCA) correlates moderately with L2 learners' knowledge of collocations, although he found that the strength of this correlation varies widely depending on the language corpora. Frequency-data retrieved from COCA were more strongly related to learners' knowledge than that of the BNC. This is likely to be the result of the characteristics of the corpora. Unlike the BNC, which contains texts produced in the 1980s and 1990s, the COCA database (and its frequencies) is updated every year, which makes it a more adequate guide to investigate collocations as they are "a highly context-sensitive phenomenon" (Gries & Durrant, 2020:149). Certainly, prior studies indicate that learners are sensitive to the frequency-driven distribution of collocations (e.g., González-Fernández & Schmitt, 2015; Puimège & Peters, 2019; Wray, 2002), although there are exceptions (e.g., Vu & Peters, 2021). For example, González Fernández and Schmitt (2015) found moderate to strong correlations between learners' existing knowledge of collocations and corpus frequency (as indicated by the COCA). More specifically, they found that corpus frequency was the best predictor of learners' recall of collocations (compared association measures), accounting for 20% of the variance. Puimège and Peters (2019, 2020) found a small correlation between COCA frequencies and the learning of collocations (as well as other multiword combinations). In a similar vein, an audiovisual study by Majuddin, Siyanova-Chanturia, and Boers (2021) found that corpus frequency predicted learners' knowledge of collocations at the level of meaning recognition. On the other hand, Vu and Peters (2021) did not find an effect for corpus

frequency. However, this contradicting finding might have resulted from the different type of L2 exposure (reading vs. viewing television), or type of target items (different types of MWUs vs collocations only).

3.5.3.4. Association strength

The association strength is the strength of co-occurrence of two or more words, which is often assessed on the basis of corpus-derived association measures (AMs) such as mutual information (MI), t-score, and log Dice. AMs play a major role in the identification of collocations as they combine information about corpus frequency with other collocational features that can be statistically expressed (e.g., Evert, 2004; Hunston, 2002; McEnery & Hardie, 2012). More importantly, they have implications for L2 collocational research (Gablasova, Brezina & McEnery, 2017). Gablasova et al. (2017) highlight the importance of choosing the right measure as some seem to favour different lexical properties of the collocations. For example, the MI score appears to feature infrequent and more specialised collocations, whose constituent elements may be low-frequency items themselves (Schmitt, 2012). For example, technical terms and idiosyncratic phrases are highlighted by MI scores (e.g., "post-grad") (ibid). Proper names (e.g., "Noam Chomsky") will also yield a high MI because the chances of seeing either word without the other are small. Similarly, wordstrings of low frequency lexical items (e.g., "wreak havoc") will inflate MI scores (Boers, 2020). MWUs with high MI scores tend to be more difficult to acquire because they often contain low-frequency words (Nguyen & Webb, 2017). A weakness of the MI score is that it does not work well with very low frequencies. On the other hand, the t-score takes frequencies into account while it also yields significant collocates that tend to co-occur. This is important because not all frequent collocations have a high t-score, but collocations identified by the t-score are frequent (Gablasova et al., 2017). Following the logic of corpus-based studies, most research on L2 collocations have used the MI score or the t-score (e.g., Durrant & Schmitt, 2010; Nguyen & Webb, 2017; Siyanova & Schmitt, 2008; Siyanova-Chanturia, 2015; Vu & Peters, 2022a; Webb & Chang, 2020), but the rationale for this selection is not always clear, which suggests that these association measures are not fully understood (González Fernández & Schmitt, 2015). The t-score corpus-derived measure that indicates the confidence with which two words are associated, i.e., their co-occurrence is not random. The t-score is also regarded as a reliable measure in identifying collocations as it is dependent on the corpus size (Hanston 2002; McEnery, Xiao, Tono, 2006). More recent investigations have introduced the log Dice as an alternative to the MI-score. The log Dice is thought to be better suited for SLA

research because, unlike other measures, it does not rely on random occurrence and can be used to compare across different language corpora while still testing collocation strength and its exclusivity (Gablasova et al., 2017). log Dice is also different from the other measures because it has a theoretically fixed maximum value of 14, but it is usually less than 10 (Rychlý, 2008).

There is some evidence that the association strength of a collocation can affect its learnability. For example, Nguyen and Webb (2017) found that association strength (as measured by MI scores) significantly predicted learners' form recognition (receptive knowledge) of verb-noun and adjective-noun collocations. This result was somewhat expected given that the MI score is a measure of tightness (González-Fernández & Schmitt, 2015), which thus seems to capture the degree to which collocations are recognised as coherent wholes (Nguyen & Webb, 2017). Puimège and Peters (2019) also found an effect for MI scores and the incidental uptake of collocations from audiovisual input. On the other hand, Vu and Peters (2022a) did not find a relationship between MI scores and form recall of collocations from reading. These mixed findings are likely to be the result of the differences between studies (one-off vs longitudinal), type of target items (different types of collocations), and type of input modality (reading vs audiovisual input).

3.6. Summary

This chapter has provided an overview of the role of MWUs and collocations in L2 vocabulary learning and the different learning conditions in which they can be acquired, with a focus on incidental learning. The two main approaches used for the identification of MWUs, and collocations have been presented, and we have seen that classifying collocations can have benefits in theory and practice, as these are useful in identifying collocational patterns. The learning difficulties that adult L2 learners may have with collocations have also been discussed, particularly in classroom environments. Finally, we have seen the various factors that can affect the acquisition of collocations.

The remainder of this PhD thesis will examine in more detail the issues discussed in this chapter. While there is some evidence for the benefit of combination of input modalities (e.g., RWL) on incidental collocation learning, the effectiveness of combining input modes for collocation learning from academic texts is still not clear, which motivated Experiment 1, reported in Chapter 4. A second area to be researched concerns the effect of repetition for the acquisition of different types of collocation knowledge, as the inconclusive findings about the effect of repetition presented in this chapter call for

further research. Experiment 2 (Chapter 5) focuses on this factor. Finally, Experiment 3 (Chapter 6) examines the role of context informativity in gaining declarative knowledge (form recall and meaning recall) of collocations, as it is not yet known how context can affect the extent to which collocations can be learnt from meaning-focused input.

4. The effect of reading mode

This chapter presents Experiment 1, which compares learning collocations from reading-only (RO) with learning collocations from reading-while-listening (RWL) to an academic text. Before presenting the experiment, I explain the background of this experiment focusing on the importance of academic input as a source of incidental collocation learning among university EFL learners.

4.1. Background

As noted in Chapter 3, there is potential for collocation to be incidentally learnt through RO (e.g., Pellicer-Sánchez, 2017; Szudarski & Carter, 2016) and RWL (e.g., Vu & Peters, 2022a; Webb & Chang, 2020). The few studies that have compared gains across RO and RWL have mostly focused on narrative stories (i.e., non-academic input) as a source of incidental collocation learning (e.g., Vu & Peters, 2022a; Webb & Chang, 2020). While those studies provide some evidence that RWL is more beneficial than RO, its effectiveness when learners are exposed to academic input is unclear (e.g., Dang et al., 2022). Dang et al. (2022) investigated the effect of various input modes (RO, listening-only, RWL, viewing, and viewing with captions) on the incidental learning of collocations using an academic lecture as the source of input. The findings revealed that RWL did not result in greater collocation gains relative to RO. Although this study only investigated academic texts, the results suggest that different types of input (academic vs non-academic) may modulate the efficiency of RWL, which have produced contrasting findings. On the one hand, studies using graded readers have reported that RWL facilitates the incidental learning of collocations better than RO (e.g., Vu & Peters, 2021; Webb & Chang, 2020). Dang et al.'s study (2022), on the other hand, found that RWL did not result in increased gains compared to RO using an academic lecture. RWL to academic input can be more challenging due to the intensity of processing more complex texts. It is thus plausible that the informational density of these academic texts relative to narrative ones may also affect the learners' ability to integrate information from different channels. Academic and expository texts could also be more difficult to understand if learners lack knowledge of the specialised vocabulary that occurs in a given academic subject (Coxhead; 2020; Nation, 2016).

The following section discusses academic input as a source of incidental collocation learning, as well as the few studies that have investigated this type of input.

4.1.1. Academic input and collocations

An increasing number of European Higher Education Institutions (HEIs) offer degrees taught in English as a result of the completion of the Bologna Process (Maiworm & Wächter, 2014). A 2021 report published by the British Council identified 27,874 English-taught degree programmes outside anglophone countries. This represents a growth of 77% compared to January 2017, showing a growing trend of English-medium programmes in international HEIs. As more courses are offered in English, academic reading has become an integral component of L2 learners' study at universities where English is the medium of instruction.

Academic lexis is more specialised when compared to vocabulary found in general English texts such as fiction or newspapers (Coxhead, 2000; Gardner & Davies, 2014). Nation (2001) classifies specialised vocabulary into academic vocabulary and technical vocabulary. The former refers to vocabulary that is rare in non-academic texts but frequent across a wide range of academic subjects (e.g., "evident") whereas the latter refers to specialised vocabulary that is essential to learning a particular discipline (e.g., "insulin" in medicine) (Webb & Nation, 2017). L2 learners studying in English-medium institutions need to know a substantial amount of specialised lexis to be able to understand academic input (Coxhead, 2020). Knowledge of specialised vocabulary is not only central to learning content, but also to cope with understanding and acquiring new vocabulary (Nation, 2006). A concern regarding academic input is that L2 learners might find the process of reading too demanding due to the relatively large proportion of subject-related vocabulary and the fact that academic texts tend to be densely packaged with content. However, in the case of university students, learners are likely to spend time on in-depth reading for their academic subjects, with sustained attention on texts (Coxhead, 2018).

Research on the incidental learning of collocations through academic sources is limited. Vilkaitė (2017) examined the incidental learning of nonadjacent and adjacent verb-noun collocations (e.g., "generate debate"). Her participants read two academic texts (~2,000 words long) containing 15 target collocations in one reading session. She found that reading academic texts can lead to the incidental learning of collocations, at least to the recognition level of knowledge. Sonbul and Schmitt (2013) compared the effects of enriched (incidental), enhanced (collocations presented in red font and bolded),

and decontextualized input (collocations taught in isolation) on the acquisition of collocations. Sixty-two advanced ESL speakers read shot passages containing medical collocations. Each collocation was repeated three times. Results showed that learners' explicit knowledge of collocations (form recognition and form recall) increased under all treatment conditions, but the enhanced condition resulted in greater gains at immediate learning than the enriched condition. However, long-term retention was significantly better in the enriched treatment, where collocations were embedded in rich, meaningful contexts. In a conceptual replication of this study, Toomer and Elgort (2019) investigated at the incidental learning of medical collocations (e.g., "stone heart") embedded in nine academic texts. The results showed that RO led to significant gains at explicit cued recall and form recognition of collocations, although this was likely due to the number of encounters with the items (nine exposures across three sessions over two days). Dang et al. (2021, 2022) conducted two recent studies to investigate the value of academic lectures as a source of input. Dang et al. (2021) found that viewing an academic lecture could lead to significant collocation gains at the form recognition level. Dang et al.'s (2022) study (mentioned in Section 4.1) investigated the effect of various input modes. They found that while RO, viewing, and viewing with captions all resulted in similar amounts of collocation gains at the form recognition level, RWL did not result in collocation learning. This study also shows that academic input contains a variety of collocations (e.g., non-specialised items, technical collocations, or academic collocations).

Such variation strengthens the ecological validity of looking at academic input as a source of collocation learning because L2 learners are likely to come across collocations with varying degrees of specialisation. If collocation learning from academic texts can be demonstrated, it opens up a large potential source of learning for both advanced learners as well as university-level EFL students who are already reading and engaging with these type of texts as part of their studies.

4.2. Aims and research questions

The first purpose of this experiment is to investigate whether reading mode (RO vs RWL) affects the degree to which collocations are incidentally learnt from an academic text. The second goal is to examine some of learner- and item-related factors that affect the learning of collocations. Therefore, the present experiment asks:

1. To what extent do adult L2 learners incidentally learn collocations through RO and RWL to an academic text?

2. Do learner-related (prior vocabulary knowledge) and item-related factors (congruency, compositionality, corpus frequency) influence incidental collocation learning?

4.3. Methods

4.3.1. Participants

Sixty-eight participants (44 female, 24 male) were recruited from a Spanish University. They were all native speakers of Spanish[8] and were between the ages of 18 and 24 (M=21.49, SD=2.11). Their English proficiency was at least upper-intermediate or at B2 level in the Common European Framework of Reference (CEFR), as determined by an external official certificate they provided before being admitted at university. All participants had formally studied English for at least ten years (M = 13.32, SD=3.00). To evaluate their prior vocabulary knowledge, I administered three levels of the revised version of the Vocabulary Levels Test (VLT) (Schmitt, Schmitt, & Clapham, 2001). The 3,000, 5,000, and 10,000 word frequency bands were determined to be the most relevant for the purpose of this experiment. The 2K level will have been too easy while the academic band might have been redundant or uninformative. The 10K level was included because a number of Latin root words in the 3K and 5K levels might have been familiar to L1-Spanish natives, possibly skewing the results. Each of the three sections assessed (3K, 5K and 10K) was scored counting the number of correct answers out of the possible 30. Participants scored an average of 26.55 out of 30 on the 3K level *(SD*=3.35). The average score was 26.08 out of 30 (SD=3.69) on the 5K level, and 20.73 out of 30 (SD=4.89) on the 10K level. The mean VLT score (for all three levels administered, MAX = 90) was 74.32 (SD = 7.94). Participants self-rated their English listening, reading, writing, and speaking skills at pretest. All participants rated[9] their reading (MAX=10, MIN=6, M=8.5, SD=0.960) and listening skills (MAX=10, MIN=6, M =7.14, SD =0.91) above 6 in a 10-point scale. Participants were then randomly assigned to either the RO (N=34) or the RWL condition (N =34). There was no statistically significant difference between the prior vocabulary knowledge of participants in the two conditions determined by a Wilcoxon Rank Sum Test (W=736.5, p=0.28). Table 4 includes the means and standard deviations of the VLT composite score

---

[8] For all the experiments presented in this dissertation, only peninsular Spanish speakers were recruited to avoid the variability found in different Spanish varieties (Lozano, 2016; Montrul & Rodríguez-Louro, 2006).

[9] I am only reporting listening and reading because as there are the relevant skills for the experiment.

as well as gender ratio for each condition. Participants who completed all components of the experiment were entered into a raffle where they had the opportunity to win an e-gift voucher in compensation for their participation.

Table 4. *VLT mean score and gender ratio per condition*

	VLT composite score M (SD)	Gender ratio	
		Female	Male
Reading-only	71.02 (6.51)	24	10
Reading-while-listening	71.47 (7.69)	20	14

4.3.2. Design

The purpose of Experiment 1 was to determine the extent to which reading mode affects the degree to which collocations are learnt incidentally. Reading mode (RO vs RWL) was manipulated between subjects. Three types of outcome measure (form recall, form recognition, meaning recognition) were collected, to obtain a finer-grained picture of the degrees of learning (Webb, 2005). These were measured at two delays (immediate posttest, delayed posttest), resulting in a total of six outcome variables.

4.3.3. Reading materials

A slightly modified version of an IELTS academic text (*Sweet and Other Sweeteners*, length: 1,001 running words) was used (see Appendix 1A). The text was selected against other two academic texts which, based on our intuitions and prior conversations with the teachers of the participants, we thought would be interesting for our participants (as later confirmed in the retrospective questionnaire). The topic of the reading was more research-focused compared to the topics of non-academic texts, which tend to cover matters of general interest. The content of the reading was descriptive and contained factual information.

The text was modified with the assistance of four English native speakers who were experienced EFL teachers. A frequency analysis using LexTutor was performed to assess the vocabulary frequency profile of the modified text. Of the words in the text, 96.2% belonged to the first 4,000 most frequent words in the *British National Corpus* (BNC) (determined by LexTutor, Cobb, n.d.).

Of the remaining words, 32 (2.90%) were from the 5K-9K frequency bands, 2 from the K-11, 1 from the K-12, and 2 from the K-14 band. The analysis of the words in the text showed that 5.5% was specialised vocabulary (e.g., "papilla", "olfactory") and 0.44% was low frequency non-specialised vocabulary (e.g., "adjust", "substitute"). A distinction between academic and specialised vocabulary was made following Nation (2013, 2016). The text included both academic words, i.e., words that occur in a wide range of academic disciplines (e.g., "exploited", "evidence"), and technical words, i.e., words that frequently occur in a specific discipline or subject area (e.g., "glucose", "calories"). A closer inspection of the text revealed that several low-frequency off-list words would not be problematic for the students because they were either cognates with Spanish (e.g., "predispose" - *predisponer*) or words which had Latin roots (e.g., "fiber" - *fibra*).

The text was assessed to be at the appropriate proficiency level for the learners, and its vocabulary profile indicated an adequate percentage of word knowledge for our participant group.

The text was divided into 10 screens of similar length using Arial size 18 as the font. The mean number of running words per screen was 97 (*MAX*=103, *MIN*=91). It was not possible for participants to go back to previous screens. Reading time was controlled by automatic advancement. Participants' reading times were monitored and recorded to ensure that treatment groups spend approximately the same amount of time on a screen and to minimise any confounding variables that might influence the results of the study. Timing was piloted with a small group of participants in the piloting to refine the experimental design and adjust the screens. No significant differences were found in the reading times in the actual experiment. The audio version of the text was recorded by a native speaker of British English in a lab environment to ensure good sound quality. The audio was 8 minutes long with a speech rate delivered at 130 words per minute (wpm), which was considered appropriate for the proficiency level of the learners.

4.3.4. Collocations

The procedure for identifying collocations was operationalised in two steps. An example of the parameters used for extracting the collocations for the experiment are presented in Figure 2. Firstly, collocations were identified using corpus-based methods of extraction (frequency-based approach). The collocational strength in this experiment is indicated by the t-score. Following a phraseological

approach, the identified collocations were then classified into distinct types (e.g., Henriksen, 2013; Li & Lu, 2011; Peters, 2016) as explained below.

Figure 2. Search parameters for collocation retrieval.

Fourteen target collocations were finally selected and embedded in a single text. The items included were noun-noun and adjective-noun lexical collocations. The experiment was limited to adjacent lexical collocations (Bahns, 1993; Henriksen, 2013) where the node word of the collocation and collocate occurred next to each other. Ten collocations were found in the authentic text and 4 were inserted in the modified version. Raw corpus frequencies were extracted from the COCA. Importantly, as is customary in collocation learning studies, the frequency of the individual collocations constituents was controlled by means of association measures. In the present experiment, the t-score was used to consider the expected frequency of a collocation controlling for individual word frequency, which helps identifying genuine collocational patterns that are not solely driven by the frequency of individual words. All target items had a high t-score (all above 6.00). It should be noted, however, that there are numerous methods for determining individual word frequency, as demonstrated by earlier research using online measures (see Carroll & Conklin, 2014; Jiang & Siyanova-Chanturia, 2020). The collocations occurred between 1 and 7 times in the text, although most of the items (11 out of 14) appeared only once. Three items appeared more than once (2, 4, and 7 times, respectively). The items varied along two dimensions: congruency and compositionality (see Table 5). The items were first piloted with 12 EFL learners with a similar language level and background to that of the participants in the experiment, using

a superset of the eventual target items. Two items were found to be relatively well known ("eating habits" and "junk food"), and were therefore excluded.

For congruency, Nesselhauf's (2003) operationalisation of congruency was used. A congruent collocation can be translated literally into the L2 (e.g., "weight loss" - *pérdida de peso*), whereas a non-congruent collocation cannot be translated word-for-word (e.g., "taste buds" - *papilas gustativas*). To yield a reliable congruency status, five experienced Spanish-English translators were consulted. They all had a Master's degree in Translation Studies, and 7-9 years of experience as translators. They were asked to label the collocations' congruency: incongruent (0) or congruent (1). Only items which received the same rating were included in the study. Half of the target collocations (7) were rated as congruent collocations and half as incongruent (7). What should be remarked though, is the overlap between congruency and transparency in studies researching collocations (see Conklin & Carroll, 2018). Many of the collocations that are often labelled as congruent items could be also considered more transparent. This is further problematised by the fact that transparency is not considered as a factor, although it is often confounded with the congruent/incongruent classification. However, it is important to note that congruency and transparency are not entirely synonymous, and there can be collocations that are congruent but not transparent or vice versa. For instance, the collocation "morning sickness" is highly congruent because "sickness" often combines with "morning", but its meaning is not entirely transparent from the individual words. Therefore, while there might be some overlap between these two concepts, they can represent distinct dimensions in collocation studies. Disentangling effects of L1-L2 congruency from transparency is not straightforward. For this reason, the present experiment includes semantic compositionality as a factor.

With regards to compositionality, a norming study was conducted to determine the degree of compositionality of the target collocations. A transparency task adapted from Macis and Schmitt (2017) was distributed to 25 raters, all native speakers of Spanish with an advanced level of English and university degrees. The raters received an explanation of collocations and compositionality and were asked to rate the meaning of the collocations on a scale from 1 to 4, with 1 being very difficult to guess (very opaque/non-compositional) and 4 being very easy to guess (very transparent/compositional). Ratings were normalised to minimise any spurious effects of raters tending to use different portions of the scale. The collocations, presented in Table 5, showed diversity in terms of their transparency (*MIN*=0.52, *MAX*=1, *SD*=0.20). The semantic transparency task is available in Appendix 1A.

Table 5. List of target collocations and their properties.

Collocation	Congruency	Compositionality	t-score	COCA frequency
morning sickness	congruent	0.57	7.89	482
chocolate bar	congruent	0.93	8.93	432
marketing ploy	congruent	0.57	7.84	162
sugar cane	congruent	0.99	9.92	861
artificial sweeteners	congruent	1	8.46	236
saturated fat	congruent	1	8.5	2541
hearty meal	congruent	1	9.91	88
withdrawal symptoms	incongruent	0.51	10.06	354
taste buds	incongruent	0.59	11.02	765
side effects	incongruent	0.92	9.31	7143
gold standard	incongruent	0.52	7.77	2339
soft drinks	incongruent	0.6	8.36	1110
fond memories	incongruent	0.71	9.38	924
sugar cravings	incongruent	0.92	7.48	37

4.3.5. Measures

4.3.5.1. Pretests

A number of learner-related factors were assessed for their contribution to learning success. This data was gathered during pretest assessments two weeks before the treatment. In the first session, participants signed consent forms, filled out a background questionnaire, and completed a vocabulary test as well as a test on the target collocations.

4.3.5.1.1. Background questionnaire

A slightly modified version of the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld, & Kaushanskaya, 2007) was used to gather language background and

demographic information. The LEAP-Q is a validated questionnaire tool that is widely used in a variety of disciplines (Neuroscience, Linguistics, Education) (Kaushanskaya et al., 2007). The information collected from participants included: years of formal English instruction, self-reported English language proficiency level in speaking, listening, reading, and writing, use of English outside the language classroom, official English language qualifications, time spent living in an English-speaking country, and other details that have contributed to their learning, e.g., private tuition, reading books, watching English-language TV programmes. The age, gender, and educational levels of participants were also collected.

4.3.5.1.2. Vocabulary test

The revised version of the Vocabulary Levels Test[10] (VLT) (Schmitt, Schmitt, & Clapham, 2001) was used to assess the participants' vocabulary size as an approximate objective measure of their language proficiency. The VLT total score was computed adding up the scores from the three individual sections (the 3K , 5K, and 10K word frequency bands) and thus it was scored out of 90 (30 words x 3 levels). While this test was not designed to sum up all the scores in the different levels, I chose this approach to obtain an approximate measure of the participant's general vocabulary proficiency while acknowledging that this score does not fully represent the proficiency level of participants. Previous studies have followed this procedure (Sonbul & Schmitt, 2013; González-Fernández, 2018).

Participants scored an average of 26.55 out of 30 (SD=3.35) on the 3K level, 26.08 out of 30 (SD=3.69) on the 5K level, and 20.73 out of 30 (SD=4.89) on the 10K level. The VLT average score out of 90 was 74.32 (SD=7.94).

4.3.5.1.3. Collocations pretest

A collocations pretest was created to assess whether a participant already knew each target collocation. A matching task tapping into form recognition was adapted from Webb and Chang (2020). Participants were asked to match words from Column A with words from Column B to form two-word combinations, by dragging and dropping the words from Column A to Column B (see Figure 3). All items appear on the screen at the same time. The node word of each collocation was randomly put in one list

---

[10] The VLT is freely available at https://www.lextutor.ca/tests/.

(Column A) and the other words in the sequences in the second list (Column B). Participants were instructed to select "I don't know" if they did not know any of the word combinations to prevent guessing. Where participants were able to form the two-word combinations, they were asked to translate them into Spanish (meaning recall). Randomisation of the component words put in the boxes increased the difficulty of the test. The test contained 12 distracters that were added to minimise the chances of successful guessing and avoid drawing attention to the target items. All distracters were semantically related to target items to prevent participants from ruling out options on semantic grounds (Greidanus & Nienhius, 2001). Data from corpora was used to ensure that distracters were adequate. They consisted of words that are more frequent than the most frequent 4,000-word level in Nation's (2012) BNC/COCA lists, and thus were likely to be known by all the participants. Five experienced EFL teachers confirmed the suitability of the distracters in the piloting phase. This test was piloted with native and nonnative speakers who did not take part in the experiment and found to work as intended.

Figure 3. Screenshot of the matching task.

4.3.5.2. Immediate posttests

In the second session, which took place two weeks after the first session, all participants completed the learning block, where they read the text in one of the reading mode conditions (RO or RWL) and took four unannounced immediate posttests: a reading comprehension test, and three tests measuring different aspects of collocation knowledge: a form recall test, a form recognition test, and a meaning recognition test, in that order.

The reading comprehension test, which consisted of 8 true/false statements based on the text, was used to assess participants' understanding of the content. This test was followed by the three vocabulary tests. To minimise a potential test effect, the form recall test was administered first, followed by the form recognition test, and finally the meaning recognition test. The distracters were kept in the posttests to minimise testing effects and maintain the integrity and validity of the assessment, i.e., it would have been easier to complete the tasks if the posttests had only included the target items.

In the form recall test, participants were asked to provide the form of the collocation in the L2 based on a Spanish translation given (in the example below, the target collocation is "taste buds"). They were provided with the first letter of each component word to avoid elicitation of other word pairs.

Example of the form recall test.

papilas gustativas

t_____ b_____

The form recognition test was identical to the one used at pretest, although the items were randomised and participants were not instructed to provide the Spanish equivalents at this stage. Using identical test formats is recommended for studies with a pretest-posttest design so that learning gains are not confounded by the use of different test formats (Nation & Webb, 2011).

The meaning recognition task measured the participants' ability to recognise the correct meaning of the given target collocation among five potential Spanish translations: the correct meaning, three fillers, and one "I'm not sure" option. Five experienced Spanish-English translators were consulted to find appropriate fillers for this test.

Example of the meaning recognition test.

side effects

   a) *efectos secundarios*

   b) *efectos nocivos*

   c) *efectos colaterales*

   d) *efectos adversos*

   e) I'm not sure

4.3.5.4. Delayed posttests

In the third and final session, which took place two weeks after the previous session, participants completed the delayed posttests, which were identical to the immediate posttests. In addition, participants completed a retrospective protocol questionnaire. The questionnaire gathered qualitative information about whether learners were familiar with the topic of the text, their perception of difficulty, engagement with the text, perception of learning, and usefulness of the audio support for the participants in the RWL group (see Appendix 1A). The questions were answered using different rating scales. Participants were also given the opportunity to leave specific feedback if they wanted to.

To reduce potential testing effects and minimise the chances of the learners remembering the test items for the delayed posttests, distracters were included in all tests, and target items and distracters were randomised across pretest, immediate posttest, and delayed posttest. With three tests (form recall, form recognition, meaning recognition) at two delays (immediate posttest, delayed posttest), the data consisted of 6 outcome variables. The final dataset comprised 5,712 observations, from 68 participants to 14 items, and 6 outcome variables.

4.3.6. Procedure

All components of the experiment were created and distributed via SoSci Survey (Leiner, 2019 Version 3.1.0), a web-based experiment presentation interface.

The data were collected in three sessions separated by two weeks each (see Table 6). In the first session, three tasks were administered: a background questionnaire, a vocabulary test, and a collocations pretest. In the second session, participants completed the treatment where they read the text in one of the reading conditions and completed four unannounced tests: a postreading

comprehension test, and three immediate collocations posttests. In the last session, participants took the same collocations posttests as in the immediate posttest, and completed a retrospective protocol questionnaire.

Participants were monitored in a computer-based environment. While total control over the participants was not possible due to the online administration of the study, timing checks were employed as a proxy to detect inattention and monitor participants' performance.

Table 6. Data collection procedure.

Week	Procedure
1	Informed consent form
	Background questionnaire
	Vocabulary test
	Collocations pretest
3	Treatment
	Reading comprehension test
	Immediate form recall test
	Immediate form recognition test
	Immediate meaning recognition test
5	Delayed form recall test
	Delayed form recognition test
	Delayed meaning recognition test
	Retrospective questionnaire

4.3.7. Data processing and model fitting

4.3.7.1. Exclusion criteria

The responses to the postreading questionnaire were checked for all the participants. The threshold of adequate comprehension was set at answering 80% of the questions correctly. As the questions were quite easy, a lower score would indicate that the participant had not paid adequate attention to the text. No participants were excluded as no participant had more than 1 incorrect response

out of 8. Additionally, none of the participants reported having trouble understanding the text in the retrospective questionnaire.

4.3.7.2. Scoring responses

The outcome variables were binary (correct response=1, incorrect response=0). Answers in the pretest were considered correct only if learners were able to both correctly match the collocations and supply an accurate L1 translation or definition of the items. This decision was made to prevent blind guessing following similar studies (e.g., Webb & Chang, 2020). Answers in the form recall tests that contained minor spelling mistakes were scored as correct as long as the response was comprehensible (e.g., "*artifitial sweeteners"). This decision was made following previous studies (Puimège & Peters, 2019; Sonbul & Schmitt, 2013).

4.3.7.3. Model fitting

Before analysing the data, the learner-related predictors (preknowledge of collocation, prior vocabulary knowledge) as well as collocation-related factors (congruency, compositionality, corpus frequency) were assessed for correlations among variables. To reduce collinearity, prior vocabulary knowledge (i.e., VLT score) was centred on the grand mean. Preknowledge of collocations and prior vocabulary knowledge were found to be correlated, as were compositionality and congruency. To avoid collinearity in the models, preknowledge of collocations[11] and compositionality were removed from the analysis (see Appendix 1B).

Unaggregated data were fitted to logistic mixed-effects regression models (Baayen et al., 2008), using the lme4 package (Bates et al., 2015) on R version 4.0.1 (R Core Team, 2020). Six separate models were fitted for each outcome measure (form recall, form recognition, meaning recognition) and each delay (posttest, delayed posttest). In addition to reading mode, each full model included fixed effects of all potential explanatory variables: number of occurrence of the items, properties of the collocations (congruency), two-way interactions with reading mode, where the interaction would be

---

[11]. The VLT was kept in the model because it is a standardised vocabulary test. unlike the preknowledge of collocations pretest. However, pretest scores were computed to estimate the amount of collocational knowledge learners had at the time of the pretest. The mean average (M=0.24) was rather low.

interpretable, and learner-related variables (prior vocabulary knowledge). Binary variables were centred, and continuous variables were standardised. To be able to compare the magnitude of effects across the form models and separately across the meaning models, all the form models used the same set of predictors, as did the meaning models. This meant that if a predictor was significant in one form (or meaning) models, it was retained in the others, regardless of their significance in those models. This decision was made based on previous research, which shows that there is a consistent order of difficulty in learning different word knowledge aspects (see Laufer & Goldstein, 2004).

Models for each of the outcome measures were selected using the following procedure. The fixed effects structure was determined by starting from the full model and using model comparison (likelihood ratio test) to remove terms that did not significantly improve model fit (Manning, 2007), unless the predictor was significant in another model of the same type (form, meaning). After the fixed effects were determined, random intercepts and slopes were included for participants and items. The maximal random effects structure that was supported by the data was used (Barr, Levy, Scheepers, & Tily, 2013).

4.4. Results

4.4.1. Descriptive statistics

Average absolute gains were calculated for each reading mode condition and for each outcome measure (see Table 7). There were no pronounced differences between the RO and RWL conditions, suggesting that the reading mode made no difference for incidentally learning collocations. The relatively high standard deviations on the posttests indicate a great deal of variation in the scores among participants. This is common within studies of incidental vocabulary learning (e.g., Webb & Chang, 2020). The higher mean scores for the recognition measures compared to the recall measures suggest that learners recognise the collocation's meaning first (passive recall) and then its form (active recognition) before they are able to recall it (active recall). On all measures, performance improved from immediate to delayed posttest, which point to testing effects. I will return to this issue in the Discussion.

Table 7. Absolute mean gains by condition and outcome measure.

		Reading only M (SD)	Reading-while-listening M (SD)
Form Recall	Immediate posttest	3.76 (2.55)	3.44 (2.12)
	Delayed posttest	6.76 (2.00)	6.15 (2.23)
Form Recognition	Immediate posttest	7.35 (2.44)	7.18 (2.60)
	Delayed posttest	9.09 (2.08)	9.59 (1.73)
Meaning Recognition	Immediate posttest	10.23 (3.03)	10.24 (2.85)
	Delayed posttest	10.87 (2.56)	11.03 (2.35)

Note: Maximum: 14. Gains are measured relative to performance at pretest.

4.4.2. Models

This section presents immediate and delayed posttest models for each type of outcome measure. The full models are available in Appendix 1B.

4.4.2.1. Form-related models

As explained elsewhere in this thesis, knowledge of form is a key aspect in vocabulary knowledge (Schmitt & Schmitt, 2000) and it was assessed at two levels of sensitivity: recall and recognition. Form recall, also defined as active recall (Laufer & Goldstein, 2004), refers to the ability to recall or retrieve the phonological (spoken) and/or orthographical (written form of a lexical item. Form recall of the form-meaning link aspect of vocabulary is considered to be the deepest level of word knowledge (Sonbul & Schmitt, 2010) and the most difficult to acquire (Laufer & Goldstein, 2004). By contrast, form recognition or active recognition refers to the ability to recognise (perceive) the phonological and/or orthographical form of a word.

As Table 8 shows, vocabulary size and congruency are reliable predictors at immediate posttest. A larger vocabulary size increased the likelihood of a correct form recall response, as did having a congruent item as opposed to an incongruent collocation. The model also revealed a marginal positive effect of repetitions.

Table 8. Summary of fixed effects for immediate posttest form recall.

Model:

Response ~   ReadingMode + VocabSize + Congruency + Repetitions

+ VocabSize:Congruency + Congruency:ReadingMode

+ (1+VocabSize+ReadingMode|ParticipantID) + (1+Congruency|ItemID)

Predictor	Estimate	SE	z value	Pr (>|z|)
Intercept	-0.25	0.23	-1.06	0.29
ReadingMode	-0.09	0.15	-0.67	0.51
VocabSize	0.57	0.17	3.47	0.00
Congruency	0.56	0.21	2.67	0.00
Repetition	0.28	0.16	1.74	0.08
VocabSize:Congruency	-0.10	0.08	-1.19	0.23
ReadingMode:Congruency	-0.00	0.08	-0.03	0.97

Looking at the same measure at delayed posttest, the vocabulary size and congruency effects remained reliable (Table 9). An interaction between vocabulary size and congruency also emerged (slope of VocabSize for Incongruent: β=0.55, *SE*=0.15, *p*=0.00; slope of VocabSize for Congruent: β=1.06, *SE*=0.24, *p*=0.00). As shown in Figure 4, the interaction arises because, while vocabulary size and congruency both have positive effects on delayed form recall, for learners with higher vocabulary sizes, congruency has a weaker effect as those learners are already performing near ceiling (β=-0.10, *p*=0.23), in other words, those learners are also performing well on incongruent collocations.

Table 9. Summary of fixed effects for delayed posttest form recall.

Model:

Response ~   ReadingMode + VocabSize + Congruency + Repetitions

+ VocabSize:Congruency + Congruency:ReadingMode

+ (1+VocabSize+ReadingMode|ParticipantID) + (1+Congruency|ItemID)

Predictor	Estimate	SE	z value	Pr (>|z|)
Intercept	2.30	0.45	5.13	0.00
ReadingMode	-0.04	0.15	-0.31	0.76
VocabSize	0.80	0.16	4.98	0.00
Congruency	1.31	0.42	3.14	0.00
Repetition	0.69	0.53	1.31	0.19
VocabSize:Congruency	0.26	0.12	2.15	0.03
ReadingMode:Congruency	0.21	0.12	1.82	0.06
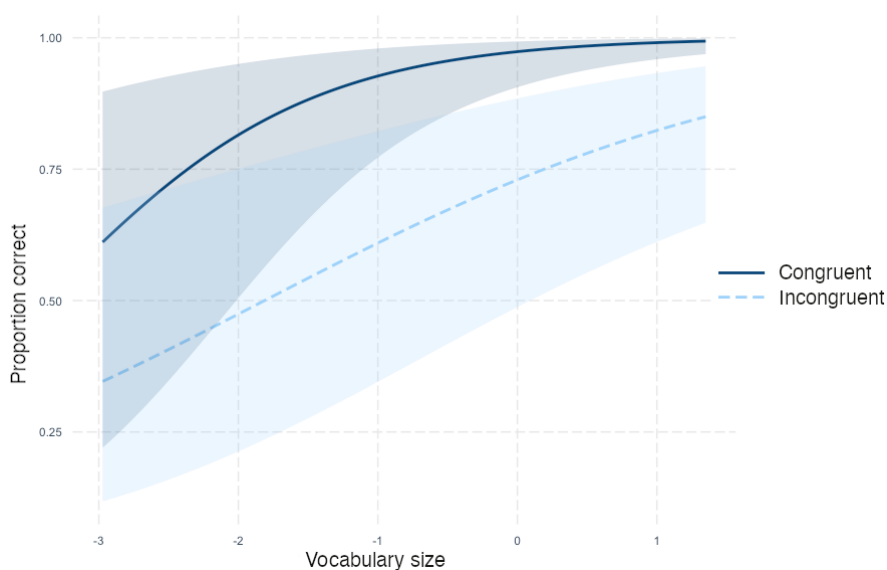
Figure 4. Interaction between vocabulary and congruency (form recall, delayed posttest).

In the immediate posttest model (Table 10), vocabulary size was the only reliable predictor of form recognition, with no main effect of congruency, suggesting that the individual vocabulary knowledge of learners played a more prominent role in form recognition than intrinsic properties of the collocations. Reading mode did not affect form recognition either (p=0.67).

Table 10. Summary of fixed effects for immediate posttest form recognition.

Model:

Response ~   ReadingMode + VocabSize + Congruency

+ Congruency:ReadingMode

+ (1+Congruency|ParticipantID) + (1+Congruency|ItemID)

Predictor	Estimate	SE	z value	Pr (>|z|)
Intercept	1.92	0.29	6.65	0.00
ReadingMode	0.06	0.16	0.43	0.67
VocabSize	0.97	0.17	5.82	0.00
Congruency	0.27	0.26	1.02	0.31
ReadingMode:Congruency	0.00	0.11	0.03	0.97

As Table 11 shows, the delayed posttest model included only a main effect of congruency (p=0.00). This effect emerges at delayed posttest, while the effect of prior vocabulary knowledge is no longer significant. We will return to this pattern of effects in the Discussion.

Table 11. Summary of fixed effects for delayed posttest form recognition.

Model:

Response ~   ReadingMode + VocabSize + Congruency + Congruency:ReadingMode

+ (1+Congruency|ParticipantID) + (1+Congruency|ItemID)

Predictor	Estimate	SE	z value	Pr (>|z|)
Intercept	6.59	1.38	4.77	0.00
ReadingMode	0.44	0.58	0.76	0.45
VocabSize	0.88	0.54	1.63	0.10
Congruency	0.72	0.25	2.94	0.00
ReadingMode:Congruency	-0.57	0.24	-2.34	0.01

The model also revealed a negative interaction between reading mode and congruency: RWL improved delayed form recognition when compared to RO, but selectively, for incongruent collocations only (slope of Congruency for RO: β=2.86, SE=1.42, p<0.05; slope of Congruency for RWL: β=1.38,

SE=1.20, p=0.25). The fact that this interaction emerges only in the delayed posttest suggests that, while reading mode initially seems to have no impact on learning, factors may have different timecourses, with some only yielding detectable effects after a sufficiently long delay after the treatment.

4.4.2.2. Meaning-related models

Because words are the first and foremost units of meaning, it was decided to assess the form-meaning link at the level of meaning recognition. Meaning recognition refers to the ability to recognise the form of a lexical item and retrieve its meaning or meanings. Meaning recognition is the easiest form-meaning link aspect to acquire as it only requires passive recognition (Laufer & Goldstein, 2004). As Table 12 shows, the final model indicated that congruence significantly affected meaning recognition at immediate posttest, with congruent items being recognised better than incongruent ones (β=2.79, z=0.25, p=0.00). Repetition also contributed to the model. Thus, the higher the frequency of an item, the better its meaning is recognised. Conversely, vocabulary size came out as a non-significant predictor of this type of knowledge. Thus, the final model includes congruency and repetition as predictors of meaning recognition.

Table 12. Summary of fixed effects for meaning recognition at immediate posttest.

Model:				
Response ~	ReadingMode + VocabSize + Congruency + Repetition			
	+(1+VocabSize+Repetition|ParticipantID)+(1+VocabSize+ Repetition|ItemID)			
Fixed effects	**Estimate**	**SE**	**z value**	**Pr (>|z|)**
Intercept	-0.53	0.89	-0.56	0.57
ReadingMode	-0.19	0.27	-0.69	0.48
VocabSize	0.07	0.17	0.45	0.64
Congruency	2.79	0.70	3.96	0.00
Repetition	1.68	0.68	2.45	0.01

The best-fitting model for the meaning recognition component at delayed posttest is presented in Table 13. As in the immediate posttest model, congruency proved to be the main significant predictor of meaning recognition at delayed posttest. On the other hand, the effect of repetition was no longer

significant in the delayed posttests. This interesting result will be discussed in the next section. Again, no evidence of vocabulary size was found in the delayed posttests. Thus, the best model fit included only L1 congruency as direct effect (*p*=0.00).

Table 13. Summary of fixed effects for meaning recognition at delayed posttest.

Model:				
Response ~	ReadingMode + VocabSize + Congruency + Repetition			
	+ (1+VocabSize+Repetition|ParticipantID)+(1+VocabSize+RepetitionItemID)			

Predictor	**Estimate**	**SE**	**z value**	**Pr (>|z|)**
Intercept	-0.63	1.97	-0.32	0.74
ReadingMode	-0.31	0.39	-0.79	0.42
VocabSize	-0.13	0.24	-0.70	0.47
Congruency	4.29	1.15	3.72	0.00
Repetition	2.82	1.85	1.51	0.12

4.4.2.3. Summary of the models

Table 14 summarises the results of the models for form recall and recognition, at immediate and delayed posttests.

Table 14. Summary of model results.

Predictor	**Form recall**		**Form recognition**		**Meaning recognition**	
	Immediate	Delayed	Immediate	Delayed	Immediate	Delayed
ReadingMode	−	−	−	−	−	−
VocabSize	+	+	+	−	−	−
Congruency	+	+	−	+	+	+
Repetition	marginal	−	N/A	N/A	+	−
VocabSize:Congruency	−	+	N/A	N/A	N/A	N/A
ReadingMode:Congruency	−	marginal	−	+	N/A	N/A

Plus signs (+) indicate a significant effect, minus signs (-) show no significant effect.

The word marginal indicates that the effect was significant at .1 level.

4.5. Discussion

4.5.1. The (non-)effect of reading mode

To what extent do adult L2 learners incidentally learn collocations through RO and RWL to an academic text? The non-significance of reading mode in the models did not reveal a significant effect of reading mode, indicating that there was not a reliable difference between RO and RWL. The results of the current experiment are at odds with those of Vu and Peters (2021) and Webb and Chang (2020), who found that RWL was more effective than RO, but is consistent with Dang et al. (2022), who found that RWL to an academic lecture did not contribute to incidental collocation learning. One possible explanation for the lack of a comparable effect in this experiment and in Vu and Peters (2021) and Webb and Chang (2020) may be due to the nature of the intervention and the materials used. First, Vu and Peters (2021) conducted a longitudinal study, as opposed to Webb and Chang (2020), Dang et al. (2022), and this experiment, which were one-off interventions. Furthermore, the participants in Webb and Chang's (2020) study experienced greater quantity of learning input. In that study, the treatment was conducted in six 50-minute classes over a 3-week period, whereas the present study was conducted in a computer-lab environment in three sessions over a 5-week period. The auditory input in the present experiment was 8 minutes long, compared to 156 minutes in Webb and Chang (2020).

Secondly, the materials used by Vu and Peters (2021) and Webb and Chang (2020) were (upper)-intermediate-level graded readers of approximately 17,000-19,000 running words. While reading narrative stories may have utility in the EFL classroom, studies that use such materials do not directly assess how advanced EFL learners incidentally learn vocabulary from reading more challenging texts (Vilkaitė, 2017). This was a key motivation behind my decision to use academic texts, focusing on the characteristics of university-level EFL learners, their proficiency levels, and their educational background. How might we reconcile the inefficacy of RWL over RO for advanced learners reading academic texts, with the benefits observed for fictional reading materials?

The findings suggest that academic listening might be particularly difficult due to the intensity of processing more complex texts, which is in line with Dang et al. (2022). The informational density of this type of input might detrimentally affect the learners' ability to integrate information from different channels. Academic input can also be more challenging to understand if L2 learners lack knowledge of the specialised vocabulary that occurs in a given academic subject (Coxhead; 2020; Nation, 2016). Viewed together with the findings of Web and Chang (2000) and Vu and Peters (2021), our findings

suggest a more nuanced landscape for L2 vocabulary learning at various levels, and with varied degrees of input difficulty. While additional information in the input (as in RWL) might benefit learners by providing additional cues (e.g., prosodic cues to constituent groupings) when a text is sufficiently easy, it is conceivable that, when faced with a more challenging text, learners do not benefit from additional input information because they are overwhelmed by the need to process two streams of information simultaneously.

The possibility of a complexity trade-off is supported by learners' responses in the retrospective questionnaire. Several learners in the RWL group reported to have found this mode distracting. Our audio was delivered at 130 wpm compared to 120 wpm in Webb and Chang (2020). The lower speech rate in Webb and Chang's study may have mitigated any extra processing cost associated with processing text and speech simultaneously, while the faster speech rate in our study may have increased processing complexity. We hypothesise that the real-time processing of RWL may have diminished its potential benefit relative to RO, where learners are able to revisit the text at their own pace (e.g., Vidal, 2011). This aligns well with the Cognitive Load Theory (Sweller, et al., 1998) and the Theory of the Limited Processing Capacity (Ellis, 2011; VanPatten, 1996).

From the perspective of the Cognitive Load Theory, dual processing (reading and listening) may have increased the cognitive load, which might have been too overwhelming for some participants, which may have been detrimental to learning. Listening to audio also involves keeping pace with the narrator, which can affect the cognitive load of the task if the learner's own reading pace is (much) slower. Some participants complained about the pace at which the narrator read (130 wpm), which was at odds with their normal reading speed in English. The general learner perception in the RWL group was that listening negatively affected concentration, which may have affected engagement and overall performance. Some participants reported having paid more attention to the audio. It is likely that attention was allocated to comprehending the auditory input only, particularly if listening was perceived as demanding. These split-attention effects have been reported in previous studies (e.g., Vu & Peters, 2020). Future studies could consider investigating whether increased cognitive load is caused by dual-task complexity.

While the additional cognitive burden associated with RWL may have counteracted any benefit from the from additional audio "support", it is important to note that the reading task in this experiment was not too challenging content-wise so that it would prevent successful processing and comprehension

of the material. This is confirmed by the postreading comprehension questions: scores did not differ by reading condition, and learners in the RWL condition were not more likely to be excluded due to low comprehension scores than those in the RO condition. As indicated by responses to the retrospective protocol questionnaire, the participants were accustomed to reading on-screen. They also had prior online learning experience during the most restrictive times of the Spanish lockdown during the pandemic. The text was short enough (1,001 running words) that it was unlikely to cause major problems in terms of fatigue effects or boredom, and the required reading time was approximately 10 minutes (considering that participants were asked to read each screen only once). Importantly, the participants of this experiment were likely to be highly proficient readers, since they were undergraduate students. This finding, thus, may not generalise to other L2 populations, such as immigrant labourers. It would be interesting to investigate if the L2 reading proficiency of this particular sample is generally higher than the L2 aural comprehension level, especially in light of the unique characteristics of EFL environments (compared to ESL contexts, for example).

The results also revealed an interaction of reading mode and congruency at delayed form recognition. For congruent collocations, reading mode did not reliably impact learning, while RWL improved learning of incongruent collocations relative to RO. Reading mode may therefore have selective effects on collocation learning, depending on several factors, including testing delay and aspects of learning discussed above (learner proficiency, type of L2 input, quantity of input, cognitive load during treatment).

4.5.2. Absolute gains

This experiment provides further evidence suggesting that academic input is a valuable source for incidental L2 collocational learning, at different levels. On average, the number of collocations learnt to the form recognition level in the immediate posttests by the RO group was 7.35 out of 14 collocations and 7.17 by the RWL group. The learning gains of the groups were thus fairly similar. However, these gains seem substantially greater than the average gains by the RO group in Dang et al. (2022) and Vilkaitė (2017) - 4.88 out of 19 collocations and 1.9 out of 15 collocations, respectively. One reason why greater recognition gains were found in this study could be that the text was shorter compared to the lecture in Dang et al. (5,674 running words) and the texts in Vilkaitė (between 2,012 and 2,230 words). Further, Vilkaitė examined verb-noun combinations (e.g., "pose a danger"), which may be more

challenging due to the abstract and polysemous nature of verbs (Altenberg & Granger, 2001). Boers et al. (2014:48) argue this type of collocations are a source of difficulty because learners find difficult to select the appropriate verb as it is the noun "that usually carries most of the semantic weight of the unit" whereas the verb lacks semantic transparency (Garnier & Schmitt, 2016).

Regarding the form recall gains, students learnt on average 3.76 collocations in the RO group and 3.44 in the RWL group. The learning gains at the recall level were thus much lower. The superior gains in form recognition are in line with previous research (Pellicer-Sánchez, 2017; Peters, 2016; Szudarski, 2012; Vilkaitė, 2017; Webb et al., 2013; Vu & Peters, 2020), which suggests that receptive knowledge aspects are more easily learnt than recall aspects. This study also provides further evidence that learners can acquire knowledge of collocations at a rate similar to that of other lexical components such as single words, from as little as a single exposure (Durrant & Schmitt, 2010). However, it should be noted that this study examined lexical collocations only. Investigation of other types of collocations in future studies would allow for a clearer understanding of the effectiveness of reading mode manipulations for the incidental acquisition of different knowledge aspects with different types of collocations.

The scores in the meaning recognition posttests showed the greatest improvement of collocation gains of all the vocabulary measures. However, this increase is likely due to a ceiling effect caused by the multiple-choice test employed to measure meaning recognition. This result was not completely unexpected as some of the translators consulted for the test design expressed concerns about the fillers provided as felicitous translations, though less natural sounding. Consequently, the results relating to meaning recognition should be treated with caution as they cannot be attributed solely to the treatment.

Taken together, the findings of the experiment indicate a relatively small amount of collocational learning, which supports the results of previous research investigating incidental learning of collocations (e.g., Dang et al., 2022; Pellicer-Sánchez, 2017; Vilkaitė, 2017; Webb et al., 2013; Webb & Chang, 2020). The gains recorded in this experiment occurred through encountering most collocations only once. It is likely that knowledge of the different aspects of collocations develops through continuous exposure, progressing along different stages until complete mastery is achieved.

4.5.3. Retention of collocations

Delayed posttests were included in the experiment to assess the extent to which immediate gains were retained over an interval of two weeks. On all measures, mean scores were higher at delayed testing. These findings should be interpreted cautiously for a variety of reasons that do not apply to the immediate posttest results. For one thing, multiple tests were used: while participants were encountering the posttests for the first time at immediate posttest, at delayed posttest, they have already seen the tests before. For example, target forms may have become more noticeable to participants after the immediate posttests. Testing effects - testing itself contributing to learning - might thus explain any increase in the scores recorded in the delayed posttests, which will have been the result of not just the treatment condition and/or prior vocabulary knowledge of the items. Testing effects have also been reported in previous studies (e.g., Peters & Webb, 2018; Webb et al., 2013, Webb & Chang, 2020). The delayed form recognition scores must be interpreted in light of the limitation pertaining the layout modification, rather than as direct evidence of incidental collocation learning. The results of the meaning recognition posttests should be taken very cautiously, given the aforementioned issues with the test format.

A related potential source of learning outside the factors considered is that participants may have looked up the items at home (Hulstijn, 2003). One limitation of this experiment is that we did not ask participants to report whether they had done so in the retrospective protocol questionnaire. Future studies should consider asking learners to report whether they looked up the items in-between testing sessions, as this is a relatively common methodological issue in L2 vocabulary studies. In sum, the results of the delayed posttests should be viewed with the above considerations in mind.

The delayed results show nonetheless that, regardless of the source of learning, L2 learners can establish collocation memory traces from the input they are exposed to, which is consistent with the findings of Durrant and Schmitt (2010). It should be noted that the present experiment assessed knowledge of form recall, form recognition, and meaning recognition of collocations at two posttest delays (immediate, delayed) whereas Durrant and Schmitt (2010) focused on the immediate retention of collocations embedded in sentences. Future studies can build upon the present research and explore other aspects of collocation knowledge (e.g., meaning recall).

In sum, while mean scores were higher at delayed testing on all measures, factors beyond those manipulated or assessed as fixed predictors of learning may have contributed to these elevated

scores. Nonetheless, the delayed posttest results broadly suggest that aspects of collocations (e.g congruency) may have differential effects based on a learner's proficiency level.

4.5.4. Learner- and item-related factors

A second objective of this experiment was to assess the contributions of learner- and item-related factors in incidental collocation learning.

Proficiency in this experiment was conceptualised as the learners' vocabulary size and measured by means of a VLT composite score. Results show that vocabulary size is a good predictor of collocational learning at the level of form recognition and form recall, indicating that a larger vocabulary size leads to better form recognition and recall of collocations. This points to a Matthew effect (e.g., Elgort, Perfetti, Rickles & Stafura, 2015; Horst, Cobb, & Meara, 1998; Stanovich, 1986) and echoes findings from previous studies that have shown that the learners' prior vocabulary knowledge is a significant predictor of incidental learning of collocations (e.g., Puimège & Peters, 2019, 2020; Vilkaitė, 2017; Vu & Peters, 2022a). However, an unexpected finding was that vocabulary size did not predict meaning recognition. This contradicts the assumption that the higher the level of proficiency of the learner, the better the meaning guessing skills (Nation, 2001). To put it another way, a larger vocabulary is more likely to allow learners to infer new meanings from context (Liu & Nation, 1985). The non-significance of vocabulary size on meaning recognition is most likely due to the test format, which may have resulted in a ceiling effect. Multiple-choice tests are also susceptible to guessing effects, which can lead to score overestimation (Gyllstad, Vilkaitė, & Schmitt, 2015). Consequently, as mentioned earlier, the results of meaning recognition must be interpreted with caution due to potential test effects. To avoid guessing effects, future research should consider testing meaning recognition using different tests.

Item-related properties of the target collocations were also important in explaining participants' performance. Congruency played a role at the levels of form recall and meaning recognition at both the immediate and delayed posttests and of form recognition at the delayed posttest: congruent collocations in which there was a high degree of L1-L2 overlap in the form-meaning link (e.g., "saturated fat" - *grasa saturada*) were easier to recall and recognise compared to incongruent collocations (e.g., "fond memories" - *buenos recuerdos*). This was an expected result as congruency reflects the degree to which the L2 collocation is plausible in the L1 and supports previous studies that have shown that

congruency has a positive effect on the incidental learning of collocations (e.g., Vu & Peters, 2022a). This finding also corroborates the ample evidence that learners rely on the L1 pattern of single words to infer word meanings (e.g., Granger, 1998; Irujo, 1988; Peters, 2016). This was manifested in some of the erroneous responses for form recall (e.g., "withdrawal syndrome") and inappropriate matching in the delayed form recognition test (e.g., "artificial sugar").

Unexpectedly, congruency did not affect immediate form recognition in the data, but this may be related to the layout of the matching task used in the online platform, which several participants reported finding taxing, as they had to monitor which collocations' constituents had already been used. As a result, the layout of this task was slightly modified at delayed testing so that learners were able to see the collocations' components and track their responses simultaneously. Learners were also encouraged to use pen and paper at delayed testing. Although minor, we cannot rule out the possibility that these changes might have made the delayed form recognition test easier than the immediate test. Consequently, comparisons of the immediate and delayed form recognition scores should be made with caution. Taken together, however, the results are consistent with other research that has reported congruency as a strong predictor of incidental collocational learning (e.g., Nguyen & Webb, 2017; Vu & Peters, 2022a), although caution should be exercised in interpreting the findings about congruency. As previously mentioned, congruency is likely correlated with other linguistic dimensions (such as semantic transparency), which are frequently poorly controlled for in studies of incidental learning. Consequently, itis difficult to attribute effects solely to it. More research is thus warranted so as to explore the amount of overlap between congruency and other linguistic factors.

Results also revealed a significant negative interaction between prior vocabulary knowledge and congruency in the delayed form recall model. This indicates that congruency benefits delayed form recall to a greater extent for learners whose vocabulary size is smaller than for learners with a larger vocabulary size. I hypothesise that there might a be a "sweet spot" for the beneficial effect of cross-linguistic congruency; that is, for learners whose proficiency level is already high, there could be a limit on the ability of congruency to facilitate the acquisition of collocations (at least as shown by the items used in this study). This finding suggests that the strength of the impact of L1-L2 congruency might be determined by the level of proficiency of the learners. While this experiment does not challenge the involvement of L1 influence in the congruency effect, it shows that collocational competence is shaped by several factors that can interact with other variables such as individual differences among learners.

In the preliminary analysis, compositionality was found to be highly correlated with congruency, lending support to the proposition that both parameters seem to be closely related. Yamashita (2018) found that most congruent collocations are classed as compositional, while incongruent collocations are non-compositional, based on an analysis of previous studies that looked at these two factors. She concludes that both factors influence collocational processing and acquisition. This experiment revealed a confound between the two, indicating that the association appears to be strong, not only from a processability standpoint (Gyllstad & Wolter, 2016; Yamashita, 2018), but also from a learning perspective. Further research into compositionality and congruency at the same time could shed light on the role of meaning in collocation acquisition. Yet, we have seen that measuring compositionality is not a straightforward task as it is difficult to quantify in absolute terms. Moreover, not all collocations are equally (non)compositional in meaning (Bartsch, 2004), despite being classified as compositional because they are typically compared to idioms (fully opaque items) (Benson et al., 1986; Skandera, 2004). There is, however, great practical and pedagogical value in researching this variable as learners encounter collocations with varying degrees of compositionality. Future studies could compare collocations with varying degrees of compositionality using different operationalisation methods, such as think-aloud protocols or follow-up interviews (Macis & Schmitt, 2017).

Finally, a comment on the effect of repetition, a factor that was not intentionally manipulated in this experiment. Only three of the fourteen collocations were repeated 2, 4, and 7 times. Surprisingly, repetition predicted knowledge of meaning recognition at immediate posttest, despite its narrow range. This result has two plausible explanations. On one hand, it may be that repetition only affected meaning recognition because it is the easiest collocational aspect to learn (Laufer & Goldstein, 2004). It would have been interesting to see if this finding was consistent on a productive test (meaning recall). On the other, it is possible that repeating the same collocations in different contexts aided their learning. Varied repetition, or encountering collocations in various contexts, can enhance vocabulary learning (Nation & Webb, 2017). The combination of varied repetition within the same short text could thus explain its immediate effect. However, the benefits of varied repetition may have faded after a two-week interval. It would be interesting to investigate how repetition could be manipulated to have a long-lasting effect on incidental collocation acquisition.

In terms of practical implications, the findings of this experiment show that academic reading is a potential source for L2 learners to acquire collocations. However, given the unique perspective that

every learner brings to the challenge of making sense of L2 collocations, special attention should be given to the impact of different treatments in relation to the learners themselves, their learning experiences, and their individual differences. In addition, learners' familiarity (and preferences) towards a specific English accent might affect learning when speech is part of the input. Some of the participants reported being less familiar with a British English accent, which might have detrimentally affected their collocational learning in this reading mode. Moreover, as shown by previous studies (e.g., Vu & Peters, 2021), it is clear that the needs of specific groups of learners remain a key factor that should guide our pedagogical approaches to L2 collocational learning. More specifically, the results of this experiment suggest that there might be a threshold for the impact of congruency, that is, the higher the proficiency level of the learners, the less of an effect of congruency when it comes to the L1-L2 collocational equivalence.

4.6. Limitations

All the findings presented must be interpreted in light of the inevitable limitations of this experiment. The use of authentic collocations increases the ecological validity of the experiment, but it also limits the generalisability of the results as it is difficult to control for prior knowledge of the individual components of the items. For example, some of the component words that made up the collocations in this experiment were high-frequency words (e.g., "chocolate", "meal"), which may have affected the learning gains. However, some items included less frequent constituents that were more likely to be unknown (e.g., "hay", "ploy"). Word frequency of the individual component words is a significant variable to consider, particularly when using authentic collocations. Future studies could look into how gains are impacted by prior knowledge of the component words. It is possible that component word unfamiliarity makes collocations less salient, decreasing their learning potential, or that component word familiarity makes collocations easier to infer (Webb & Chang, 2020). It is also important to note that the difficulty of the collocations selected for this experiment was likely to vary, as well as with the difficulty levels of collocations in previous studies.

Another limitation is the lack of norming of the collocations. I noticed that some items were familiar to the participants, most notably "soft drinks". This item was recognised at pretest (i.e., the components were correctly matched) by 45 participants, but several failed to provide an accurate Spanish translation of the item. Toomer and Elgort (2019) suggest norming for familiarity with a comparable group of L2

learners to help identify familiar collocations in the piloting, allowing a chance to replace them with less familiar ones. Another solution would have been to use a different, non-dichotomous scoring system at the pretest. Some of the partially correct answers on the pretest revealed that participants had either intuitions or partial knowledge of the items, even if their translations proved to be incorrect. In my view, answering based on partial knowledge and perceptions differs from blind guessing (see McDonald & Asaba, 2016). Future studies could also include subjective certainty ratings to examine participants confidence in their responses (e.g., Pellicer-Sánchez, 2017).

As noted in the Discussion, test effects may have been inevitable. This is typically a concern for studies with a pretest-posttest design where learners are tested after each session as they could have been alerted about the target items (Laufer, 2003), although I tried to minimise test effects by adding (and keeping) distracters in the pretests and posttests. In the future, it would be useful to add a control group so that treatment effects can be isolated. Future studies should also consider adding follow-up interviews or questionnaires to ask learners whether they looked up the collocations or attempted to memorise the items from the tests.

The next limitation pertains to the measurement of the receptive knowledge of meaning (i.e., meaning recognition) which should be improved upon as the test employed in this experiment may have led to an overestimation of the scores due to a ceiling effect. Future studies should consider alternative measurements of form and meaning to tap into different degrees of collocation knowledge. Further, the minor modification between the layout of the form recognition task at the immediate and delayed posttests may have also affected the scores. The experiment is also limited by the fact that form recall and meaning recognition were not pretested, although the collocations pretest assessed meaning recall to prevent blind guessing and ensure that learners did not know the target items.

Finally, there are some constraints when doing web-based research. Because the data was collected in a computer-based environment, the teacher was unable to monitor the screen of every single learner, although invigilation was in placed to oversee what participants were doing. Further, the experimental platform recorded timing checks, which were eventually employed as a proxy to detect inattention and monitor participants' performance.

4.7. Conclusion

Experiment 1 adds to the body of evidence on the effectiveness of RO and RWL on incidental learning of collocations. The findings indicate that RWL did not cover an advantage over RO, which contradicts previous research that used fictional stories as reading materials but supports Dang et al. (2022), who focused on academic lectures. This highlights the importance of considering the complexity of both the type of input and the properties of the reading materials when developing effective pedagogical strategies. This experiment confirms previous findings that learner- and collocation-related properties (i.e., prior vocabulary knowledge and congruency) influence incidental collocation learning.

5. The effect of repetition

As seen in Chapter 3, a collocation's learning burden is determined not only by its properties, but also by the difficulty of encountering it repeatedly in a relatively short span of time of natural discourse (Webb & Nation, 2017). The quantity (i.e., number of encounters) and quality of the input are important aspects of collocation learning, especially in incidental learning, where authentic input does not provide sufficient exposures to collocations (Macis et al., 2021).

This chapter presents Experiment 2, which investigates the effect of repetition on the incidental learning of collocations through reading multiple expository texts. I begin by explaining the background of the experiment, emphasising the importance of examining repetition in conjunction with the variables that may interact with it.

5.1. Background

As noted in Chapter 3, repetition regulates incidental collocation learning in various ways. Several studies have found that repetition has a positive effect (Durrant & Schmitt, 2010; Toomer & Elgort, 2019; Vilkaitė, 2017; Webb et al., 2013; Webb & Chang, 2020), while others have not (Macis, 2018; Pellicer-Sánchez, 2015; Szudarski & Carter, 2016).

The varying effect of repetition across studies is likely due to the differences in the frequency range of the target items (1-15 repetitions), the reading materials (e.g., graded readers vs short stories vs expository texts), type of input (academic vs non-academic), and the collocations examined (lexical vs grammatical), among other factors. A crucial variable to consider is the type of repetition (i.e., massed vs spaced), which can influence the effectiveness of activities for incidental collocation learning (e.g., Macis et al., 2021). As noted in Chapter 3 (see Section 3.5.1.3.), massed repetition is typically employed in incidental learning conditions to "flood the input" with several encounters of the target item. For example, Macis et al. (2021) found that massed repetition (as opposed to spaced practice) was more effective for incidentally acquiring the form of collocations to a level needed for recall tasks through the reading of narrative texts. Previous studies also indicate that encountering the same collocations multiple times contributes to collocation retention (e.g., Durrant & Schmitt, 2010; Sonbul & Schmitt,

2013). Sonbul and Schmitt (2013) showed that after encountering technical collocations three times in a medical-related text, advanced learners developed both recall and recognition knowledge of the items. Using sentence contexts, Durrant and Schmitt (2010) found a positive effect of verbatim repetition for retention (i.e., participants read the same sentence twice). However, Szudarski and Carter (2016) did not find a positive effect of input flood. They found that six or twelve encounters with 20 target collocations did not lead to learning gains after two weeks. The lack of an effect in this study is likely due to the types of collocations used. The verb-noun collocations contained delexicalised verbs, while the adjective-noun items contained low-frequency nouns (e.g., "retort"). As a result, the beneficial effect of repetition on learners' knowledge of collocations requires further examination.

Taken together, the studies that have looked at repetition show that the impact of other factors, particularly those affecting the properties of the reading materials, may outweigh the effect of repetition. Some researchers contend that the frequency-based approach to collocations may be insufficiently nuanced to account for the sheer complexity of factors that can influence collocation learning (Pellicer-Sánchez & Boers, 2019). While repetition is important, the evidence suggests its impact interacts with a variety of other variables (e.g., Peters, 2016; Webb & Chang, 2020; Pellicer-Sánchez, 2020). More empirical studies are thus required to establish how different factors can shape the contribution of repetition to incidental collocation learning.

5.2. Aims and research questions

This experiment examines the extent to which repetition can affect incidental collocation learning, considering the type of texts used as well as the contextual support of the materials. Factors related to the learners (prior vocabulary knowledge) and the collocations (association strength, corpus frequency, congruency, and compositionality) will be also investigated in order to assess whether the effect of repetition is modulated by them, and if so, for which aspects of collocational knowledge (e.g., form recall, form recognition). The following research questions were addressed in this experiment:

1. Does repetition of collocations in supportive contexts promote incidental collocation learning? Are specific aspects of collocation knowledge (form recall, form recognition) developed further?

2. Do learner-related (prior vocabulary knowledge) and item-related factors (congruency, compositionality, corpus frequency, association strength) influence incidental collocation learning through reading?

5.3. Methods

5.3.1. Participants

Eighty-two participants (47 female, 35 male) were recruited from a Spanish University. They were native speakers of Spanish and were between the ages of 18-26 (M=21.34,SD=2.38). Their proficiency was upper-intermediate at least or a B2 level according to the CEFR. All participants had formally studied English for at least ten years (M=14.72, SD=4.66). To evaluate their prior vocabulary knowledge, I administered the revised version of the VLT (Schmitt et al., 2001). Participants scored an average of 27.34 out of 30 on the 3K level *(SD*=3.03). On the 5K level, the average score was 26.01 out of 30 (SD=3.89) and a mean score of 23.63 (SD=4.53) was revealed on the 10K level. The VLT average score out of 90 was 77.13 (SD=8.71). Participants self-rated their English listening, reading, writing, and speaking skills at pretest. All participants rated[12] their reading (MAX=10, MIN=7, M=9, SD= 0.96) above 7 in a 10-point scale. As in Experiment 1, participants who completed all components of the experiment were entered into a raffle where they had the opportunity to win an e-gift voucher in compensation for their participation.

5.3.2. Design

A pretest-posttest-delayed posttest, Latin-square design was adopted. Two types of outcome measure (form recall and form recognition) were collected, to obtain a finer-grained picture of the degrees of learning (Webb, 2005). These were measured at two times: immediately after the treatment and two weeks after (delayed posttest), resulting in a total of four outcome variables.

5.3.3. Reading materials

Twenty-four short texts were extracted from a range of different sources: language corpora (BNC and COCA), medical resources (the *Collins dictionary [of] medicine*, Youngson, 2004; the *Mosby's Medical, Nursing, and Allied Health Dictionary*, Anderson, Keith, and Novak, 2002; and *Stedman's Medical Dictionary for the Health Professions and Nursing*, Stedman, 2005), as well as Internet medical

---

[12] Only reading is reported as this is the relevant skill for the experiment.

resources (the NHS website for England and the John Hopkins medicine website, among others). Eighteen of them contained the target collocations (see Section 5.3.4.) and the remaining 6 were included as distracters.

To increase the comprehensibility of the texts, word frequency was controlled for. Each text had at least 96% lexical coverage using the most frequent 4,000 words of the British National Corpus (as calculated with the Vocabprofile tool of the Compleat Lexical Tutor-Cobb, n.d.; Nation, 2006, 2012). While 98% coverage might have been more adequate for an optimal comprehension of the texts (and for promoting guessing from context) (Laufer & Ravenhorst-Kalovski, 2010; Nation, 2013), achieving such coverage proved to be impossible without removing several medical terms required for content comprehension and compromising the texts' coherence (e.g., "transplant", "disease"). However, based on experience with learners with a similar background to the participants of the experiment, it was predicted that they would have known about 4,000-5,000 words of English receptively. This was later corroborated by the participants' average score of 26.53 out of 30 (SD=1.69) on the 5K of the VLT. Therefore, the comprehensibility of the texts used was judged to be appropriate for adequate (if not optimal) successful L2 reading comprehension, which was eventually confirmed in the piloting.

A modified version of Webb's (2007a) context specification scale was used to rate the level of contextual support of the texts (see Appendix 2A). Five English native speakers with a TEFL (teaching EFL) background were asked to read the passages and rate the level of contextual support on a scale from 1 (contextually unsupportive) to 3 (contextually supportive). The experiment only included texts that received a rating of 3. This gave me some control over the types of contexts I chose as reading materials. In Beck et al.'s (1983) terms, the contexts in which the collocations appear are pedagogical contexts: they include context cues about the target collocation's meaning.

The reading materials were modified to include 1, 3, and 5 repetitions of the target collocations. The changes made to the texts were minimal to ensure they remained as authentic as possible. For example, where possible, the target items were replaced with pronouns. The final versions of the texts were approximately 100 words long (range 85-110) and were presented as single paragraphs. The materials were checked by two native speakers of English to ensure that they sounded natural after the modifications. All texts are available in Appendix 2A.

The number of repetitions in this experiment (1, 3, 5) were chosen considering findings from previous studies. Durrant and Schmitt (2010) found that a single exposure can result in an initial memory

trace which can later be strengthened through subsequent encounters. The 3-repetition range was chosen because Peters (2014) found that 3 exposures stimulated collocation learning when using decontextualised vocabulary activities, so I was interested in examining whether this would be the case in learning collocations through meaning-focused input. Five repetitions were included because previous research has indicated that 5 encounters can lead to incidental collocation learning (Macis et al., 2021; Webb et al., 2013).

A factor that has not been thoroughly examined is the relative proximity of the repetitions, which could potentially increase the likelihood of successful guessing from context (Carnine et al., 1984; Nation, 2013). Taking this into account, the encounters with the collocations were seeded into short texts (~100 words) presented in a single paragraph. The frequency conditions (1, 3, and 5 repetitions) were distributed as follows: the first encounter with the items was embedded in the first sentence in all conditions, therefore in the 1-repetition condition learners saw the target collocation at the beginning of the text. In the 3-occurrence condition, the subsequent encounters with the collocation were placed in the middle of the paragraph (second encounter) and at the end, almost always in the final sentence (third encounter). Wherever possible, on the 5-repetition condition, the stimuli paragraphs consisted of 5 sentences where learners encountered the target items once in each sentence. Increased repetition was employed in creating the 3- and 5-repetition conditions. This decision was made to indirectly test how the proximity of repetitions could enhance the noticing and retention of the target items. Examples of the different conditions are provided below.

Example 1-repetition condition.

Unlike a cold, **hay fever** isn't caused by a virus, it is caused by an allergic response to outdoor or indoor allergens. If you're someone who regularly gets it at particular times of the year, you might be able to spot your typical symptoms when they start. Having a runny or blocked nose is a very common symptom of this allergic reaction. Some people may also experience wheezing or breathlessness, particularly if they also have asthma. There's currently no cure for this condition, but you can do things to ease your symptoms when the pollen count is high.

Example 3-repetition condition.

Unlike a cold, **hay fever** isn't caused by a virus, it is caused by an allergic response to outdoor or indoor allergens. If you're someone who regularly gets it at particular times of the year, you might be able to spot your typical symptoms when they start. Having a runny or blocked nose is a very common symptom of **hay fever**. Some people may also experience wheezing or breathlessness, particularly if they also have asthma. There's currently no cure for **hay fever**, but you can do things to ease your symptoms when the pollen count is high.

Example 5-repetition condition.

Unlike a cold, **hay fever** isn't caused by a virus, it is caused by an allergic response to outdoor or indoor allergens. If you're someone who regularly gets **hay fever** at particular times of the year, you might be able to spot your typical symptoms when they start. Having a runny or blocked nose is a very common symptom of **hay fever**. Some people with **hay fever** may also experience wheezing or breathlessness, particularly if they also have asthma. There's currently no cure for **hay fever**, but you can do things to ease your symptoms when the pollen count is high.

The texts were presented one at a time (one passage per screen) in Arial size 18. Participants were unable to return to previously read screens. Participants were required to read the passages once (see Appendix 2A).

5.3.4. Collocations

Except for the operationalisation of collocational strength, the same procedure used in Experiment 1 (see Chapter 4, Section 4.3.4.) was used in this experiment to operationalise collocations. Instead of the t-score, the log Dice score is used in this experiment to determine whether a word pair is a strong collocation The log Dice score is a statistical measure used to calculate the association strength of collocations in a corpus. It is based on the concept of observed and expected co-occurrence frequencies of word pairs. It is calculated using the frequency of collocates, nodes, and collocational pairs. The log Dice value ranges from negative infinity to positive infinity. A positive log Dice score indicates a strong collocational association, while a negative score indicates a weak or non-collocational association. Overall, this measure is a valuable tool in collocation studies as it has been shown to yield reliable results for collocation candidates (see Gablasova, 2017). This score is designed to include technical and/or restricted collocations ranging from very strong ("bone marrow") to fairly strong ("heat

rash") (see Chapter 3, Section 3.6.3.5. for a more in-depth explanation of the selection of strength association measures). The significance cut-off threshold was set at 5 or higher, given that log Dice is usually less than 10 (see Rychlý 2008).

Eighteen target collocations were selected for this experiment. Raw corpus frequencies were obtained from the COCA, using Davies's (2008) interface. COCA frequencies ranged between 14 and 3,917 occurrences. As in Experiment 1, only noun-noun and adjective-noun lexical collocations were included. log Dice scores ranged between 5.6 and 12.4, indicating a medium-high collocational strength (Rychlý, 2008). The target collocations included both congruent (9) and incongruent (9) collocations. Collocations also varied in compositionality, whose ratings ranged from 0.12 to 0.98 (*SD*=0.26) (see Table 15).

Table 15. List of target collocations and their properties.

Collocation	Congruency	Compositionality	log Dice	COCA
gum disease	congruent	0.74	7	248
hay fever	congruent	0.45	10.8	252
head lice	congruent	0.74	6.1	181
cerebral palsy	congruent	0.86	12.4	903
acute pain	congruent	0.94	7.7	167
kidney failure	congruent	0.98	8	652
yeast infection	congruent	0.72	6.1	233
wasp sting	congruent	0.89	9.4	14
stomach cramps	congruent	0.82	9	145
bone marrow	incongruent	0.48	11.3	2317
hot flushes	incongruent	0.74	7.9	74
bedside manner	incongruent	0.14	6.6	292
cold sore	incongruent	0.12	6.6	105
heat rash	incongruent	0.50	5.9	42
stem cells	incongruent	0.6	8.2	3917
chief complaint	incongruent	0.24	5.6	76
rib cage	incongruent	0.82	11.08	806
stretch marks	incongruent	0.69	9.6	334

5.3.5. Measures

5.3.5.1. Pretests

Learner-related information was collected in pretest assessments two weeks before the treatment. In this session, all participants signed the consent forms, filled out a slightly modified version of the LEAP-Q (Marian et al., 2007) and completed the 3K, 5K, and 10K sections of the revised version of the VLT (Schmitt et al., 2001) as well as the pretests on the target items. The LEAP-Q and revised version of the VLT used in this experiment are described in Chapter 4, Section 4.3.5.1.

The preknowledge test of the target collocations consisted of a form recall and a form recognition tasks. The form recall pretest measured productive knowledge of the form-meaning link using a translation format in which the L1 meanings cued the responses. Participants were provided with the first letter of each component word of the collocations to avoid elicitation of other word pairs and asked to provide the L2 collocation (see example below).

Example of form recall test.

sarpullido:

h_____ r_____

A second pretest was administered to tap into the form recognition of the items, and therefore, account for partial knowledge of the collocations. For example, learners may be familiar with one of the words in the collocation but not know which word it commonly occurs with. This methodological decision was made taking into account that learners could have different degrees of knowledge of the items, which can have an effect on the amount of knowledge that is gained (see Webb et al., 2013). This test had a multiple-choice format that measured the learners' ability to recognise the correct form of the target items among five options: the correct form, three fillers, and one "I'm not sure" option. To prevent participants from blind guessing, they were advised to select the "I'm not sure" option if they had no knowledge of the answer. L2 learner corpora was used to find appropriate fillers for this test: the *Universidad Autónoma de Madrid Corpus de Interlenguas Escritas* (extracted using the UAM Corpus tool, O'Donnell, 2008) and the Spanish subcomponent of the *International Corpus of Learner English* (ICLE), SPICLE. Incorrect responses given by the participants in the piloting phase were eventually added as options of the final multiple choice test to increase the difficulty of the test.

Example of the form recognition test.

heat _____

a) spot

b) fever

c) rash

d) feeling

e) I don't know

The pretests contained 16 distracters to draw participants' attention away from the target items. The distracters were semantically related to the collocations (e.g., "biological clock", "health insurance", "period pain", "vocal cords") but considered substantially easier. The distracters were chosen by researching the English textbooks and materials that participants were using as part of their studies, as well as consulting with the teachers.

5.3.5.2. Immediate posttests

Postreading comprehension

To assess the participants' overall comprehension of the texts, a postreading reading comprehension test was distributed immediately after the treatment. Participants were asked to rate their understanding of the passages on a 5-point scale (from *I struggled to understand* to *I had no trouble understanding the texts*) and they could leave specific feedback. They were also asked to rate the perceived difficulty of the texts on a scale from very difficult to very easy . A multiple-choice question was included to assess whether participants had paid attention to the texts, as checking for understanding of every single text was logistically unworkable. Participants had to select the topics they had encountered in the texts from a 10-item list, where only 5 were correct. This question aimed at identifying participants who might have skipped some passages as identifying the right topics was an easy task. The full comprehension test is available in Appendix 2A.

Form recall and form recognition posttests

Two vocabulary tests followed the postreading comprehension test: a form recall test and form recognition test, in that order. The tests were identical to the ones administered at pretest, although the items were randomised. The distractors were kept in the posttests to minimise testing effects.

5.3.5.3. Delayed posttests

At delayed testing, the same form recall and form recognition tests were administered, with the items re-randomised.

As in Experiment 1, a retrospective questionnaire was distributed at the end of the experiment. The questionnaire collected qualitative data on participants' overall engagement with academic texts,

perception of learning (both for content and vocabulary learning), and vocabulary learning strategies (if employed). The questions were answered using a mixture of different rating scales and multiple-choice questions (see examples below). Participants were also given the opportunity to leave specific feedback if they wanted to (see Appendix 2A).

Example question of the retrospective protocol questionnaire.

2. Did you learn any words or terms that you didn't know before participating in this study?

| I didn't learn any words or terms | I learnt one or two new words or terms | I learnt a few words or terms | I learnt a lot of new words or terms I didn't know before |

Example question of the retrospective questionnaire.

4b. [If you answered yes to the previous question] Please tell us which of the following strategies you use when learning new vocabulary:

a) Guessing from context

b) Bilingual dictionary use (Spanish-English)

c) Monolingual dictionary use (only English)

d) Morphemic analysis, i.e., inferring the meaning of new words by examining their meaningful parts (prefixes, suffixes, etc.)

e) Others: please specify _____

5.3.6. Procedure

The same procedure as in Experiment 1 was used, which is described in Chapter 4 (see Section 4.3.4). As in Experiment 1, participants completed the 3 components of the experiment over 3 sessions, separated by approximately 2 weeks each (see Table 16). All components were distributed online via SoSci Survey (Leiner, 2019, Version 3.1.06). The final dataset comprised 6,048 observations, from 84 participants to 18 items, with 4 outcome variables.

Table 16. Data collection procedure.

Week	Procedure
1	Informed consent form Background questionnaire Vocabulary test Collocations pretests
3	Treatment Reading comprehension test Immediate form recall test Immediate form recognition test
5	Delayed form recall test Delayed form recognition test Retrospective questionnaire

5.3.7. Piloting

The experiment was piloted with a pool of 12 EFL learners with a similar proficiency level and language learning background to that of the participants of the actual experiment. Following the criteria specified in Section 5.3.4. an initial pool of 36 potential target collocations were selected for piloting. This pilot study was carried out to maximise the chances that the participants of the experiment would not already know the target items. Participants in the pilot study were asked to read the texts, fill out the collocation pretests (form recall and form recognition) as well as the reading comprehension test so as to ensure that the level of difficulty of the texts was appropriate. One-to-one interviews were conducted to evaluate whether learners had partial knowledge of the items. The collocations that were found to be relatively well known (e.g., "blood clot", "sprained ankle") or problematic due to its level of specialisation (e.g., "flail chest") were removed. The final 18 target items were those that were clearly unknown by the learners in the piloting. The participants showed no difficulty in answering comprehension questions. Therefore, the texts were judged to be appropriate for the target population.

5.3.8. Data processing and model fitting

5.3.8.1. Exclusion criteria

The answers to the postreading comprehension test indicated that the participants had no trouble understanding the reading materials and that they had found them reasonably easy to understand. Eighty per cent of the learners reported understanding the main ideas and most details. All participants identified the topics correctly, which indicated that they had paid attention to the content of the texts. No participant was excluded from the data set.

5.3.8.2. Scoring responses

The outcome variables were binary (correct response=1, incorrect response=0). Responses in the form recall tests that contained minor spelling mistakes were scored as correct as long as the response was comprehensible (e.g., "bone *marro*").

5.3.8.3. Model fitting

Before analysing the data, the predictors (compositionality, congruency, preknowledge of collocations, prior vocabulary knowledge) were assessed for correlations among variables. To avoid collinearity in the models, only congruency and prior vocabulary knowledge[13] were used in the analysis.

As in Experiment 1, the data for the present experiment was analysed using the R software environment (version 4.1.1, R Core Team, 2021). Unaggregated data were then fitted to logistic mixed-effects regression models (Baayen et al., 2008; Jaeger, 2008). Four separate models were fit for each outcome measure (form recall, form recognition) and each delay (immediate posttest, delayed posttest). Each model included fixed effects of all the potential explanatory variables: frequency of occurrence, item-related factors (congruency, compositionality, corpus frequency, association strength), two-way interactions with repetition, where the interaction would have been interpretable, and learner-related variables (prior vocabulary knowledge). To be able to compare the magnitude of effects across the form models, all models have the same set of predictors (see Chapter 4 for rationale).

---

[13] As in Experiment 1, pretest scores were computed to estimate the amount of collocational knowledge learners had at the time of the pretest. The mean average (*M*=0.35, *SD*=2.05).

Models for each of the outcome measures were selected as described for Experiment 1, with fixed effects determined first, and including the maximal random effects structure supported by the data.

5.4. Results

5.4.1. Descriptive statistics

Absolute mean gains by repetition for form recall and form recognition on the immediate and delayed posttests are reported in Table 17. Overall, the tables show that increased repetition resulted in more raw gains regardless of the knowledge aspect examined. The scores are lower at the level of form recall when compared to the scores of form recognition, which suggests that learners first recognise the collocation's form (active recognition). The lower scores in the form recall indicate that the ability to recall the collocation's form (active recall) is more effortful. This result shows that active recall is more difficult than passive recognition (Sonbul & Schmitt, 2013). The mean scores in the delayed postttests are lower than in the immediate posttests, although differences are small.

Table 17. Absolute mean gains by condition and outcome measure.

	Form Recall		Form Recognition	
Repetitions	Pre-Post $M\,(SD)$	Pre-Delay $M\,(SD)$	Pre-Post $M\,(SD)$	Pre-Delay $M\,(SD)$
1	1.78 (0.65)	1.63 (1.45)	5.44 (1.29)	5.19 (2.46)
3	3.71 (0.89)	3.52 (1.38)	8.04 (1.83)	7.77 (3.08)
5	4.89 (1.37)	4.49 (1.67)	8.36 (2.05)	8.12 (2.71)

Note: Maximum: 18. Pre-Post refers to gains accrued from pretest to immediate posttest, Pre-Delay refer to the gains from pretest to delayed posttest.

5.4.2. Models

This section presents immediate and delayed posttest models for each type of outcome measure: form recall and form recognition. All full models are available in Appendix 2B.

5.4.2.1. Form recall models

As Table 18 shows, repetition and vocabulary size were significant predictors in the immediate form recall model. Congruency and association strength (as measured by log Dice) were also significant predictors of collocation learning at form recall, implying that congruent collocations with strong association strength were answered significantly more accurately than incongruent items with lower association strength scores. Two interactions marginally influenced these main effects. The negative interaction between prior vocabulary size and repetition suggests that the effect of repetition decreases for learners with higher existing vocabulary knowledge. To put it another way, the effect of repetition is stronger for learners whose vocabulary size is smaller. The negative interaction between association strength and repetition suggests that repetition has less of an effect for items with a higher log Dice score, which is most likely due to a ceiling effect for repetition: the amount that repetition can aid learning may be limited for collocations that are already easier to learn due to their high association strength.

Table 18. Summary of fixed effects for immediate posttest form recall.

Model:				
Response ~	VocabSize + Repetition + Congruency + Association			
	+ Association:Repetition + VocabSize:Repetition			
	+ (1+VocabSize|ParticipantID) + (1+VocabSize|ItemID)			

Predictor	**Estimate**	**SE**	**z value**	**Pr (>|z|)**
Intercept	-2.79	0.49	-5.60	0.00
VocabSize	0.75	0.16	4.63	0.00
Repetition	0.48	0.11	4.23	0.00
Congruency	0.54	0.21	2.59	0.00
Association	0.14	0.05	2.52	0.01
Association:Repetition	-0.02	0.01	1.75	0.08
VocabSize:Repetition	-0.07	0.04	1.90	0.05

The main effects of vocabulary size, repetition, and association strength remained reliable in the delayed model, but congruency is no longer significant (see Table 19). The lack of significance of congruency suggests that other factors play a more significant role in delayed collocation form recall.

This finding will be discussed in the next section. The negative interactions between association strength and repetition, and between vocabulary size and repetition reached significance in the delayed model. The former indicates that there is less of an effect of repetition on items with a higher collocational strength. The latter suggests that the effect of repetition strengthens at delayed postttest for learners who have smaller vocabularies.

Table 19. Summary of fixed effects for delayed posttest form recall.

Model:				
Response ~	VocabSize + Repetition + Congruency			
	+ Association:Repetition + VocabSize:Repetition			
	+ (1+VocabSize|ParticipantID) + (1+VocabSize|ItemID)			

Predictor	**Estimate**	**SE**	**z value**	**Pr (>|z|)**
Intercept	-3.65	0.61	-5.99	0.00
VocabSize	0.63	0.17	3.62	0.00
Repetition	0.57	0.12	4.85	0.00
Congruency	0.48	0.29	1.62	0.10
Association	0.18	0.07	2.79	0.00
Association:Repetition	-0.03	0.01	2.17	0.02
VocabSize:Repetition	-0.09	0.04	2.10	0.03

5.4.2.2. Form recognition models

As shown in Table 20, the immediate form recognition model revealed that vocabulary size and repetition significantly affected the odds of a correct answer in this test. Repetition enabled learners to strengthen their (passive) knowledge of collocation forms, and their prior vocabulary knowledge facilitated performance. However, neither congruency nor association strength had a significant effect on the recognition of collocations at the immediate posttest. This suggests that the main difficulty of incongruent items lies in their production rather than in their recognition.

Table 20. Summary of fixed effects for immediate posttest form recognition.

Model:

Response ~   VocabSize + Repetition + Congruency

+ (1+VocabSize|ParticipantID) + (1+VocabSize|ItemID)

Predictor	Estimate	SE	z value	Pr (>|z|)
Intercept	0.64	0.37	1.73	0.08.
VocabSize	0.47	0.10	4.43	0.00
Repetition	0.16	0.04	3.87	0.00
Congruency	0.18	0.21	0.83	0.40
Association	0.04	0.39	1.01	0.31

Looking at the delayed postttest model (Table 21), vocabulary size and repetition remained reliable predictors of form recognition. The effect of repetition was thus durable. Unlike in the immediate model, congruency (and marginally association) emerged as predictors in this model, although their effect is not as significant as the effect of vocabulary size and repetition. This result will be discussed in the next section.

Table 21. Summary of fixed effects for delayed posttest form recognition.

Model:

Response ~   VocabSize + Repetition + Congruency + Congruency + Association

+(1+VocabSize+Repetition|ParticipantID)+ (1+VocabSize+Repetition |ItemID)

Predictor	Estimate	SE	z value	Pr (>|z|)
Intercept	0.07	0.36	0.19	0.84
Vocab.Size	0.45	0.08	5.65	0.00
Repetition	0.16	0.04	3.72	0.00
Congruency	0.45	0.20	2.19	0.02
Association	0.07	0.03	1.89	0.05

5.4.3.3. Summary of the models

Table 22 summarises the results of the models. As shown in the table, repetition and vocabulary size are the most reliable predictors of form recall and form recognition of collocation at both testing times. This indicates that repetition had a significant effect on learning the form of collocations: as the number of encounters increased, so did the knowledge gained. The effect of vocabulary size indicates that a larger vocabulary size leads to greater form recall and form recognition of collocations. The finding that congruency only affected form recall at the immediate posttest is somewhat surprising. This finding will be discussed in the next section. Finally, the negative interactions between occurrences and vocabulary size indicates that as the proficiency of the learners' decreases the effect of repeated encounters increases.

Table 22. Summary of model results.

	Form recall		Form recognition	
Predictor	Immediate	Delayed	Immediate	Delayed
Repetition	+	+	+	+
Vocabulary size	+	+	+	+
Congruency	+	-	-	+
Association strength	+	+	-	marginal
Vocabulary size*Repetition	marginal	+	N/A	N/A
Association strength*Repetition	marginal	+	N/A	N/A

Note: Plus signs (+) indicate a significant effect, minus signs (-) show no significant effect.

The word marginal indicates that the effect was significant at .1 level.

5.5. Discussion

5.5.1. The effect of repetition

The main goal of this experiment was to investigate whether repetition had a positive effect on incidental collocational learning through reading. The experiment showed collocation learning at form recognition and form recall. The amount of knowledge gained increased with the number of repetitions, which suggests that repeated exposure of collocations may help L2 learners notice unknown items and

form-meaning links in the first place. It would be wrong to assume, however, that repetitions need to be seeded in a relatively short span of time. To ensure valid interpretation of the results, a distinction needs to be made between massed and spaced repetition conditions. Future studies could investigate the effect of specific repetition techniques on incidental collocational learning and the relationship between repetition and collocation retention. The results of this experiment suggest that embedding collocations in rich, meaningful contexts may help reduce the vocabulary load of a text, allowing learners to notice unfamiliar linguistic units other than individual words.

The findings revealed a significant effect of repeated exposure regardless of the time of posttest administration. The gains at the immediate posttest were slightly higher than the gains at the delayed posttest. This was expected, as the retrieval routes for retaining newly learnt vocabulary are weaker in delayed testing than in immediate testing and so delayed testing frequently results in lexical knowledge decay (Barclay & Pellicer-Sánchez, 2021). However, the differences were very small, which was somewhat surprising given the fact that the delayed posttests were administered two weeks after the treatment, so a larger collocational "loss" was expected. One possibility is that learners verified their answers after the treatment session. Participants were asked to report whether they had looked up the collocations after each testing session. Six participants reported looking up the meaning of some of the items after the treatment because they were unsure about the meaning of the collocations or had never heard of them before. Nonetheless, an analysis of their responses showed that this additional exposure had no effect on their scores.

It is worth examining the negative interactions found at the form recall level between repetition and prior vocabulary knowledge, as well as between repetition and association strength. The former indicates that the effect of increased repetitions is modulated by overall proficiency, with a stronger relationship between recall knowledge and repetition as proficiency decreases. Put differently, the higher the proficiency of the participants, the lower the effect of repetition at the level of form recall. One possible explanation for this result is that more proficient learners require fewer repetitions than learners with smaller vocabularies. Analysed further, this interaction demonstrates that vocabulary/collocation knowledge is incremental in nature, and thus lower proficiency learners benefit the most from increased repetitions. Additionally, it is more likely for learners with a larger vocabulary size to already have some partial knowledge of unknown collocations as they know more words. The negative interaction between repetition and association strength indicates that the effect of repetition decreases for collocations with

a high log Dice score. This effect is not very large though and it is likely due to a ceiling effect. Taken together, these interactions show that repetition interacts collectively with other variables (Pellicer-Sánchez & Boers, 2019), so it is reasonable to speculate that other factors may be moderating the role of repetition to different degrees.

These findings contribute to a body of studies showing that repetition enhances the incidental learning of collocations through reading (Durrant & Schmitt, 2010; Macis et al., 2021; Toomer & Elgort, 2019; Vilkaitė, 2017; Webb et al., 2013; Webb & Chang, 2020). They also support the idea that for incidental learning collocations to occur, repetitions need to be seeded in a relatively short span of time (Macis et al., 2021; Tsai, 2015), and to Schmitt's (2008:340) claim that "virtually anything that leads to more exposure, attention, manipulation, or time spent on lexical items adds to their learning". In this experiment, the effect of repetition was significant within a single learning session. It would be interesting to examine spacing repetitions in a similar treatment, as there is some evidence that massed repetition leads to greater collocations gains compared to spaced repetition. Macis et al. (2021) argue that encountering collocations repeatedly is particularly relevant for incidental learning given the relative rarity of collocations in unmodified texts. Future research could compare the effects of massed vs spaced repetition when learners are exposed to different contexts, i.e., context that are more or less "pedagogical". This type of research would allow us to determine how much contextual support affects the effectiveness of spaced and massed repetition on incidental.

5.5.2. Absolute gains

In terms of absolute gains, gains at form recall were significantly smaller than mean gains at form recognition. Absolute gains increased with the number of repetitions for both measures. Absolute gains for all conditions increased from pretest to immediate posttest but decreased from immediate to delayed testing.

Because of fundamental differences in the treatment, the gains reported here are not easily comparable to the gains reported in previous studies. In Macis et al.'s (2021) study, participants gained form recall knowledge of an average of 5 collocations out of 25. Macis et al. (2021) used a massed repetition technique, but they focused on fictional short stories, which are not directly comparable to the texts in this experiment. Dang et al. (2020) reported mean gains of 4.88 out of 19 collocations to the recognition level, but the frequency range (1-9) in that study was not directly manipulated because they

used an academic lecture as a source of input, i.e., authentic input. Vilkaitė (2017) and Toomer and Elgort (2019) used academic texts (2 and 9 texts, respectively), but Vilkaitė (2017) focused on verb-noun collocations (e.g., "assume duties"), which resulted in lower gains compared to this experiment. Toomer and Elgort (2019) investigated the effects of spaced repetition over two consecutive days, as opposed to this experiment, which focused on massed repetition in a single session. They found that spaced repetition was beneficial for collocation learning, although in this study each collocation occurred nine times, and the items were 15 low-frequency medical collocations (e.g., "split hand", "cloud baby").

Taken together, the larger gains in the present experiment are most likely due to the treatment, which included repeated exposure to the target items in short expository texts which showed a good amount of contextual support as displayed by the ratings. Importantly, findings from this experiment are compatible with the evidence passive knowledge aspects (e.g., form recognition) are easier to acquire than active knowledge aspects (e.g., form recall). This pattern of results in relation to the order of difficulty of collocational competence are in line with incidental collocation studies (e.g., Vilkaitė, 2017; Webb et al., 2013; Webb & Chang, 2020).

It is important to note that higher scores in form recognition posttests may be overestimated. The form recognition test assessed the ability to select the correct answer in a multiple-choice test, which is arguably problematic (Dang et al., 2020; Vilkaitė, 2017). Blind guessing was discouraged, but it is not possible to know whether participants followed the instructions. However, the majority of the participants did not select the "I don't know" option. The form recognition scores could also be explained by an order effect. Although this is unlikely, it cannot be ruled out that some learners were able to recall the cues given in the preceding form recall posttests, allowing them to choose to select the right answers in the multiple-choice test. For example, the first letter of each component word was provided.

5.5.3. The impact of learner- and item-related factors

As in Experiment 1, overall language proficiency in this experiment was conceptualised as the learners' vocabulary size and measured by means of a VLT composite score. As expected, the findings showed that the higher the learners' prior vocabulary knowledge, the better their level of recall and recognition of collocations. The result of this experiment corroborates the results of Experiment 1 in showing that vocabulary size is a good predictor of collocation learning at the levels of form recognition

and form recall. This finding is also in line with previous research (e.g., Puimège & Peters, 2019, 2020; Vilkaitė, 2017; Vu & Peters, 2022a).

Congruency had an inconsistent effect on learning. At the immediate posttest, only form recall but not recognition revealed an effect, which may suggest that the main difficulty of incongruent collocations lies in their short-term retrieval and not in their recognition. This finding is consistent with Peters (2016), who found that congruency only played a role in the form recall of collocations but not in the form recognition. Incongruent items are typically less likely to be recalled due to L1-L2 mismatches (e.g., Bahns, 1993; Nesselhauf, 2003, 2005). This result can be explained by considering the phenomenon of deceptive compatibility, which occurs when learners rely on L1 patterns to supply word combinations in the L2 (Laufer, 2011). There may be collocations that appear to be relatively congruent with the L1, but, upon closer inspection, they do not have literal equivalents in the L2. For example, in my data set, the collocation "cerebral palsy" induced learners to recall the infelicitous collocation "cerebral paralysis" because its Spanish equivalent is "*parálisis cerebral*". Previous studies have found such an influence of the L1 on collocational acquisition as well as better recall knowledge of congruent collocations (e.g., Gyllstad & Wolter, 2016; Peters, 2016; Wolter & Gyllstad, 2011).

Counterintuitively, congruency had the opposite effect in the delayed posttests, affecting form recognition but not form recall. There are two plausible explanations for this result. On the one hand, it cannot be ruled out that learners relied on different strategies at delayed testing. Closer examination of the individual responses in the form recognition test revealed that some participants made incorrect choices due to the availability of felicitous options in their L1. For example, 6 participants "gum disorder" over "gum disease" despite having previously supplied the correct form in the form recall test. This demonstrates how participants can fall victim to deceptive compatibility as long as this option is available in the test. However, I believe that some participants developed a level of collocational awareness that was counterproductive during delayed testing. It is possible that a learner rejected the L1 equivalent, which would normally correspond to the L2 congruent item, because they had previously noticed that some collocations were misleading. Some participants commented that they were puzzled by the presence of two "correct" answers in the immediate posttests, despite being told there was only one correct answer. They also reported that in the delayed posttests, they could recall the alternative "possible" option that they previously thought it was correct but was not the "right" answer.

The lack of a congruency effect at the level of form recall in the delayed posttest was unexpected. To try to explain this result, I examined the individual responses in the immediate and delayed form recall posttests and discovered some patterns. While some participants failed to recall some incongruent items at the immediate posttest, they were able to recall them in the delayed posttest. As previously mentioned, one explanation for this behaviour concerns the presence of L1-related fillers in the multiple-choice form recognition test, which may have aroused noticing of L1-L2 mismatches. A high-level conflict selection among competing options may have improved performance at recalling incongruent items after the treatment session, particularly if learners paid attention to the posttests. Though speculative, this interpretation is supported by Pulido and Dussias (2020), who found that L1-inference using plausible L1 distracters during lexical selection enhanced recall of incongruent collocations. From this standpoint, the authors argue that L1-interference can be a desirable difficulty during language learning (see Pulido & Dussias, 2020). More empirical research is needed to determine whether L1-distracters can improve recall knowledge of collocations. Moreover, since congruency is likely to be correlated with other linguistic factors (such as semantic transparency), its effect cannot be isolated due to lack of control of other variables. For theses reasons, congruency-related results should be interpreted with caution. These findings indicate that more research is needed before making firm assumptions about the effect of L1 congruency on different aspects of collocational knowledge.

Association strength, understood as the degree of statistical significance or frequency with which specific words co-occur together, emerged as a reliable predictor of collocation learning for form recall at both testing times. One possible interpretation for the effect of association strength in this experiment is that a collocation with a high log Dice score stands out as being more salient, which in turn affects its learnability. For example, one of the collocations with the highest log Dice score in this study was "bone marrow" (11.3). This item may have been more noticeable than other collocations whose log Dice score was lower (e.g., "heat rash", 5.9) because "bone marrow" is unlikely to appear in non-medical contexts, and one of its constituents ("marrow") is also unlikely to co-occur with other lexical items. These results align well with the characterisation of the log Dice score as reflecting both the association strength of the word components that make up a collocation, and its exclusivity (Gablasova et al., 2017). Nguyen and Webb (2017) similarly found that association strength (measured by MI scores) significantly predicted learners' knowledge of collocations.

The results and insights provided in the retrospective questionnaire added qualitative information to the learners' engagement and behaviours. As previously stated, six participants reported looking up the meaning of some collocations after the treatment. This shows that some learners were intrinsically motivated to learn new vocabulary. One participant stated that she checked the items because she was embarrassed to make spelling errors. The fear of making mistakes (or "language perfectionism") is a well-documented issue in SLA, particularly among adult learners (e.g, Gregersen & Horwitz 2002). Eight participants admitted paying more attention to the repeated items because they noticed their incidence: "*I think I noticed new words more easily because they were repeated, and I think I could recall them more quickly*" (Participant 34). In terms of engagement, 85% of the learners responded that they enjoyed reading the texts. Ten participants disliked the topic of the texts, but they all admitted that they had learnt new vocabulary that they did not know before the experiment. Participants reported using guessing from context as the most common vocabulary strategy, which supports that the idea that L2 learners engage in lexical inferencing when encountering new vocabulary (e.g., Schmitt & Schmitt, 2020; Webb & Nation, 2017).

This experiment makes theoretical and pedagogical contributions on incidental collocation learning, by showing that increased repetition enhances collocation learning when the level of contextual support is controlled for. This suggests that practitioners and textbook writers should not only consider the quantity of the input and the distribution of the encounters with the collocations, but also the quality of the input (contextual support) and how this relates to the learners' background and experience with the FL. This experiment demonstrates that incidental collocation learning can occur while reading academic texts, just like when reading fiction, as long as the learner's characteristics and prior language exposure are taken into account. Short texts of an expository kind may be thus a shortcut way to boost collocational development. Finally, the task of seeding a text with collocations may be more appropriate for material developers than for non-native EFL teachers, who do not always have the creativity or time to adapt and manipulate reading materials (Boers, 2020; Macis, 2018). Future research should take this into account as a key design principle to help promote incidental collocation learning, particularly in EAP programmes and/or English-medium universities.

5.6. Limitations

This experiment is limited in that it only measured the effect of repeated exposure at the level of form. Future studies could examine the effect of repetition on other types of collocations and knowledge aspects (e.g., meaning recall, meaning recognition).

Regarding the measures used, it is important to note that the tests used in this experiment may not have been sensitive enough to reveal the knowledge that students may have accumulated. An analysis of the individual responses in the form recall posttests revealed that some participants gained partial knowledge of the form of the items. For example, in some cases learners were able to recall the collocation's node, i.e., the central word or concept that carries the core meaning. Other instances showed that learners were able to recall the collocate, i.e., the word that commonly appears alongside the node, which adds meaning or context to the collocations. This methodological flaw has been noted in previous studies (e.g., González-Fernández & Schmitt, 2015; Peters, 2006). Future studies should include several productive and receptive tests to determine the extent of the knowledge acquired. As earlier mentioned, the scores in the form recognition posttests may have been distorted due to testing effects and should be carefully interpreted. Although the participants were placed at an upper intermediate or advanced level of proficiency, using a more generalised measure of language proficiency would have allowed for a more comprehensive analysis of the role of proficiency and vocabulary knowledge on the collocation gains, especially since the VLT assesses form recognition only, and L1-Spanish learners may find it easier to recognise English low-frequency words due to the influence of their L1.

Some participants reported having looked up the items after the treatment. It is likely that some participants did so without disclosing it. As a result, the gains measured in the delayed posttests must be interpreted with caution. In addition, immediate testing itself may also contribute to learning. Finally, the findings of this experiment do not necessarily apply to other EFL contexts where learners may not have prior experience reading academic texts.

Despite its limitations, this experiment makes some theoretical and pedagogical contributions on incidental collocation learning. This experiment shows that increased repetition enhanced collocation learning when the level of contextual support is controlled for. As a result, practitioners and textbook writers should not only consider the quantity of the input and the distribution of the encounters with the collocations, but also the quality of the input (context informativity) and how this relates to the learners'

background and experience with the FL. An additional aspect that future studies could explore may be the way in which readers normally engage with different types of text: narrative stories tend to be linear, and so they are read them in that way, while factual and expository texts are not linear and denser, and it is likely that readers, particularly L2 learners reread more often.

5.7. Conclusion

Experiment 2 adds to the body of evidence supporting the beneficial effect of repeated exposure under incidental learning conditions (e.g., Macis et al., 2021). Repetition enhanced collocation knowledge at the form recall and form recognition levels. This can be accomplished by manipulating the frequency of the items to ensure that collocations appear repeatedly in a short period of time (Boers & Lindstromberg, 2009). It is also important to consider the contextual support of the reading materials, which may modulate the gains accrued from reading different types of texts, e.g., highly informative (or pedagogical) vs less informative. The relatively large gains in this experiment suggest the benefits of providing meaningful context cues to make collocations easier to infer. How contextual support affects the incidental leaning of collocations through reading was explored in Experiment 3.

6. The effect of contextual support

The results of Experiment 2 indicated that repeated exposure improved collocation learning. They also suggested that pedagogical contexts may facilitate acquisition. However, it should be noted that the texts used in Experiment 2 were manipulated in such a way that learners could infer the meanings of the items based on the presence of contextual cues. Despite the importance of contextual support in incidental vocabulary learning (Nation, 2001), its impact on collocation acquisition has received little attention. To the best of my knowledge, no research has been conducted to explore the effect of varying levels of contextual support on incidental collocation learning, motivating the experiment presented in this chapter.

The focus of Experiment 3 was to investigate the efficacy of different degrees of informativity for acquiring explicit collocation knowledge at the levels of form recall and meaning recall. Before presenting the experiment, I review key findings from research on incidental vocabulary learning from context.

6.1. Background

As noted in Chapter 2, incidental learning from context is the most important of all the sources of vocabulary growth (Nation, 2001). The role of context has not been studied as thoroughly as the impact of repetition, partly because it is difficult to separate the effects of the two (Nation & Webb, 2017). As a result, past research has mainly focused on the effect of repetition without accounting for the effect of context as a potential confounding factor. Yet, a small number of studies have shown that contextual information can help learners infer the meaning of words (e.g., Teng, 2016; Webb, 2008). Presumably, the presence of meaningful contextual cues may enhance the relevance and saliency of a collocation. The limited research on context effects in MWUs has primarily focused on the understanding and/or learning of idioms (e.g., Ackerman, 1982, Cooper, 1999; Liontas, 2003). Since collocations can also be relatively semantically opaque (e.g., "win respect") (Taylor, 2004), context may also play a role in understanding and acquiring them.

The following section describes different approaches to context effects in vocabulary acquisition, including its role in incidental vocabulary learning.

6.1.2. Context and incidental vocabulary acquisition

Prior research has adopted different approaches to study the effect of context (Haastrup, 1989; Mondria 2003; Mondria & Wit-de Boer, 1991; Paribakht & Wesche,1999). From the perspective of Haastrup (1989), encountering words in less supportive contexts incurs more cognitive effort, increasing the chances of lexical retention. This is based on the findings of Jacoby et al. (1979) and Cairns et al. (1981), who found that decision-making difficulty leads to a stronger memory trace (Nation, 2001). Similarly, de Bot et al. (1997) argued that informative contexts are effective for initial meaning-guessing, but not for retention. Words surrounded by rich contextual cues are easy to guess, but also easy to forget because the word and its properties are processed relatively quickly, resulting in poor retention (Pulido, 2009). Previous research backs up this theory (e.g., Mondria & Wit-de Boer, 1991; Hu & Nassaji, 2012; Hulstijn & Laufer, 2001; Zahar, Cobb, & Spada, 2001). Mondria and Wit-de Boer (1991) examined word learning and retention using specially constructed, isolated sentences that included a range of contextual clues. They found that successful guessing did not guarantee retention. Similarly, Hu and Nassaji (2012) found a negative effect of ease of inferencing on vocabulary retention suggesting that word retention is poor when inferring the meaning is easy.

Relatively few studies have examined the role of context in incidental vocabulary learning (Nation, 2013). Previous research has focused on different types of context (short vs long) and measuring methods, making it difficult to assess the effect of context on incidental word learning (Teng, 2016). For example, Webb (2008) used short contexts (one or two sentences) that were rated according to his context specification 4-point scale on informativity (see Webb, 2007a). He asked 50 Japanese EFL learners to read three sets of 10 sentences, each sentence containing 1 of 10 target words. Learners were then administered four tests that measured recall of form, recognition of form, recall of meaning, and recognition of meaning. The results showed that more informative contexts resulted in greater gains in learning word meaning, but context had little effect on the acquisition of form. Teng (2019) also found that contextual richness led to the acquisition of word meanings using short texts, but results also indicated a significant effect for learning word forms. This disparity in results is most likely due to Teng (2019) including three different frequencies of word exposure (1, 5, and 15 repetitions). The findings of

Webb (2008) and Teng (2019) suggest that any facilitation associated with richer contexts could mean that words in uninformative contexts are ignored, or that the meanings inferred from uninformative contexts are less accurate (Hulstijn et al., 1996).

Other researchers have used longer texts, such a short stories, employing Beck et al.'s (1983) contextual support model alone or in combination with other techniques. Zahar et al. (2001) first measured the degree of contextual richness of the immediate context of 30 target words using Beck et al.'s model. They found that rich and directive contexts did not consistently lead to greater gains, leading them to argue that words are best learnt through variable contexts, i.e., a mix of contexts containing both more and less informative cues. According to Zahar et al. (2001), one mechanism for word learning involves "an unclear or semi-clear context open[ing] up a learning need, or conceptual gap, which is then reactivated when the word is eventually met in a clear context" (2001: 556). Past research has used graded readers (Horst, Cobb, & Meara, 1998), short narrative and expository texts (Nagy et al., 1985; Teng & He, 2015), specially constructed passages (Jenkins, Stein, & Wysocki, 1984) and single sentences (Dempster,1987; Laufer & Shmueli,1997) (qtd in Webb, 2008).

In summary, the single-word learning studies reviewed above attempted to quantify how useful a context was for word learning, whether the measure was framed in terms of contextual richness, supportiveness, or informativity. In extending this research to collocation learning, it is important to be precise about how context is operationalised, as well as to control for aspects of the reading materials, such as text length, which can affect vocabulary learning (Laufer & Shmueli, 1997).

6.2. Aims and research questions

Presenting collocations in contexts with varying levels of context informativity will allow us to examine whether context influences the incidental learning of collocations. This experiment uses lower congruency collocations so that learners are more likely to infer meaning from context rather than relying on the L1.

Therefore, this experiment is guided by the following questions:

1. What is the effect of context informativity on the incidental learning of collocations through reading short texts?

2. Do learner- (prior vocabulary knowledge) and item-related variables (congruency, association strength, corpus frequency) affect the learning gains?

114

6.3. Methods

6.3.1. Participants

Ninety-four participants (64 female, 30 male) were recruited from a Spanish University. They were all native speakers of Spanish and were between the ages of 18 and 23 (M=20.21, SD=2.03). All participants were at least of an upper-intermediate or B2 English proficiency level in the CEFR, as they had all met this language level before university admission. All participants had formally studied English for at least ten years (M=11.47, SD=3.15), and they were undertaking the English-taught degree of English Studies. They had experience reading English academic texts on diverse topics such as English linguistics, English literature, or English stylistics. The reading materials used in this experiment were not related to any of the participants' academic background. To evaluate their prior vocabulary knowledge, three sections of the revised version of the VLT (Schmitt, et al., 2001) were administered. Participants scored an average of 27.03 out of 30 on the 3K level (SD=1.27). On the 5K level, the average score was 26.43 out of 30 (SD=2.50), and a mean score of 23.46 (SD=3.68) was revealed on the 10K level. The VLT average score out of 90 was 78.41 (SD=7.45). Participants self-rated their English listening, reading, writing, and speaking skills at pretest. All participants rated[14] their reading above 7 in a 10-point scale (MAX=10, MIN=7, M=8.88, SD=0.86). Participation was voluntary, and participants were entered into a raffle where they could win one of ten e-gift cards in compensation for their participation.

6.3.2. Design

A pretest-posttest-delayed posttest, Latin-square design was adopted. Two outcome measures, form recall and meaning recall, were measured at two delays: immediately after the treatment (immediate posttest) and two weeks after (delayed posttest), resulting in a total of four outcome variables.

---

[14] Only reading is reported as this is the relevant skill for the experiment.

6.3.3. Reading materials

Sixty-six short texts (80-100 running words) were extracted from a range of different sources: language corpora (COCA and BNC), medical resources (the *Collins dictionary [of] medicine*, Youngson, 2004; the *Mosby's medical, nursing, and allied health dictionary*, Anderson, Keith, and Novak, 2002; and *Stedman's medical dictionary for the health professions and nursing*, Stedman, 2005), as well as Internet medical resources (e.g., the NHS website for England and the John Hopkins medicine website). Sixty passages were used as target texts in the treatment (i.e., they contained the target collocations), while the other six were used as control passages (i.e., they contained control items).

To increase the comprehensibility of the passages, the frequency profile of the texts was controlled for using the VocabProfile of the Lextutor website (Cobb, 2015). Based on experience with learners with a similar background to the participants, a receptive knowledge of 4-5,000 words was assumed for the analysis. Prior conversations with the teachers of the participants confirmed this. All the texts had at least 95% lexical coverage using the 4,000 most frequent word families (BNC-COCA list) when accounting for cognate words that were initially identified as Off word list (e.g., virus - *virus*, trauma - *trauma*, gastrointestinal - *gastrointestinal*). A separate list was then compiled to assess the cognateness status of these words. Five native speakers of Spanish identified the items as cognates by means of a translation task (see Appendix 3A). Spanish-English cognates accounted for 8-12% of the texts and were thus highly likely to be known by the L1-Spanish participants due to their common origin, i.e., these items were semantically and orthographically and/or phonologically similar in English and Spanish. It was expected that the participants would have receptive knowledge of Graeco-Latin medical terms (e.g., electrocardiogram - *electrocardiograma*) as many are also cognates. Therefore, the texts were judged to be appropriate for the target population.

As in Experiment 2, the passages were presented one at a time (one text per screen) in Arial size 18. Participants were unable to return to previously read screens, and they were asked to read the passages only once (see Appendix 3A).

6.3.3.1 Context informativity

Context informativity in this experiment refer to the presence of contextual cues in the text that could be used to infer the meanings of the target items. For example, the terms "menopause", "body heat", "hormone levels", and "biological process" were identified as contextual cues for "hot flushes" in

one of the texts, because they are informative of the core meaning of this collocation. Put differently, the availability of these terms in the text strengthens the chances of guessing the meaning of "hot flushes" and inferring it from context.

To observe an effect of context informativity, three conditions were created: highly informative (high), mid-informative (mid), and low-informative (low). To establish the level of informativity of the selected passages, I first examined the texts in search of contextual cues so as to distinguish between the three conditions. Six passages for the high condition were taken from Experiment 2, and the other 54 passages were extracted from the abovementioned sources. To ensure that high and mid conditions differed in informativity, the total number of informative cues in the high condition was greater than in the mid condition. The context cues in the high condition were also thought to stimulate meaning-inferring that could help learners fine-tune the comprehension of the items. Hence, the high and mid informativity conditions differed both in terms of quantity and quality of contextual cues. As for the low informativity condition, a number of texts had to be slightly modified to ensure they remained contextually "unsupportive". The texts were modified with the assistance of two English native speakers who were also linguists. This involved replacing informative cues with vague, less informative ones or mentioning the item only in passing, i.e., embedding it in an uninformative context. Below are three examples of the passages used (high/mid/low) for the target item *bone marrow*. All passages are provided in Appendix 3A. What constitutes informativeness in each passage is indicated by underlying for clarification purposes.

Example of a high informativity text

The goal of this <u>transplant</u> is to <u>cure</u> serious blood diseases and types of <u>cancer</u> including leukaemia, lymphoma and myeloma. A **bone marrow** transplant takes healthy blood-forming <u>cells</u> from a <u>donor</u> and infuses those cells into the <u>patient's</u> bloodstream, where the cells may begin to grow and produce healthy red blood cells and white blood cells. The preparations for a transplant like this vary depending on the type of transplant, the <u>disease</u> needing transplant, and your tolerance for certain <u>medicines</u>.

Example of a mid informativity text

This year, more than 130,000 Americans will be <u>diagnosed</u> with it. In 2020, 4,864 unrelated and 4,160 related **bone marrow** <u>transplants</u> were performed in the United States. Although millions of

people are already included in the global <u>donor</u> registry, it still lacks ethnic and racial diversity and is of limited benefit to <u>patients</u> of non-Northwestern European descent. This makes it harder for some patients to get the <u>treatments</u> they need. The global registry hopes to increase its diversity by recruiting through social media and community outreach.

Example of a low informativity text

These <u>procedures</u> aren't as rare as you might think. In fact, surgeons perform nearly 30,000 of these procedures each year in the US. The majority are kidney <u>interventions</u>, followed by **bone marrow**, liver, heart, and lung surgical procedures. However, recent data shows that the most frequent procedures performed in the US are now <u>blood</u> transfusions, respiratory intubation and ventilation. While blood transfusions are a very common medical procedure, the increased demand of invasive and non-invasive ventilation was triggered by the COVID-19 outbreak.

As a first step for validation, 5 English native speakers were asked to read the passages and highlight the words they considered to be contextual cues of the target items. They were shown randomised passages (high, mid, low) so that they only encountered the target collocations in one of the conditions. This first stage allowed me to confirm my own intuitions about the level of informativity of the selected texts.

A second stage involved conducting a norming study to determine an adequate and consistent level of informativity within each informativity condition. Three different counterbalanced surveys were distributed among 24 raters. This was done to prevent raters from comparing varying degrees of informativity for the same target items. All raters were native speakers of English and had a background in (applied) linguistics. They were asked to rate from 1 (extremely unlikely) to 5 (extremely likely) how easy or difficult it was to infer the meaning of the target items from the passages (see Appendix 3A). The scale was adapted from Webb (2008) for the present experiment. Eight passages (out of 60) were modified as they showed ratings that were too similar between the low/mid or mid/high conditions. The modified passages were then rated by a different pool of 6 English native speakers and were found to be consistent. An inter-rater reliability check was carried out to find how much agreement there was between the raters. The raters agreed on 457 (out of 480) of the cases (95%). The mean rating for the high condition was 4.8, followed by 3.02 in the mid and 1.58 in the low condition. It is important to note that raters were not as extreme in their low ratings (1s and 2s) compared to the high ratings (almost

always 5). This may be because even in the low informativity condition, the passage had to cohere to some extent for the target items to be used felicitously. It would therefore have been possible to get a vague sense of the meaning of a collocation embedded in a text without knowing exactly what it meant. The passages were informally rated for naturalness by two native speakers of English. Minimal changes had to be made to ensure they sounded natural.

6.3.4. Collocations

Twenty target collocations were selected for this experiment. As the participants were university students undertaking the English-taught degree of English Studies, I decided to focus on subject-specific collocations related to medicine, as they were not related to their field of studies. The item selection started from checking the medical resources mentioned in the previous section. A few items from the Experiments 1 and 2 were added to the initial list of potential target items. This final list consisted of 15 noun-noun and 5 adjective-noun lexical collocations (see Table 23). The target items varied in their degree of specialisation; some were non-specialised collocations (e.g., "stomach bug") and others technical collocations (e.g., "tooth decay"). This variation in degree of specialisation corresponds to what learners are likely to encounter in natural reading of academic and non-academic input.

Raw corpus frequencies were extracted from the COCA. All items had a relatively high log Dice scores (all above 5.00). Each target collocation appeared once per text. The target items were placed roughly in the same position (second or third sentence) without any enhancement.

The items varied in their degree of congruency, though they were all lower congruency items. To establish their congruency status, the meanings of each collocations' components was first checked in the Oxford English-Spanish dictionary (Rollin, 2008) in order to decide if a word-by-word translation was a plausible collocation in Spanish. This stage served as confirmation of intuitions about the congruency of the collocations that were initially considered as target items for the experiment and to remove any congruent ones. This experiment focuses on lower congruency collocations so that learners are more likely to infer meaning from the contextual support of the reading materials rather than using the L1. High congruency items were thus discarded from the target list. Congruency levels for the remaining items were then established in a small-scale norming experiment. Eight Spanish native speakers took part in the experiment. They were all English-Spanish translators with university degrees

in Translation and at least 5 years of experience as professional translators. Ratings were collected using a web-based survey. The participants were presented with a congruency task where they had to rate the congruency status of the items on a scale from 1 (incongruent) to 3 (congruent). An explanation of congruency was provided at the beginning as well as examples of incongruent and congruent items in English-Spanish (see Appendix 3A). Participants were also asked to translate the collocations into Spanish. This allowed me to confirm the possible translations of the items and accept the most and least technical ones in the actual testing phase (e.g., motion sickness can be translated into *cinetosis* (technical term) or *mareo por movimiento* (common usage)).

Table 23. List of target collocations and their properties.

Collocation	Spanish equivalent	Congruency	COCA	log Dice
hot flushes	sofocos	1.12	74	7.9
cold sore	herpes	1	105	6.6
heat rash	sarpullido	1.25	42	5.9
dull ache	dolor sordo	2	22	8.72
sore spot	llaga	1	229	6.72
bone marrow	médula ósea	2.75	2317	11.3
stretch marks	estrías	1	334	11.3
tooth decay	caries	1.25	267	10.0
hay fever	alergia al polen	1	252	10.8
general practitioner	médico de cabecera	1.37	816	9.81
flesh wound	arañazo	1.12	169	7.42
eye strain	fatiga visual	1.37	74	5.24
labour pain	contracciones	2.12	122	5.2
shoulder blade	omóplato	1.25	286	11.72
stomach bug	virus estomacal	2.37	67	7.23
outpatient care	atención ambulatoria	1.12	146	10.92
motion sickness	mareo por movimiento	1.25	482	6.13
withdrawal symptoms	síndrome de abstinencia	1.37	354	9.24
rib cage	caja torácica	1.75	807	11.08
shin bone	tibia	1	149	10.05

6.3.5. Measures

6.3.5.1. Pretests

A number of learner-related variables were measured to assess their contributions, if any, to learning success. This information was collected in the pretest assessments conducted two weeks before the treatment. In this first session, all participants signed the consent forms, filled out a slightly modified version of the LEAP-Q (Marian et al., 2007) and completed the 3K, 5K, and 10K sections of

the revised version of the VLT (Schmitt et al., 2001) as well as the pretests on the target collocations (form recall and meaning recall) and a distractor task so as to avoid potential testing effects. The LEAP-Q and the revised version of the VLT used in this experiment are the same as in Experiments 1 and 2 (see Chapter 4, Section 4.3.5.1). The testing sequence was also designed to avoid testing effects (see Section 6.3.6).

The pretest on the target collocations consisted of a form recall and a meaning recall tasks. The form recall test format was the same as the immediate form recall tests used in Experiments 1 and 2. This pretest used a translation format in which the L1 meanings cued the responses. Participants were provided with the first letter of each component word of the collocations to avoid elicitation of other word pairs and asked to provide the L2 collocation. The meaning recall test was a translation task in which the L2 forms cued the L1 responses. Participants were asked to provide a Spanish translation, an equivalent (e.g., a synonym), or a definition of the L2 form.

A relatively short distractor task was included after the postreading questionnaire to further divert the participants' attention away from the task at hand. The distractor task was taken and slightly adapted from a website of the International House World Organisation (*https://www.ihbristol.com/free-english-exercises/test-esol/esol-wf-cctv/submitted*), where it was suggested as an activity for upper-intermediate learners of English. This task was a multiple choice test where learners had to select the conversational phrases that best matched a short dialogue. This task is available in Appendix 3A.

6.3.5.2. Immediate posttests

Immediately after the treatment, participants took the same form recall test they took at pretest followed by a postreading questionnaire, which contained six questions that tapped into the learners' overall comprehension of the texts as well as their perception of difficulty, general engagement, reading habits, and preferences (i.e., whether they preferred reading short or long texts, or both), and the use of vocabulary strategies when they read in English (e.g., guessing from context). This questionnaire used a series of rating scale and multiple-choice questions (see Appendix 3A). Two rating scale questions tapped into the learners' perceptions of difficulty of the texts. A multiple-choice question was used to check for overall content comprehension, as checking for understanding of every single text was logistically unworkable. This question was employed to ensure that learners had paid to the content of the texts. They had to tick the topics they had encountered in the texts from a list of 10 medicine-related topics, where only 5 were correct. Identifying the right topics was a very easy task. This was

confirmed in the piloting. Then, participants completed the same distractor task they had taken at pretest before filling out the meaning recall posttest, which was identical to the one used at pretest, with the items re-randomised.

6.3.5.3. Delayed posttests

The same form recall and meaning recall immediate posttests were administered two weeks after the treatment (delayed posttests), with the items were re-randomised for each test. Participants took the form recall posttest first, followed by the distractor task so as to prevent them from memorising the items. Then, participants completed the meaning recall posttest, and, finally, they filled out a simplified version of the retrospective protocol used in Experiment 2 (see Appendix 3A).

All tests (pre- and posttests) contained distracters (15 per test) to draw participants' attention away from the target items. The distracters were chosen by researching the English core textbooks that participants were using as part of their studies, as well as consulting with the teachers. The distracters were thematically-related to the target items (e.g., "period pain", "birth control", "hearing loss", etc.) but considered substantially easier relative to the target ones.

6.3.6. Procedure

Participants completed the 3 components of the experiment over 3 sessions, separated by 2 weeks each (see Table 24). All parts were taken online via SoSci Survey, the same web-based experiment presentation platform used in Experiments 1 and 2 (Leiner, 2022, Version 3.4.12). The final dataset comprised 1,880 observations, from 94 participants to 20 items, with 4 outcome variables.

Table 24. Data collection procedure.

Week	Procedure
1	Informed consent forms
	Background questionnaire
	Form recall pretest
	Vocabulary test
	Distractor task
	Meaning recall pretest
3	Treatment
	Immediate form recall posttest
	Reading comprehension test
	Distractor task
	Immediate meaning recall posttest
5	Delayed form recall posttest
	Distractor task
	Delayed meaning recall posttest
	Retrospective questionnaire

The experiment was first piloted with 12 L1-Spanish EFL learners with a similar language level and background to that of the participants of the experiment, using a superset of the eventual target items. This pilot experiment was carried out to maximise the chances that the actual participants of the experiment would not already know the target items. Participants of the piloting were asked to fill out all the pretests (LEAP-Q, VLT, form recall, and meaning recall). The two items that were found to be relatively well known were removed from the list of target items and incorporated as distracters (i.e., "runny nose", "wasp sting"). The final 20 target items were those that were clearly unknown by the learners in the piloting. The distractor task was included upon the realisation that some participants in the piloting were able to recall some target items from the form recall task when they completed the meaning recall task. The number of distractors was also increased (from 10 to 15).

6.3.7. Data processing and model fitting

6.3.7.1. Exclusion criteria

To start with, the answers to the postreading comprehension test were checked for all the learners. The threshold of adequate comprehension was set at identifying at least 4 of the 5 topics correctly, and selecting no more than 7, as this would have indicated that the learner had either not paid attention to the texts or was blind guessing the answer. None of the participants had to be excluded as they were able to identify the right topics. The learners' overall understanding of the texts was also checked to evaluate whether they had struggled with understanding the reading materials. All learners reported understanding the overall content of the texts and a majority (88 out of 94 participants) reported understanding the main ideas and most details. Additionally, the responses to the perception of difficulty of the texts revealed that all the learners were at ease with the relatively complexity of the academic texts used in this experiment. Therefore, it was assumed that participants had no trouble understanding the passages.

Participants' answers to the distractor task were not analysed as conversational phrases are not of interest in this experiment.

6.3.7.2. Scoring responses

The outcome variables were binary (correct response=1, incorrect response=0). Responses in the form recall tests that contained minor spelling mistakes were scored as correct as long as the response was comprehensible (e.g., "bone *marro*"). This decision was made following previous studies (Sonbul & Schmitt, 2013).

6.3.7.3. Model fitting

Before analysing the data, the predictors (congruency, preknowledge of collocations, prior vocabulary knowledge, corpus frequency, and association strength) were assessed for correlations among variables (see Appendix 3A). To avoid collinearity in the models, only congruency and prior vocabulary knowledge were used in the analysis.

As in Experiments 1 and 2, the data for the present experiment was analysed using the R (R Core Team, 2022). Unaggregated data were fitted to logistic mixed-effects regression models (Baayen et al., 2008; Jaeger, 2008). Four separate models were fit for each combination of outcome measure (form recall, meaning recall) and delay (immediate posttest, delayed posttest). Each model included

fixed effects of all the potential explanatory variables: informativity, item-related factors (congruency), two-way interactions with congruency, where the interaction would have been interpretable, and learner-related variables (prior vocabulary knowledge). Congruency and prior vocabulary knowledge were centred, and informativity was Helmert-coded (contrast 1: low vs mid; contrast 2: low and mid vs high). To be able to compare the magnitude of effects across the form models, the immediate and delayed models had the same set of predictors for each outcome measure.

Models for each of the outcome measures were selected using the same procedure as for Experiments 1 and 2, with fixed effects determined first, and using the maximal random effects structure supported by the data.

6.4. Results

6.4.1. Descriptive statistics

Table 25 shows the average absolute gain scores for each informativity condition and for each outcome measure in the immediate and delayed posttests. Gains at posttest were measured relative to performance at pretest. The meaning recall posttests yielded higher scores relative to form recall, which suggests that eliciting the meaning of a collocation is easier than recalling its form. With regard to context informativity, the mean scores reveal that performance is consistently higher in the mid and high context informativity conditions relative to the low informativity condition, and interestingly that higher scores are consistently revealed for the mid in comparison to the high condition. In terms of testing delay, it can be seen that performance on the delayed posttests slightly improved for form recall in all conditions but not for meaning recall.

Table 25. Average absolute gains by informativity condition and outcome measure.

| | | Form recall | | Meaning recall | |
		Pre-Post M (SD)	Pre-Delay M (SD)	Pre-Post M (SD)	Pre-Delay M (SD)
Informativity	Low	1.31 (0.10)	1.92 (0.10)	2.90 (0.12)	2.52 (0.12)
	Mid	1.79 (0.13)	2.32 (0.13)	4.59 (0.13)	4.29 (0.14)
	High	1.73 (0.11)	2.25 (0.12)	4.05 (0.14)	3.45 (0.11)

Note: Maximum= 20. Pre-Post refers to gains accrued from pretest to immediate posttest, Pre-Delay refer to the gains from pretest to delayed posttest.

6.4.2. Models

6.4.2.1. Form recall models

As Table 26 shows, prior vocabulary knowledge is a reliable predictor at immediate posttest. That is, a larger vocabulary size (as measured by the VLT composite score) increased the likelihood of a correct form recall response. Congruency also came as a predictor in the model, although its effect was only marginal. With regard to context informativity, the first Helmert contrast (low vs mid informativity) shows that the mid informativity condition had a positive effect on immediate form recall relative to the low informativity condition. The second contrast (low and mid informativity vs high), however, was not reliable - there was no increase in immediate form recall performance for the high informativity condition relative to the low and mid conditions.

Table 26. Summary of fixed effects for immediate posttest form recall.

Model:				
Response ~	Informativity + VocabSize + Congruency			
	+(1+Congruency|ParticipantID) + (1+Congruency|ItemID)			

Predictor	**Estimate**	**SE**	**z value**	**Pr (>|z|)**
Intercept	-2.56	0.78	-3.26	0.00
Inf.Helmert1	0.39	0.14	2.69	0.00
Inf.Helmert2	0.17	0.12	1.44	0.14
Vocabulary	0.21	0.07	3.03	0.00
Congruency	0.91	0.55	1.65	0.09

Looking at the same measure at delayed posttest, prior vocabulary size and the first Helmert contrast (low vs mid) remained significant (see Table 27), whereas the second contrast (low and mid informativity vs high) did not have a significant effect. In contrast to the immediate model, congruency did not come out as a predictor at delayed form recall. This was not a surprising result, since the effect

of congruency was only marginal in the immediate model and considering that the items in the experiment had a small range of congruency. This finding will be discussed in the following section.

Table 27. Summary of fixed effects for delayed posttest form recall.

Model:				
Response ~ Informativity + VocabSize + Congruency				
+(1+Congruency|ParticipantID) + (1+Congruency|ItemID)				

Predictor	**Estimate**	**SE**	**z value**	**Pr (>|z|)**
Intercept	-2.70	1.19	-2.26	0.02
Inf.Helmert1	0.33	0.15	2.23	0.02
Inf.Helmert2	0.10	0.12	0.81	0.41
Vocabulary	0.41	0.08	5.00	0.00
Congruency	1.20	0.85	1.39	0.16

6.4.2.2. Meaning recall models

The immediate meaning recall model is presented in Table 28. In this model, prior vocabulary knowledge and the first contrast (low vs mid informativity) are significant predictors. On the other hand, the second contrast (low and mid informativity vs high) did not reach significance. This suggests that a higher level of informativity does not necessarily lead to better performance at recalling the meaning of collocations, as there was no increase in meaning recall performance for this condition. There was also no significant effect of congruency.

Table 28. Summary of fixed effects for immediate posttest, meaning recall.

Model:				
Response ~	Informativity + VocabSize + Congruency			
	+(1+VocabSize+Congruency|ParticipantID)+(1+VocabSize+Congruency			
	|ItemID)			

Predictor	**Estimate**	**SE**	**z value**	**Pr (>|z|)**
Intercept	0.00	0.57	0.00	0.99
Inf.Helmert1	0.96	0.13	7.13	0.00
Inf.Helmert2	0.13	0.11	1.19	0.23
Vocabulary	0.16	0.07	2.26	0.02
Congruency	0.47	0.37	1.26	0.20

Table 29 shows that the model of meaning recall at delay includes exactly the same predictors as the immediate model. That is, while prior vocabulary knowledge and the first informativity contrast (low vs mid) had a positive significant effect on recalling the meaning of collocations, the second contrast (low and mid informativity vs high) and congruency did not come out as significant predictors. This suggests that learners performed better at retaining the meaning of collocations in the mid informativity condition relative to the low and high informativity conditions. I will return to this pattern of effects in the Discussion section.

Table 29. Summary of fixed effects for delayed posttest, meaning recall.

Model:				
Response ~	InformativityHelmert1 + InformativityHelmert2 + VocabSize + Congruency			
	+(1+Congruency|ParticipantID) + (1|ItemID)			

Predictor	**Estimate**	**SE**	**z value**	**Pr (>|z|)**
Intercept	-0.40	0.63	-0.62	0.53
Inf.Helmert1	1.09	0.13	8.09	0.00
Inf.Helmert2	0.00	0.12	0.02	0.98

Vocabulary	0.26	0.05	4.74	0.00
Congruency	0.49	0.42	1.17	0.24

6.4.2.3. Summary of the models

Table 30 summarises the results of the present experiment showing the significance of the main predictors (the first Helmert contrast, the second Helmert contrast, prior vocabulary knowledge, and congruency) in the models.

Table 30. Summary of model results.

	Form recall		Meaning recall	
Predictor	Immediate	Delayed	Immediate	Delayed
Informativity contrast 1	+	+	+	+
Informativity contrast 2	-	-	-	-
Vocabulary	+	+	+	+
Congruency	marginal	-	-	-

Note: Plus signs (+) indicate a significant effect, minus signs (-) show no significant effect.

The word marginal indicates that the effect was significant at .1 level.

6.5. Discussion

6.5.1. The effect of contextual support

The first aim of this experiment was to test the effect of context informativity on incidental collocation learning. The results show that context informativity had a significant effect on recalling the meanings and the forms of collocations, both at immediate and delayed testing. The present experiment operationalised context informativity as the presence of context cues (+/- informative) in a text that learners could use to derive the meanings of collocations. Interestingly, the level of informativity improved performance only to a certain extent. Specifically, the first informativity contrast (low vs mid) was reliable for all four models, but the second contrast (low and mid informativity vs high) did not reach significance. This indicates that participants performed better at learning collocations embedded in

semi-informative contexts relative to highly informative ones. One plausible interpretation of this finding is that the mid informativity passages required greater cognitive engagement when compared to high informativity passages, prompting learners to notice the collocations, which in turn led to more sizeable gains (Haastrup, 1989). In contrast, encountering a collocation in a context that is highly informative about its meaning might have been less effective in prompting learners' cognitive effort. From the perspective of the Involvement Load Hypothesis (Hulstijn & Laufer, 2001) and related depth-of-processing frameworks (Anderson & Reder, 1979; Craik & Lockhart, 1972), cognitive effort, understood as the amount of cognitive capacity that learners allocate to learning, is needed to decode a word and process its meaning. As meaning-inferring did not require much effort in the high informativity condition, learners may have devoted their cognitive resources to other aspects of the vocabulary or on the content of the text. This interpretation is supported by the results of previous single-word studies, which have shown that words embedded in helpful contexts are easily forgotten due to the possibility of processing their properties superficially (Pulido, 2009).

Another possible explanation for the significant effect of the mid informativity condition relative to high informativity can be found in its operationalisation. In the mid informative texts, the collocations were surrounded by semi-informative context cues, i.e., cues that were not entirely explicit (or as informative as in the high informativity texts) nor entirely vague (as in the low informativity texts). Take for example the mid informativity passage used for the item "bone marrow":

Example of mid informativity text

This year, more than 130,000 Americans will be diagnosed with it. In 2020, 4,864 unrelated and 4,160 related **bone marrow** transplants were performed in the United States. Although millions of people are already included in the global donor registry, it still lacks ethnic and racial diversity and is of limited benefit to patients of non-Northwestern European descent. This makes it harder for some patients to get the treatments they need. The global registry hopes to increase its diversity by recruiting through social media and community outreach.

As can been seen, the context cues available (e.g., "diagnosed", "transplant", "donor") provide a few helpful clues that might be supportive of successful guessing, however, they are not extremely informative or explicit about the meaning of the collocation. Additionally, the text does not contain

signalled synonyms within the context (e.g., "cancer") or contextual cues that could be considered more semantically salient (e.g., "bloodstream", "cells") - cues that were available in the high informative text. This operationalisation might have revealed a lexical learning need or a conceptual gap (Brown, 1993) which was activated as learners exerted more effort into deriving the meanings of collocations. This aligns well with Zahar et al.'s (2001) claim that middle ground informative contexts, i.e., relatively unclear or semi-clear, are most likely to promote word learning, particularly if the word is encountered again in a more helpful context. The results from this experiment align well with findings from single-word studies that show word learning improvement through texts that provide quality context cues to infer the meanings of new words (e.g., Mondria & Wit-de Boer, 1991; Webb, 2008). They also lend support to Beck et al.'s (1983) claims that context is an important source of information about word meanings and that the informativity of a context is text-dependent. Altogether, it seems that the role of context informativity is more nuanced than initially thought. Future research should thus consider how context informativity affects incidental collocation learning in order to determine its effect in acquiring different collocation knowledge aspects (form recall, meaning recall, form recognition, meaning recognition).

The operationalisation of context informativity in this experiment was related to learning the meaning of the collocations, but it was not expected to affect learning of form to the same extent: learners encountered the target items only once, which is not enough to incidentally gain knowledge of form to the recall level (e.g., Vilkaitė, 2017). However, context did also improve form recall in this experiment. This finding is incongruent with Webb (2008), which showed that context had little effect on acquiring the form of L2 words, but coincides with Teng (2019), who found that context improved form recall. The reason for these conflicting results in single-word research can be attributed to the repetition manipulation in Teng's study, who included 1, 5, and 15 encounters with the target items, suggesting that the number of repetitions with the words rather than the context may have a greater effect on gaining knowledge of word forms.

While context might be less relevant for acquiring single word forms, this may not be the case for learning collocations, as many carry specific semantic meanings that are non-literal or non-compositional (e.g., "hot dog"). It could be argued that the special semantic bonds between collocating items focuses learner's attention on a collocation's contingency (the association between its form and meaning), especially when there are cues in the immediate context of the unknown item. For example,

if a learner encounters the word "bone" in "bone marrow" in a context where it is important to understand what "bone marrow" as a whole means, the chances of the learner putting effort into guessing the meaning of this item increase, potentially drawing their attention to its form-meaning link. This may be particularly true for collocations composed of words that learners already know (i.e., partial knowledge), as they may devote more attention to deriving the meaning of the newly encountered combination. The more the meaning of a lexical item is needed, the greater the likelihood that learners will also focus on its form. This might be particularly true for collocations embedded in contexts which have useful context cues, as shown by the absolute gain average scores on the form recall posttests in the mid and high informativity conditions (2.322 and 2.255 items learnt out of 20, respectively). Additionally, the fact that participants probably had partial knowledge of some of the target items due to frequency might have also affected their lexical engagement. For example, one of the participants reported paying special attention to the item "flesh wound" because she had noticed the word "wound" being frequently used on various English-language medical TV series she enjoyed watching on her spare time.

Turning to the question of collocation retention, this experiment included delayed posttests to assess the extent to which immediate gains were retained over a two-week interval. The goal of administering tests two weeks after the learning session was to prevent participants from remembering the target items. The distractors were also kept in the delayed posttest so as to make it more challenging. Average scores in the meaning recall delayed posttest were slightly lower compared to the immediate posttest, suggesting that the two-week interval might have weakened the treatment effects. Interestingly, the lowest collocational "loss" was recorded in the mid informativity level (-0.301) relative to the high and low levels (-0.596 and -0.383, respectively), although these differences are very small. On the other hand, participants performed slightly better on the form recall delayed posttest relative to the immediate posttest. Though small, the increases on the form recall delayed posttest suggest that testing itself may be contributing to learning/retention. This pattern has been observed in previous studies (Peters & Webb, 2018; Webb et al., 2013, Webb & Chang, 2020).

A closer evaluation of the form recall responses revealed that participants' performance improved when recalling collocations for which they had shown partial knowledge at immediate posttest. For example, where participants produced infelicitous collocations at immediate testing - some induced by the L1 (e.g., "movement sickness" instead of "motion sickness") - they were able to supply the correct word combination on the delayed posttests. This has two possible explanations. It is possible that

participants had remembered and looked up the items between the immediate and the delayed posttests (Nation & Webb, 2011), although they reported not having done so in the qualitative questionnaire. Some learners did report attempting to remember some of the items. Another interpretation is that this pattern of effects reflects the gradual gains that are made through exposure to meaning-focused input and indicates increased collocational awareness. This is supported by the fact that learners did not seem to fall into the trap of deceptive L1-L2 collocational equivalence (Laufer, 2011).

6.5.2. Form recall vs meaning recall

The descriptive statistics of the form and meaning recall posttests show that participants learnt between 1.3 (lowest) and 1.7 (highest) out of 20 items (9.5 %) at immediate form recall, and between 2.9 and 4.5 collocations on the immediate meaning recall posttest. Regarding collocational retention, the range of scores was slightly higher at delayed form recall, between 1.9 and 2.3 items out of 20 (or 8.5% and 11.5%, respectively), but lower for meaning recall, between 2.5 and 4.2 (12.5% and 21%). Low gains at the level of form recall were expected because learning the L2 forms of 20 items from a single exposure is quite a challenging task.

It is not easy to compare the gains registered in the present experiment to the gains in previous studies. Firstly, to the best of my knowledge, this is the first attempt to look at effect of context on the incidental acquisition of collocations. Second, differences in terms of the reading materials (e.g., text type, length), treatment variables (e.g., mode of input, repetition), and learner characteristics (e.g., L2 proficiency, institutional level) seem to have led to variation in learning gains across incidental collocation acquisition studies (e.g., Dang et al., 2022; Vilkaitė, 2017; Webb et al., 2013; Webb & Chang, 2020). However, some patterns can be observed. For example, these findings are inconsistent with Webb et al. (2013), who also showed that gaining productive knowledge of form was more difficult relative to meaning. Crucially, however, Webb at al. (2013) used a reading-while-listening task as a treatment. As previously seen in this thesis, assisted reading might lead to greater incidental collocation learning (Vu & Peters, 2022a). Therefore, Webb et al.'s (2013) findings might be not generalisable to reading-only studies. It is also important to note that Webb et al.'s (2013) learners experienced greater quantity and density of input, as they read four short stories, while the learners in this experiment read a series of shorter academic texts.

When looking at the studies that have also focused on academic input, other patterns emerge. For example, both Dang et al. (2022) and Vilkaitė (2017) registered a low level of collocational learning. Vilkaitė (2017) found that L1-Lithuanian university students acquired (and retained) knowledge to the form recognition level of 1.69 out of 15 adjacent collocations (12.6%), but she found that a majority of the participants did not acquire the form of any of the items to the recall level. The target items in that experiment were incongruent verb-noun collocations, which may have affected their learnability (e.g., Altenberg & Granger, 2001). In contrast, the present experiment focused on adjective-noun and noun-noun collocations, which show less morphological variation, yet still pose a challenge for L2 learners. More recently, a study by Dang et. al., (2022) found that on average, L1-Chinese university students acquired 4.88 out 19 collocations (25.68%) through reading a transcript of an undergraduate academic lecture. This experiment, however, only measured knowledge at the form recognition level.

Knowledge to the recall level is one of the most difficult vocabulary components to acquire (Barcroft, 2002), and it is even more difficult in an incidental approach (Schmitt, 2010). The results of the present experiment are thus encouraging, considering that acquiring collocations is particularly challenging for L2 learners as they need to learn the forms of two individual words. The findings indicate that incidental collocational learning can enhanced by optimising context effects. In this experiment, context informativity led to better chances of initial correct meaning-guessing and improved collocation retention, dependent on the informativity level. This may reveal a distinctive advantage over the texts used in previous studies, in which context informativity might have been insufficient for learners to call up the target collocations and derive their meanings. It should be noted, however, that the high and mid levels of informativity in this experiment were designed to make the meaning of the collocations inferable, particularly in the high informativity condition. Admittedly, the characteristics of the participants (L1-Spanish EFL learners) might have also facilitated the acquisition of the medical-related collocations used as target items in this experiment. Partial knowledge might also explain some of the collocation gains, although it was not addressed due to the difficulty of defining what partial knowledge is, and only full scores were given for each learnt item. Although gaining partial knowledge was not formally examined, a closer examination of the responses revealed that learners who had partial knowledge of specific items at pretest performed better on these items on the meaning recall posttests, although it was also found that some learners were unable to provide the exact meaning of the item. This suggests that learners were drawing on the meaning of the collocation's constituent they knew.

For example, some learners were able to partially recall the meaning of the item "stomach bug" as a stomach-related illness, but they failed to provide an equivalent in Spanish. This suggests that they had some general knowledge of the meaning of this collocation being related to health and this body part, but they did not know its exact meaning. Ideally, future studies might want to include scores for partial and full knowledge of collocations so as to bring a more accurate assessment of the learning gains (see Nation & Webb, 2011). In this respect, the participants' responses on the form recall tests indicated that some were able to partially recall the word forms of the target collocations, particularly at immediate delay. I first interpreted this as an indication of participants having partial knowledge of semi-congruent collocations (e.g., "labour pain", "stomach bug"). However, an informal analysis of the responses revealed that this was not necessarily the case, as learners produced errors in recalling items whose partial congruency score was moderately high. For example, the item "bone marrow" showed the highest rating (2.75 out of 3) in terms of L1-L2 congruency in this experiment, but participants did not do particularly well in recalling its form (average 0.17 from pretest to delayed posttest).

6.5.3. The effect of learner- and item-related factors

Prior vocabulary knowledge has been found to have a positive impact on the incidental learning of collocations through reading (e.g., Vilkaitė, 2017; Vu & Peters, 2022a). The results of the present experiment also revealed that a larger prior vocabulary knowledge (as measured by the VLT composite score) increased the likelihood of collocations being incidentally learnt. This effect may be explained by the relationship between vocabulary knowledge and reading comprehension (Laufer & Ravenhorst-Kalovski, 2010). A learner with a larger vocabulary may devote fewer attentional resources and processing effort to understanding L2 input, making it easier to detect new vocabulary (Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011). Lower proficiency learners are more likely to spend a greater amount of time trying to make sense of the text, at the expense of noticing unfamiliar words and collocations.

It is possible that the L1 of the participants aided reading comprehension, thereby further supporting the positive effect of prior vocabulary knowledge. As previously mentioned in this chapter, knowledge of Graeco-Latin vocabulary may have played a significant role in the learning of the medical-related target collocations in this experiment. The analyses of the vocabulary profile of the texts revealed that Spanish-English cognates accounted for 8-12% of the total running words. The utility of

cognates has important implications for reading comprehension and vocabulary learning, as shown in prior research (e.g., Lubliner & Hiebert, 2011; Ramírez, Chen, & Pasquarella, 2013). An interesting feature of Spanish-English cognate pairs is that low-frequency English words tend to be higher-frequency words in Spanish (e.g., impose and *imponer*). Thus, cognate awareness is particularly useful for learning low-frequency English vocabulary by L1 Spanish learners (Chen, Ramirez, Luo, Geva, & Ku, 2012).

The findings suggest that congruency only played a marginal role in the immediate form recall posttest, whereby incongruent items seem to be slightly more difficult to learn than partially incongruent ones. The lack of a congruency effect in this experiment is expected given the smaller range of congruency for the target collocations, which focused on partially incongruent and incongruent collocations only, whereas previous studies have typically included both congruent and incongruent collocations (e.g., Vu & Peters, 2022a; Yamashita & Jiang, 2010).

The inclusion of a qualitative questionnaire was aimed at gathering qualitative information to the findings of this experiment as well as identifying patterns in the learners' preferences and reading habits. An informal analysis of the responses revealed general patterns across participants. For example, 80% of the learners indicated that they preferred reading short texts (e.g., a short story, a piece of news) over long ones (e.g., a novel). The rationale of using short texts (80-100 running words) as reading materials was based on informal conversations with the teachers and the participants of the piloting experiment. Arguably, the length of the texts might have had an effect on collocation learning as this factor has been shown to influence incidental word learning from context (Laufer & Shmueli, 1997). Using short texts may suggest to the learner that contextual cues are present in the immediate environment of the collocation, even within the same sentence as the collocation. It is possible that guessing from context was encouraged by this inference of immediacy, in contrast to longer texts, which could have required learner to search for cues in the wider context.

6.6. Limitations

While every effort was made to ensure that the experiment was rigorous and methodologically sound, there are some issues that future studies can address. Some of the limitations of the present experiment can become the starting point for future research. Firstly, it is important to acknowledge that context manipulation in this experiment relied on human ratings to assess the level of informativity of

the texts. This is also the case of previous single-word studies looking at context (e.g., Beck et al., 1983; Teng, 2016; Webb, 2008). While I attempted to have consistent informativity ratings, assessing the informativity of contexts in this manner is necessarily subjective, and relies on a range of factors such as the perceived saliency of the cues, its explicitness, and an individual's familiarity with the topic (Nation, 2001). This was confirmed in the norming study, where a few texts had to be modified due to differences in ratings from person to person. A larger norming sample could help to even out this variability.

Importantly, this experiment is limited in interpreting differences between passages for a number of reasons. Firstly, some of the differences could have been due to differences in naturalness and cohesion of the stimuli materials, rather than their informativeness, which may have affected the pattern of results. All passages were proofread for naturalness, but this process was not systematic. Future studies should include a measure of naturalness, e.g., by using rating scales. Another important caveat is that cohesion was not controlled for. Therefore, if participants were trying to figure out a referent that was unclear, they may have less attention for other elements of the text. Future studies should construct more clearly differentiated informativity conditions. An alternative operationalisation would be using the same vignettes in all three conditions (via script orders) and then systematically either adding or removing contextual clues).

The degree of specialisation of the collocations varied from non-specialised to more technical items. This may have affected the gains, although including different types of collocations is more ecologically valid (Dang et al., 2022). Future studies could examine other types of collocations such as verb-noun collocations, as research suggests that they pose different difficulties for learners (Peters, 2016).

Testing effects could be minimised by having participants perform different tasks between the form recall and the meaning recall posttests, although participants would not be able to serve as their own controls as in the current experiment. An alternative way to prevent testing effects could have been to include two experimental groups so that one is tested for form recall and the other for meaning recall (e.g., Peters & Webb, 2018) or including attention-intensive tasks. Future studies could include different types of immediate posttests (i.e., episodic form recall, contextualised form recall test) as different tasks can affect the results. For instance, a contextualised form recall test might lead to greater learning gains as learners might be able to retrieve vocabulary more easily when context cues are available. Further,

this type of test may more accurately reflect real-life learning situations, where learners encounter new vocabulary in context.

Another potential limitation is the congruency task used to determine the congruency status of the target items, which presented raters with the intact collocations (both component words). I suspect that a few raters were unintentionally thinking of compositionality while assessing congruency, which may have influenced some of the ratings. A more fine-grained method of assessing congruency would have been to show raters the individual components of the collocations (e.g., Vilkaitė, 2017).

Finally, previous studies looking at collocation retention show an advantage of verbatim repetition (encountering a collocation in the same context) over varied repetition (encountering a collocation in different contexts) (Durrant & Schmitt, 2010). An open question is how repetition and contextual support might interact in influencing collocation acquisition.

6.7. Conclusion

Taken together, the findings of the present experiment indicate that context informativity can enhance collocational learning at the levels of form recall and meaning recall, in circumstances where congruency cannot be relied upon. The improvement in collocational knowledge in this experiment is likely due to the effect of an optimum level of context informativity, as well as the background and characteristics of the participants - relatively advanced and highly motivated L1-Spanish university EFL learners, who showed a high level of interest and engagement in the tasks and had experience reading academic texts.

This experiment may be a good reflection of actual vocabulary learning from reading for several reasons. First, it reflects the reading that university EFL learners typically engage in as part of their studies. Second, even if the medicine-related texts used as stimulus materials did not match the participants' academic backgrounds, the varying levels of informativity of the contexts represented the typical range of contexts they are likely to encounter while reading for their studies. The stimulus materials included pedagogical contexts, i.e., contexts that included very meaningful and clear context cues (high level of informativity). Such contexts are comparable to ones used in English textbooks and teaching materials. Participants were also exposed to contexts that contained a few helpful cues (mid level of informativity), which resemble general contexts in which learners can use partial information to infer the meanings of unfamiliar lexical items. Lastly, uninformative contexts (low level of informativity),

or non-directive in Beck et al.'s (1983) model, were contexts that did not contain any clear cues (Nation & Webb, 2010).

7. General discussion and conclusions

The aim of this chapter is to contextualise the main findings from the experiments reported in this PhD thesis within the framework of previous research on incidental collocation learning. I first identify the general themes that have emerged across the three experiments conducted, interpret its main findings, and discuss them in relation to the literature. Next, I discuss the pedagogical implications of my research, present potential directions for future research, and make concluding remarks.

7.1. Findings and implications

7.1.1. Incidental collocation learning from reading-(while-listening)

Previous studies have shown that learners can pick up collocations from meaning-focused input activities such as reading (e.g., Dang et al., 2022; Vilkaitė, 2017; Vu & Peters, 2022a; Webb & Chang, 2020), but the amount of learning appears to vary according to moderator variables such as learner characteristics, properties of the reading materials, and collocation-related factors. If "comprehensible" input (Krashen, 1982) is essential for learning, it follows that we should try to understand the range of interacting factors that make input comprehensible for collocation learning, including the impact of different reading modes. In terms of reading, the results of this thesis show two main trends. On the one side, Experiment 1 found new evidence that RWL does not always hold an advantage over RO, which suggests that not all types of input may benefit from additional audio "support" due to the complexity of the materials (e.g., Dang et al., 2022). On the other, the experiments of this thesis led to learning gains from reading at different degrees of collocation knowledge. For instance, reading a single text in Experiment 1 resulted in sizeable gains at the level of form recognition, while Experiments 2 and 3 showed the effectiveness of reading multiple texts for gaining knowledge at the form recall, form recognition, and meaning recall levels, provided that collocations are seeded in short texts (Experiment 2), or the reading materials are contextually supportive of learning (Experiment 3). Taken together, the results corroborate previous findings that collocations can be learnt through reading only.

Collocation learning from reading has its advantages. Reading alone may carry a lower cognitive burden compared to RWL, as learners only need to attend to the written input. This may help them stay

focused on the reading task as they do not have to integrate information from different channels. It also allows learners to pause and go back to reread parts of the text that they may not have fully understood (Dang et al., 2022). Learners can also evaluate whether an unknown lexical item needs to be guessed, which is likely to bring about greater effort through the guessing (Nation, 2013), thereby improving retention. Learners may not find as easy to do this in a RWL condition, where they have to process and interpret information in real time. Another important aspect of RO versus RWL involves learner perception of the task. Participants of the RWL group in Experiment 1 reported feeling less in control of their own reading comprehension, as they had to read at the pace of the auditory input. It is thus likely that RO offers a more self-regulated, autonomous learning experience, which can positively contribute to performance (Teng & Zhang, 2022).

Similar to previous research on incidental collocation learning, gains across experiments were relatively small, but they are still meaningful given the incremental nature of collocation development and vocabulary learning in general (e.g., Webb & Chang, 2020). The findings also revealed trends that are consistent with the literature on the differential learning challenge associated with various aspects of collocation knowledge. Form recall remained the most difficult aspect of collocation knowledge to acquire across experiments, while gains in receptive knowledge aspects (e.g., form recognition) were more easily learnt. This is consistent with previous studies (e.g., Pellicer-Sánchez, 2017; Vilkaitė, 2017; Vu & Peters, 2022a, Webb et al., 2013; Webb & Chang, 2020) and echoes evidence that receptive aspects of vocabulary knowledge are more easily acquired than productive ones (Nation, 2001).

7.1.2. Text-related factors (text type and text length)

Collectively, the results of this thesis indicate that the source of input and text characteristics may affect task difficulty and thus learning. One theme that emerged regarding text type is that the complexity of the materials can affect learning, as revealed in Experiment 1. Exploring how much incidental collocation learning occurs when learners read more challenging texts in terms of content is important, especially given that many participants in vocabulary research studies are university students who are also likely to be relatively advanced EFL learners (e.g., Vilkaite, 2017). Expository academic texts can provide a better indication of more typical incidental collocation learning among these students because they reflect the types of texts they read as part of their studies. Moreover, because university students spend a considerable amount of time reading such texts for their academic subjects (Coxhead,

2018), they may benefit from reading more than adult students enrolled in traditional FL classrooms, where sustained reading (if any) may be limited to class time. The results of this thesis indicate that EFL learners studying at university spend a significant amount of time reading academic texts, which shows that most L2 reading is utilitarian in nature. This is a meaningful finding because it suggests that reading is perceived as part of study habits, rather than a pleasurable activity. Yet, this does not necessarily mean that students are unaware of the importance of reading, but rather that they are more interested in other out-of-class activities such as interacting on social media, playing videogames, etc.). Taken together, the results emphasise the importance of examining various text-related properties (e.g., vocabulary load, informational density, complexity, formality) as they influence the extent to which learners can attend to new vocabulary under different incidental learning conditions (e.g., Dang et al., 2022). More research on different text types, however, is needed to understand how text type differences may affect incidental vocabulary/collocation learning (Gardner, 2004).

Text length is another text-related factor that has received little attention in research on incidental vocabulary/collocation learning. Previous studies have largely used short stories as reading materials, particularly (modified) graded readers, because these texts provide learners with fine-tuned vocabulary for different proficiency levels (e.g., Vu & Peters, 2022a; Webb et al., 2013; Webb & Chang, 2020). Nonetheless, there is evidence that suggests that adult EFL learners do not read extensively in the FL as they prefer other activities such as watching English TV (Peters, 2018). The students' learning habits and preferences should also be considered, because this can reflect learning in a more ecologically valid way. For instance, 80% of the participants in Experiment 2 reported that they preferred reading short texts such as book summaries and online news over long texts. This is a promising finding, particularly considering that reading news online is a rich source for incidental learning of academic and general formulas, among other types of vocabulary (Dang & Long, 2023). Dang and Long's (2023) study included news articles of various lengths (e.g., 350, 700 running words) to reflect the real-life range of news article lengths. This suggests that L2 learners could benefit from reading short texts as a way to compensate for the lack of reading long texts like narrative or nonfiction books, but only if they engage in regular reading, further highlighting the incremental nature of vocabulary learning.

Therefore, it seems that that the type of input (e.g., non-academic vs academic) and the characteristics of the reading materials (e.g., narrative vs expository texts) play a role in the learning

process, which highlights the importance of optimising the degree of complexity of the materials for developing effective pedagogical strategies.

7.1.3. Trade-offs in incidental collocation learning

A striking finding of this thesis concerns the trade-offs that unfold as a result of adopting mixed-effect modelling in the analysis, in response to suggestions that using rigorous statistical analytical approaches is important (see Siyanova-Chanturia & Omidian, 2020). The possibility of a complexity trade-off was first observed in Experiment 1, where learners in the RWL group did not benefit from the additional information in the auditory input, i.e., prosodic cues of collocations such as pauses, stress, intonation (Lin, 2012). As already stated, it is likely that RWL led to split-attention effects because learners' capacity of processing information was overburdened due to the complexity of the text and/or the properties of the auditory input (e.g., speech rate, accent familiarity, etc). This finding stresses the relevance of examining different types of input and its characteristics, as they can lead to differences in performance.

One broad theme that emerged from the findings is that properties of the collocations (e.g., congruency) and the materials (e.g., repetition) seem to have variable effects depending on properties of the learners (prior vocabulary knowledge). The results of Experiments 1 and 2 raised the question of how overall language proficiency (in this study, prior vocabulary knowledge) influences the effect of congruency. It appears that congruency can be relied on, but only within a particular band of proficiency. This suggests that if processing is effortful due to low proficiency, the strength of the congruency effect can increase. From this perspective, the utility of congruency for learning is maximal when proficiency is low. However, as previously noted in this thesis, findings about congruency are limited due to the fact that congruent and incongruent items likely differ on other linguistic dimensions. Nevertheless, this finding contributes to our understanding of collocational competence and may have implications for L2 pedagogy and classroom practise (see Section 7.2).

Further, Experiment 2 also revealed that prior vocabulary knowledge interacts with other influential factors such as repetition, as it was found that repetition had a greater effect as proficiency decreased, at least at the level of form recall. This suggests that learners with smaller vocabularies benefit the most from increased repetitions, possibly because they know fewer words and, as a result, are also less likely to have partial knowledge of the target items. This finding further emphasises the

importance of assessing the multiple interacting variables that affect collocation/vocabulary learning (Chen & Truscott, 2010).

From a methodological standpoint, this study offers further insight into the value of adopting a multivariate approach to studying collocation learning through meaning-focused input (e.g., Vu, 2022). Previous research only looked at learning gains without considering other factors that influence learning, such as prior vocabulary knowledge, congruency, or compositionality (e.g., Sonbul & Schmitt, 2013; Szudarski, 2012; Szudarski & Carter, 2016). Taken together, the findings suggest that a multivariate approach to collocation research is instrumental for better understanding the impact of moderator variables involved in incidental collocation learning.

7.1.4. Prior vocabulary knowledge, congruency, and features of the learning context

This thesis has provided new insights into a variety of variables that affect collocation learning. Overall, the findings demonstrate a strong positive effect of prior vocabulary knowledge, which is consistent with previous research (e.g., Puimège & Peters, 2019, Majuddin et al., 2021; Vilkaitė, 2017, Vu & Peters, 2022a). Broadly translated, this finding indicates that an increase in vocabulary knowledge is associated with an increase in overall language proficiency, particularly for receptive vocabulary knowledge (Milton, 2013). This is explained by the fact that learners who have larger vocabularies have better reading skills (Schmitt et al., 2011), allowing them to pay attention to unknown L2 words while engaged in a reading task. The trend that emerged is clear: prior vocabulary knowledge predicts incidental collocational learning.

Regarding collocational congruency, Experiments 1 and 2 indicated that congruency plays a role in incidental collocation learning, although its effect may vary depending on the aspect of collocational knowledge being tested (e.g., form recall, form recognition), the measurement instruments (test format), the learners' individual characteristics (prior vocabulary knowledge), and the learning experience (e.g., task involvement, salient context cues in the context, etc.). The main conclusion that can be drawn is that congruent items are easier to acquire thanks to the L1 equivalents of their component words. This finding adds to a growing corpus of research showing that congruency affects both incidental (Vu & Peters, 2022a, 2022b) and intentional (Peters, 2016) collocation learning.

Concerning features of the learning context, this thesis contributes to our understanding of how repetition promotes collocation learning. One issue that may make collocations troublesome for learners is that they are frequently composed of high-frequency words which may not be salient enough in the input (e.g., Boers, 2020). A way to draw learners' attention to collocations is through increased repetition. Results of Experiment 2 showed that repetition had a positive effect on collocation learning, which supports findings from previous studies (e.g., Macis et al., 2021). It should be noted that repetition was operationalised as increased exposure to the same collocations within the same text. Since the texts were relatively short (a paragraph long), it is likely that learners noticed the items as they were exposed to high frequency over a short period of time. The rationale behind this operationalisation of increased repetition was grounded on Schmidt's (1990) Noticing Hypothesis, which posits that factors determining noticeability include salience and frequency. From Baddeley's (1997) perspective, such operationalisation is particularly effective because participants are re-exposed to new items almost immediately after the initial encounter, which aids in reinforcing the initial form-meaning link (Baddeley, 1997; Nation, 2001). It must be noted that this experiment used contextually supportive texts, as rated using Webb's (2007a) context quality rating scale. As a result, it is possible that the presence of rich context cues stimulated salience, facilitating increased noticing of the collocations, thus leading learners to engage in more lexical inferencing.

Building on the idea that contextual support may affect incidental collocation learning, Experiment 3 was designed to address its effect in a more controlled fashion by creating three levels of informativity (high, mid, low). To my knowledge, this is the first experiment to examine the effect of contextual support on collocation learning. The results from Experiment 3 showed that contextual support was a predictor of recalling the meanings and forms of collocations, which provides initial evidence that contextual informativity influences incidental collocation learning and expands upon previous findings indicating that context affects the learning of single words (e.g., Teng, 2016; Webb, 2008). Experiment 3 also raised important questions about the role of contextual support in general, as the level of informativity increased performance to a certain degree only. Learning from mid-informative contexts resulted in greater gains than learning from low and high informativity contexts, which implies that there may be an optimal level of contextual support that triggers learners' effort in decoding the meaning of collocations. This finding is related to the Involvement Load Hypothesis (Laufer & Hulstijn, 2001), as it implies that the contextual support of a text may trigger different levels of cognitive involvement or

engagement from learners. Thus, identifying an appropriate level of contextual support is also important for facilitating vocabulary learning.

7.2. Pedagogical implications

The findings of this PhD project have pedagogical implications for course designers, material developers, and teaching methodology in general. Regarding text-related aspects, both teachers and material developers need to consider the type and length of texts they employ (e.g., fiction vs nonfiction; long vs short texts), as well as the learners' background and experience with the language. Specifically, I propose the manipulation of short texts, such as information leaflets, reviews, newspaper articles, or online news, to include multiple repetitions of collocations, so that learners can build up on the instances they encounter in meaningful input. This is a concern in L2 textbooks, where collocations (and MWUs in general) are underrepresented, or featured in a way that is not characteristic of L1 formulaic language (Koya, 2004; Northbrook & Conklin, 2019). Moreover, they are often presented in isolation using definitions, translation exercises, or match up activities (Brown, 2011). However, the quality of the contextual support may provide a better chance of gaining knowledge of collocations. For example, having learners explore varied contexts in which a collocation or MWU can occur induces more in-depth mental processing, which may in turn lead to assimilation (Schmitt, 2010). Textbooks writers should also consider the distribution of repetition. As previously mentioned in this thesis, repeated encounters with the same collocations in natural contexts are generally rare (Boers & Lindstromberg, 2009). The amount of collocation knowledge that may be gained in the first encounter is likely to be forgotten by the time learners encounter the same collocation a second time. Seeding a text with plenty of repetitions of target collocations is a shortcut to enrich the input and promote noticing. As shown in Experiment 2, increased repetition can serve to increase the salience of collocations, thereby facilitating learning.

Based on the differential effects of the reading modes investigated in this thesis, it can be claimed that RWL does not necessarily result in better learning when compared to RO. This has clear implications for vocabulary pedagogy as it involves making sure that the learning task (e.g., RO, listening, RWL) is carefully chosen to maximise collocation/vocabulary learning. Ideally, the selection of the learning activities should involve some consideration of the properties of the materials (e.g., text type and its length in a reading task, amount of lexical coverage, etc.), the background of the learners

(e.g., L2 proficiency, education level), and the way in which the activity is used (e.g., individualised vs group learning).

As for the effect of prior vocabulary knowledge, it seems that learners with smaller vocabularies might need additional support when learning collocations incidentally. One way to facilitate such learning is using attention-drawing techniques such as adding textual input enhancement (e.g., bolding) which help direct learners' attention to new items in the input. Research indicates that textually enhancing collocations is positively related to incidental learning of those items (e.g., Sonbul & Schmitt, 2013; Toomer & Elgort, 2019). In terms of congruency, the results from Experiments 1 and 2 suggest that incongruent items may require more attention from learners and teachers than congruent collocations. This could be done by text manipulation. For example, as Experiment 3 revealed, presenting incongruent collocations in pedagogical contexts can promote collocation learning, even when collocations display a low degree of congruency.

Teachers should raise awareness of the existence of collocations and their functions in academic environments, particularly in EFL contexts where learners' knowledge of collocations is frequently lower than in English as a Second Language (ESL) settings (Gaballa & Al-Khayri, 2014). Incorporating the most frequently used collocations according to the field of study may also enable learners to achieve an adequate collocational competence in academic environments (Ellis et al., 2008). Teachers could also benefit from looking at the various corpus-based collocation lists that have been compiled on specific collocation types that appear to pose difficulty to the learners (e.g., Ackerman & Chen, 2013; Durrant, 2009; Shin & Nation, 2008). It should be noted, however, that some of these lists are limited as they focus on specific types of vocabulary (e.g., The Academic Collocation List by Ackerman and Chen, 2013) or target specific language proficiency levels (e.g., Shin and Nation's (2008) list of high-frequency collocations in spoken English targets low proficiency learners). Nevertheless, these listings are pedagogically valuable for L2 collocational learning and teaching, in particular for those involved in English for Academic Purposes (EAP)and/or ESL programmes.

7.3. Directions for future research

The limitations of the experiments presented in this thesis have been already addressed in the Limitations sections of Chapters 4, 5, and 6. Based on those, this section will focus on recommendations for future research. First, future research should consider flooding longer texts with repetitions in a large

number of experimental reading sessions, over a longer period of time, to determine whether that could lead to the development of stronger collocation knowledge. It would also be beneficial to examine the extent to which partial knowledge of collocations contribute to the learning gains. In my experiments, many of the component words of the collocations were likely to be known (e.g., fever, cage), while others were more likely to be new to the learners (e.g., lice, sore). Admittedly, the degree of familiarity with a collocation's constituents may affect its learnability, particularly in inferring its meaning (Webb & Chang, 2020), so partial knowledge should be considered. Future research could also examine the acquisition of other types of collocations. Research on verb-noun collocations indicate that they pose a different level of difficulty than the adjective-noun and noun-noun collocations examined in this thesis (e.g., Vilkaitė, 2017). More research on different types of collocations, e.g., grammatical collocations, is thus needed to better understand the different challenges they present to learners.

Another line of enquiry for future research concerns the generalisability of the findings to other learning conditions. According to existing research, learners make larger and faster gains under deliberate learning conditions than under incidental learning conditions (Webb & Nation, 2017). It would thus be interesting to see whether a deliberate focus can be more beneficial for accelerating collocational learning, in particular from the availability of contextual support in short texts.

Changes in the experimental conditions may produce more informative results in future research concerning the role of context. Prior research suggests that varied contexts promote the formation of richer semantic associations (e.g., Ferreira & Ellis, 2016), but the evidence is limited to individual words. It is not yet known whether encountering collocations in varied contexts enhances learning more than collocations met in the same context. There is some evidence, however, of the better retention of the form of collocations through single contexts (fluency-oriented repetition) relative to multiple contexts (Durrant & Schmitt, 2010). It would be interesting to see if the same context advantage applies to collocation acquisition, particularly for guessing incongruent collocations' meanings, and how much this is affected by the quality/quantity of contextual cues.

Future research could also consider including different measures of L2 vocabulary proficiency while taking into consideration the participants' L1. The VLT composite score was used as the measurement instrument to assess language proficiency in this thesis. While the VLT provides a general estimate of vocabulary proficiency, it should not be interpreted as the actual vocabulary size of

the learners. Furthermore, the L1 of the learners in this thesis may have influenced their knowledge of low frequency words, which could have aided reading comprehension.

7.4. Concluding remarks

The purposes of this PhD project have been to investigate the incidental learning of collocations and factors that predict their acquisition. Its results contribute to L2 collocation learning research with the finding that RWL does not necessarily confer an advantage over RO when learners are exposed to academic texts. Further, this thesis has provided initial evidence that the degree of contextual support affects collocation learning, which opens up a large potential source for learning collocations and other types of MWUs from meaning-focused tasks, as long as researchers and material developers are willing to manipulate and control for the role contextual support.

The findings showed how the following factors can influence incidental collocation acquisition:

1. Prior vocabulary knowledge (Experiments 1, 2, and 3)

2. Congruency (Experiments 1 and 2)

3. Association strength (Experiment 2)

4. Repetition (Experiment 2)

5. Contextual support (Experiment 3)

A final, somewhat broader conclusion, relating to the findings of this thesis concerns the nature of L2 learning itself. This study offers insights into how aspects of L2 learning can occur organically, via EFL classroom-external activities that learners are already engaging with in their day to day lives. Through the lens of collocation learning, this thesis highlights the holistic nature of vocabulary learning and the L2 in general. From this standpoint, learning is concerned with the value of the activities in which learners are routinely involved, as part of their own idiosyncratic learning journeys, rather than how it may occur in the FL classroom. Such an important source of learning must not be discounted by adhering strictly to established notions about how to maximise learning, as these may not apply to the learning realities of the learners. The broad implication of the present research is that L2 learning is not only governed by the effectiveness of specific teaching practises or learning tasks, but it is also influenced by multiple factors that interact as a whole. Understanding how incidental (vocabulary)

learning occurs requires a holistic approach to unfolding the underlying mechanisms of such learning process. Learning is ultimately a lifelong, socio-contextual, and educational experience.

References

Ackerman, B. P. (1982). On comprehending idioms: Do children get the picture? *Journal of Experimental Child Psychology*, *33*(3), 439-454.

Ackermann, K., & Chen, Y-H. (2013). Developing the Academic Collocation List (ACL). A corpus driven and expert judged approach. *Journal for Academic Purposes*, 12, 235-247.

Agustín Llach, M., & Moreno Espinosa, S. (2014). Knowledge of form and meaning. In J. Milton & T. Fitzpatrick (Eds.), *Dimensions of vocabulary knowledge* (pp. 60-72). Basingstoke: Palgrave Macmillan.

Aitchison, J. (2003). *Words in the mind.* (3a). Oxford: Blackwell.

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology : Theory, Analysis, and Applications* (pp. 101-122). Oxford: Clarendon Press.

Altenberg, B. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, *22*(2), 173-195.

Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, *22*(2), 173-195.

Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L.S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory.* Hillsdale, NJ: Erlbaum.

Anderson, R.C., & Freebody, P. (1981). Vocabulary Knowledge In J.T. Guthrie (Ed), *Comprehension and teaching: Research reviews* (pp.77-117). Newark, DE: International Reading Association.

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*, 67-82. doi:10.1016/ j.jml.2009.09.05

Baayen, R.H., Davidson, D.J., & Bates D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A.D. (1997) *Human Memory: Theory and Practice*. Psychology Press, East Sussex.

Bahns, J. (1993). Lexical collocations: a contrastive view. *ELT Journal*, *47*(1), 56-63.

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, *21*(1), 101-114.

REFERENCES

Baldwin, Timothy (2006) Distributional Similarity and Preposition Semantics, In Patrick Saint-Dizier (ed.) *Computational Linguistics Dimensions of Syntax and Semantics of Prepositions*, Springer, ISBN 1402038496, pp. 197-210.

Bannard, C. & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*(3), 241-248.

Barclay, S. & Pellicer-Sánchez, A. (2021) Exploring the learning burden and decay of foreign language vocabulary knowledge: The effect of part of speech and word length. *ITL - International Journal of Applied Linguistics 10.*1075/itl.20011.bar. (In press). Green open access.

Barr, D.J., Levy, R., Scheepers, C., & Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-268.

Bartsch, S. (2004). *Structural and Functional Properties of Collocations in English: A Corpus Study of Lexical and Pragmatic Constraints on Lexical Co-occurrence*. Gunter Narr Verlag, Tübingen.

Beck, I., McKeown, M., & McCaslin, E. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal*, *83*, 177-181. doi:10.1086/ 461307

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). London: The Guilford Press.

Benson, M., Benson, E., & Ilson, R. (1997). *The BBI dictionary of English word combinations.* Amsterdam: John Benjamins.

Bentley, T. (1998). Learning beyond the classroom. *In Tomorrow's Politics: The Third Way and Beyond*. Edited by I. Hargreaves & I. Christie, 80-95. Demos: London.

Biber, D. (2009). A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, *14*(3), 275-311.

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: Studies in honor of Stig Johansson* (pp. 181-189). Amsterdam: Rodopi.

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

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371-405.

REFERENCES

Bird, S. A., & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. Applied Psycholinguistics, 23(4), 509-533.

Bishop, H. (2004). The effect of typographical salience on the look up and comprehension of formulaic sequences. In Schmitt (Ed.), Formulaic sequences: Acquisition, processing and use (pp 227-244). John Benjamins.

Boers, F. (2020). Factors affecting the learning of multiword items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp.143-157). Routledge.

Boers, F., Lindstromberg, S., & Eyckmans, J. (2014). Some explanations for the slow acquisition of L2 collocations. *Vigo International Journal of Applied Linguistics*, *11*, 41-62.

Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercises on verb-noun collocations. *Language Teaching Research*, *18*(1), 54-74.

Boers, F., & Webb, S. (2015). Gauging the semantic transparency of idioms: Do natives and learners see eye to eye? In R. R. Heredia & A. B. Cieślicka (Eds.), *Bilingual figurative language processing* (pp.368-392).Cambridge University Press.

Boers, F., Demecheleer, M., He, L., Deconinck, J., Stengers, H., & Eyckmans, J. (2017). Typographic enhancement of multiword units in second language text. *International Journal of Applied Linguistics*, *27*(2), 448-469. doi:10.1111/ijal.12141

Boers, F., & Webb, S. (2018). Teaching and learning collocation in adult second and foreign language learning. *Language Teaching*, *51*, 77-89.

Boone, G., De Wilde, V., & Eyckmans, J. (2023). A longitudinal study into learners' productive collocation knowledge in L2 German and factors affecting the learning. *Studies in Second Language Acquisition*, *45*(2), 503-525.

Bradford, P. (2010). *The acquisition of colloquial speech and slang in second language learners of English in El Paso, Texas. Open Access Theses & Dissertations*. 2443.

Brown, C. (1993). Factors affecting the acquisition of vocabulary: Frequency and saliency of words. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 263-286). Norwood, NJ: Ablex.

Brown, D. (2011). What aspects of vocabulary knowledge do textbooks give attention to? *Language Teaching Research*, *15*(1), 83-97.

REFERENCES

Brown, J. (2014). Knowledge of collocations. In J. Milton & T. Fitzpatrick (Eds.), *Dimensions of Vocabulary Knowledge* (pp. 123-139). Basingstoke: Palgrave Macmillan.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, *20*(2), 136-163.

Buerki, A. 2020. (How) is formulaic language universal? Insights from Korean, German and English. Piirainen, E., Filatkina, N., Stumpf, S., & Pfeiffer, C., eds. *Formulaic Language and New Data: Theoretical and Methodological Implications, Formulaic Language*, vol. 2. Berlin: De Gruyter, pp. 103-134.

Bybee, J. (1995). Phonology of the lexicon: Evidence from lexical diffusion. Paper presented at the Rice University Symposium on Usage-based Models of Language. March1995, Houston, TX.

Bybee, J. (2002). Phonological evidence for exemplar storage of multiword sequences. *Studies in Second Language Acquisition*, *24*, 215-221.

Carey, S. (1978). *The child as word learner*. In Halle, M., J. Bresnan, and G. A. Miller, editors, Linguistic Theory and Psychological Reality. The MIT Press.

Carrol, G., & Conklin, K. (2020). Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. *Language and Speech*, *63*(1), 95-122.

Carrol, G., & Conklin, K. (2014). Getting your wires crossed: Evidence for fast processing of L1 idioms in an L2. *Bilingualism: Language and Cognition.* *17*(4), 784-797.

Çekiç, A., & Demirezen, M. (2020). Comparison of the impacts of different multimodalities on incidental l2 vocabulary learning. *Moderna Språk*, *114*(2), 109-138.

Chacón-Beltrán, R., Abello-Contesse, C., & Torreblanca-López, M. del Mar (Eds.). (2010). *Insights into non-native vocabulary teaching and learning*. Bristol, UK: Multilingual Matters.

Chan, T., & Liou, H. (2005). Effects of web-based concordancing instruction of EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning*, *18*(3), 231-251. doi:10.1080/09588220500185769

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293-332.

Chapelle, C. (1998). Construct definition and validity inquiry in SLA research. In L. Bachman & A. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp. 32-70). Cambridge: Cambridge University Press.

REFERENCES

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics*, *31*(5), 693-713.

Chen, X., Ramirez, G., Luo, Y. C., Geva, E., & Ku, Y.-M. (2012). Comparing vocabulary development in Spanish-and Chinese-speaking ELLs: The effects of metalinguistic and sociocultural factors. *Reading and Writing*, *25*(8), 1991- 2020.

Choi, S. (2017). Processing and learning of enhanced English collocations: An eye movement study. Language Teaching Research, 21(3), 403-426.

Chomsky N. (1959). Review of Skinner's Verbal Behavior. *Language 35,* 26-58.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology*, *11*(3), 38-63.

Cobb, T. (2015). *Web Vocab Profiler Compleat*. Retrieved from https://www.lextutor.ca/vp/comp/

Cobb, T. *Compleat lexical tutor v.8* [Computer Program]. Retrieved from http://www. lextutor.ca/ concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning 50*: 1-56.

Conklin, K., Alotaibi, S., Pellicer-Sánchez, A., & Vilkaitė-Lozdienė, L. (2020). What eye-tracking tells us about reading-only and reading-while-listening in a first and second language. *Second Language Research*, *36*, 257-276.

Conklin, K., & Carrol, G. (2018). First language influence on the processing of formulaic language in a second language. In: Siyanova-Chanturia, A. and Pellicer-Sánchez, A., eds., *Understanding formulaic language: A second language acquisition perspective* (pp 62-77). Routledge.

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, *32*, 45-61.

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. Journal of Memory and Language, *26*(3), 341-361. 208

Cooper, T. C. (1998). Teaching idioms. *Foreign Language Annals*, *31*(2), 255- 266.

Coulmas, F. (1981). Introduction: conversational routine. In F. Coulmas (Ed.), *Conversational routine: explorations in standardized communication situations and prepatterned speech* (pp. 1-17). The Hague: Mouton Publishers.

REFERENCES

Cowie, A. P. (1981a). Lexicography and its Pedagogic Applications: An Introduction. *Applied Linguistics*, *2*(3), 203-206.

Cowie, A. P. (1981b). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics*, *2*(3), 223-235

Cowie, A. P. (1994). Phraseology. In R.E. Asher & J.M.Y. Simpson (Eds.), *The encyclopedia of language and linguistics* (Vol. 6, pp. 3168-3171).

Cowie, A.P. (1988a). Stable and creative aspects of language use, in Carter, R. and McCarthy, M. (eds), *Vocabulary and Language Teaching*. London: Longman, pp. 126-39.

Cowie, A. P. (Ed.) (1998). *Phraseology: theory, analysis, and applications.* Oxford: Oxford University Press.

Coxhead, A. (2000). A new academic word list, *TESOL Quarterly*, *34*(2), 213-238.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behaviour*, *11*(6), 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.

Croft, W. (2000). *Explaining language change: An evolutionary approach*. Harlow, Essex: Longman.

Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *The Journal of Educational Research*, *36*(3), 206-217.

Crossley, S., Salsbury, T., & McNamara, D. (2010). The development of polysemy and frequency use in English second language speakers. *Language Learning*, *60*(3), 573-605.

Cruse D. A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

Dang, T. N. Y., Lu, C., & Webb, S. (2021). Incidental learning of single words and collocations through viewing an academic lecture. *Studies in Second Language Acquisition*. Advance online publications. https://doi.org/10.1017/S0272263121000474.

Dang, T. N. Y., Lu, C., & Webb, S. (2022). Incidental learning of collocations in an academic lecture through different input modes. Language Learning. Advance online publication. https://doi.org/10.1111/lang.12499.

Dang, T.N.Y., Lu, C., & Webb, S. (2022). Open access academic lectures as sources for incidental vocabulary learning: Examining the role of input mode, frequency, type of vocabulary, and elaboration. *Applied Linguistics*. https://doi.org/10.1093/applin/amac044.

REFERENCES

Dang, T.N.Y., & Long, X. (2023). Online News as a Resource for Incidental Learning of Core Academic Words, Academic Formulas, and General Formulas. *TESOL Quarterly*, Early view: February 2023 https://doi.org/10.1002/tesq.3208.

Davies, M. (2008). The Corpus of Contemporary American English. http://www.english-corpora.org/coca.

Davies, M., & Gardner, D. (2011). *A frequency dictionary of contemporary American English*. London: Routledge.

de Bot, K., Paribakht, T.S., & Wesche, M.B. (1997). Toward a lexical processing model for the study of second language vocabulary acquisition: Evidence from ESL reading. *Studies in Second Language Acquisition*, *19*(3), 309-329.

de Bot, K., Lowie, W., and Verspoor, M. (2005). *Second Language Acquisition: An advanced resource book.*

De Groot, & Keijzer, R. (2000) What is hard to learn is easy to forget: The roles of word

de Jon, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional Science 38,* 105-134.

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf, & G. Appel (Eds.), *Vygostkian approaches to second language research* (pp. 33-56). New Jersey: Ablex.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.

Dunn, L. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test (PPVT-IV)* (4th ed.). Circle Pines, MN: American Guidance Service.

Durrant, P. (2008) *High frequency collocations and second language learning.* [Doctoral thesis, University of Nottingham].Nottingham e-Theses. https://eprints.nottingham.ac.uk/10622/.

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, *28*(3), 157-169.

Durrant, P. (2014). Corpus frequency and second language learners' knowledge of collocations: a meta-analysis. *International Journal of Corpus Linguistics*, *19*(4), 443-477.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations?. *IRAL - International Review Of Applied Linguistics In Language Teaching*, *47*(2).

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, *26*(2), 163-188.

Durrant, P. & J. Mathews-Aydınlı (2011). A function-first approach to identifying formulaic language in academic writing. *English for Specific Purposes*, *30*(1): 58-72.

Eckerth, J., & Tavakoli, P. (2012). The effects of word exposure frequency and elaboration of word processing on incidental L2 vocabulary acquisition through reading. *Language Teaching Research*, *16*(2), 227-252.

Elgort, I., Perfetti, C.A., Rickles, B., & Stafura, J.Z. (2015). Contextual learning of L2 word meanings: Second language proficiency modulates behavioural and event-related brain potential (ERP) indicators of learning. *Language, Cognition and Neuroscience*, *30*(5), 506-528.

Ellis, N. C. (1994). Introduction. Implicit and explicit language learning: An Overview. In N. Ellis (Ed.), *Implicit and Explicit Learning of Languages* (pp. 1-32). London; San Diego: Academic Press.

Ellis, N. C. (1995). Consciousness in second language acquisition: A review of field studies and laboratory experiments. *Language Awareness*, *4*, 123-146.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, *19*, 91-126.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition 24*(2), 143-18.

Ellis, R. (2005). *Instructed second language acquisition: A literature review*. Wellington: Learning Media Ltd.

Ellis, N. C. (2008). Implicit and explicit knowledge about language. In J. Cenoz & Hornberger, *Encyclopaedia of Language and Education* (2nd ed., Vol. 6: Knowledge about Language, pp. 1-13). New York: Springer Science+Business Media LLC.

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2007). *The processing of formulas in native and second-language speakers: psycholinguistic and corpus determinants*. Paper presented at the The 25th UWM Linguistics Symposium.

Ellis, R. & Shintani, N.(2013). *Exploring language pedagogy through second language acquisition research*. Routledge. London.

Ellis, N. & Wulff, S. (2015). Usage-based approaches to SLA. In B. VanPatten, & J. Williams (Eds.), Theories in second language acquisition (pp. 75-93). London: LEA.

Elman, J. L. (2005). Connectionist models of cognitive development: Where next? *Trends in Cognitive Sciences*, *9*, 111-117.

REFERENCES

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text - Interdisciplinary Journal for the Study of Discourse*, *20*(1).

Evert, S. (2004). *Computational approaches to collocations*. Retrieved 18 December, 2021, from www.collocations.de.

Evert, S. (2009). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics. An international handbook (Volume 2)* (pp. 1212-1248). De Gruyter Mouton.

Farghal, M., & Obeidat, H. (1995). Collocations: a neglected variable in EFL. *International review of applied linguistics in language teaching*, *33*(4), 315- 331

Feng, Y., & Webb, S. (2020). Learning vocabulary through reading, listening, and viewing: Which mode of input is most effective?. *Studies in Second Language Acquisition*, *42*(3), 499-52

Ferreira, R. A., & Ellis, A. W. (2016). Effects of contextual diversity on semantic decision and reading aloud: Evidence from a word learning study in English as a second language. *Studies in Psychology, 37*, 162-182.

Fioravanti, I., Senaldi, M. S. G., Lenci, A., & Siyanova-Chanturia, A. (2021). Lexical fixedness and compositionality in L1 speakers' and L2 learners' intuitions about word combinations: Evidence from Italian. *Second Language Research*, *37*(2), 291-322.

Firth, J. R. (1957). Modes of meaning. In *Papers in linguistics* 1934-1951 (pp. 190- 215). Oxford: Oxford University Press.

Foster, P. (2001). Rules and routines: a consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: second language learning, teaching and testing* (pp. 75-94). London: Longman.

Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, *40*(2), 273-293.

Gaballa, H. E. M., & Al-Khayri, M. A. (2014). Testing collocational knowledge of Taif University English seniors. *IOSR Journal of Humanities and Social Science*, *19*(11), 63-90.

Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in corpus-based language learning research: Identifying, comparing, and interpreting the evidence. *Language Learning*, *67*(S1), 155-179.

Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics*, *25*(1), 1-37. doi:10.1093/applin/25.1.1

Garnier, M., & Schmitt, N. (2016). Picking up polysemous phrasal verbs: How many do learners know and what facilitates this knowledge?. *System, 59*, 29-44.

Gass, S. (2017). *Input, Interaction, and the Second Language Learner* (2nd ed.). New York: Routledge.

Gass, S. (1988). Integrating research areas: a framework for second language studies. *Applied Linguistics 9,* 198-217.

Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten, & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175-200). London: LEA.

Gass S., Behney J., & Plonsky L. (2013) *Second Language Acquisition: An Introductory Course* (4th ed.). Routledge, New York.

Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. *Memory & cognition*, *16*(2), 110-119.

Gibson, S. (2008). Reading aloud: A useful learning tool? *ELT Journal*, *62*(1), 29-36.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*, 313-331.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago: Chicago University Press

González-Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? *ITL - International Journal of Applied Linguistics*, *166*(1), 94-126.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In: Cowie, A.P. (Ed.), *Phraseology: Theory, analysis, and applications*. Oxford: Oxford University Press.

Granger, S. (2019). Formulaic sequences in learner corpora: Collocations and lexical bundles. In Siyanova-Chanturia A., Pellicer-Sánchez A. (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 228- 247). New York: Routledge.

Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 27-49). John Benjamins.

REFERENCES

Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics*, *25*(1), 38-61.

Gregersen, T., & Horwitz, E. K. (2002). Language Learning and Perfectionism: Anxious and Non-Anxious Learners' Reactions to Their Own Oral Performance. *Modern Language Journal, 86*, 562-570.

Gregg, K. (2001). Learnability and second language acquisition theory. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (Cambridge Applied Linguistics, pp. 152-180). Cambridge: Cambridge University Press.

Greidanus, T., & Nienhuis, L. (2001). Testing the quality of word knowledge in a second language by means of word associations: Types of distractors and types of associations. *The Modern Language Journal*, *85*, 567577.

Gries, S.T. & Durrant, D. (2020). Analyzing co-occurrence data. In Magali Paquot & Stefan Th. Gries (eds.), A *practical handbook of corpus linguistics*, 141-159. Berlin & New York: Springer.

Gyllstad, H. (2013). Looking at L2 Vocabulary Knowledge Dimensions from an Assessment Perspective - Challenges and Potential Solutions. In C. Bardel, B. Laufer, & C. Lindqvist (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (Vol. Eurosla Monographs Series, 2, pp. 11-28).

Gyllstad, H., Vilkaite, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats : Issues with guessing and sampling rates. *ITL: Institut Voor Toegepaste Linguistik*, *166*(2), 278-306.

Gyllstad, H. & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning*, *66*(2), 296-323.

Haastrup, K. (1989). The Learner as Word Processor. *International Association of Applied Linguistics Review*, *6*, 34-46. .

Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, *10*, 221-240.

Halliday, M. A. K. (1966). Lexis as a linguistic level. In C. E. Bazell, J. C. Catford, M. A. K. Halliday & R. H. Robins (Eds.), In *Memory of J.R.Firth* (pp. 148-162). London: Longmans, Green and Co. Ltd.

Hawkins, J. (2004). *Efficiency and Complexity in Grammars*, Oxford University Press, Oxford, UK.

REFERENCES

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, *21*(02), 303-317.

Henriksen, B. (2013). Research on L2 learners' collocational competence and development. In C. Bardel, C. Lindquist & B. Laufer (Eds.), *L2 vocabulary acquisition, knowledge, and use: New perspectives on assessment and corpus analysis. EuroSLA Monograph Series 2* (pp. 29-56). European Second Language Association.

Hill, J. (2000). Revising priorities: from grammatical failure to collocational success. In M. Lewis, (Ed.), *Teaching Collocation* (pp. 47-69). Language Teaching Publications.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. New York: Routledge.

Hoover, D. L. (2002). Frequent words sequences and statistical stylistics. *Literary and Linguistic Computing 17*(2):157-180

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a Clockwork Orange: acquiring second language vocabulary through reading. *Reading in a Foreign Language*, *11*, 2:207-223.

Howarth, P. (1998). The phraseology of learners' academic writing. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 161-186). Oxford University Press.

Hsu, J.Y. & Chiu, Ch.Y. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *Asian EFL Journal, 10*(1), 181-204.

Hu, H. M. (2013). The effects of word frequency and contextual types on vocabulary acquisition from extensive reading: A case study. *Journal of Language Teaching and Research*, *4*(3), 487-495.

Hu, H.C., & Nassaji, H. (2012). Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings inferred from context. *Canadian Modern Language Review*, *68*(1), 54-77.

Hulstijn, J.H. (1992) Retention of inferred and given word meanings: experiments in incidental vocabulary learning. In P. Arnaud and H. Bejoint (eds.) *Vocabulary and Applied Linguistics Macmillan*, London.

Hulstijn, J. H. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge University Press.

Hulstijn, J.H. (2003). Incidental and Intentional Learning. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 349-381). Malden, MA: Blackwell.

163

REFERENCES

Hulstijn, J., Hollander, M., & Greidanus, T. (1996) Incidental vocabulary learning by advanced foreign language students: the influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *Modern Language Journal 80*(3), 327-339.

Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Irujo, S. (1986). A piece of cake: learning and teaching idioms. *ELT Journal*, *40*(3), 236-242.

Isbilen, E. S., McCauley, S. M., & Christiansen, M. H. (2022). Individual differences in artificial and natural language statistical learning. *Cognition*, 225, 105-123.

Janssen, N., & Barber, H. A. (2012). Phrase frequency effects in language production. *PLoS ONE, 7*(3), Article e33202.

Jiang, N. (2002). Form-meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 24*(04), 617 - 637.

Jin, Z., & Webb, S. (2020). Does writing words in notes contribute to L2 vocabulary learning? *Language Teaching Research*. https://doi.org/10.1177/13621688211062184

Joe, A. (2010). The quality and frequency of encounters with vocabulary in an English for Academic Purposes programme. *Reading in a foreign language*, *22*, 117-138.

Kasahara, K. (2011). The effect of known-and-unknown word combinations on intentional vocabulary learning. *System*, 39, 491-499.

Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, *12*(3), 365-386.

Kecskes, I. (2007). Formulaic language in English Lingua Franca. In I. Kecskes & L. Horn (Eds.), *Explorations in Pragmatics: Linguistic, Cognitive and Intercultural Aspects* (pp. 191-219). Berlin/New York: Mouton de Gruyter. Placement Test Upstream Enterprise. EGIS & Express Publishing.

Keshavarz, M.H., & Salimi, H. (2007). Collocational Competence and Cloze Test Performance: a Study of Iranian EFL Learners. *International Journal of Applied Linguistics, 17*(1), 81-92.

Koya T. (2004). Collocation research based on corpora collected from secondary school textbooks in Japan and in the UK. *Dialogue*, 3, 7-18.

Krashen, S. D. (1982). Principles and practice in second language acquisition. Oxford: Pergamon Press.

REFERENCES

Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, *73*(4), 440-464.

Krashen S. D. (2004). *The power of reading*, 2nd edition. Portsmouth, NH: Heinemann.

Kremmel, B., Brunfaut, T., & Alderson, C. (2017). Exploring the role of phraseological knowledge in foreign language reading. *Applied Linguistics*, *38*(6), 848-870.

Kuppens, A. H. (2010). Incidental foreign language acquisition from media exposure. Learning, *Media and Technology*, *35*(1), 65-85. https://doi.org/10.1080/17439880903561876.

Labov, W. (1973). The boundaries of words and their meanings. In C. J. Bailey & R. Shuy (Eds.), *New ways of analyzing variation in English*. Washington, D. C: Georgetown University Press.

Lakoff, G. (1987). Cognitive models and prototype theory. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 63-100). Cambridge University Press.

Langacker, R. W. (1987). Foundation of Cognitive Grammar (Vol. 1). Theoretical Prerequisites. Stanford: Stanford University Press.

Lantolf, J. P. (Ed.). (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford: Oxford University Press.

Lantolf, J., Thorne, S. L., & Poehner, M. (2015). Sociocultural Theory and Second Language Development. In B. van Patten & J. Williams (Eds.), *Theories in Second Language Acquisition* (pp. 207-226). New York: Routledge.

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science*, *21*(2), 284-291.

Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.

Laufer, B. (1997) What's in a word that makes it hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition. In N. Schmitt and M. McCarthy (eds.) *Vocabulary: Description. Acquisition and Pedagogy*. Cambridge University Press, Cambridge.

Laufer, B. (1998) The development of passive and active vocabulary: same or different? *Applied Linguistics 19* (2), 255-271 .

REFERENCES

Laufer, B. (2000). Avoidance of idioms in a second language: The effect of L1- L2 degree of similarity. *Studia Linguistica*, *54*(2), 186-196.

Laufer, B. (2005). Focus on form in second language vocabulary learning. *EuroSLA Yearbook*, *5*(1), 223-250.

Laufer, B., & Nation, P. (1995). Lexical richness in L2 written production: Can it be measured? *Applied Linguistics, 16*(3), 307-322.

Laufer, B. & Shmueli, K. (1997) Memorizing new words: Does teaching have anything to do with it? *RELC Journal*, *28*, 189-108.

Laufer, B., & Paribakht, T.S. (1998) The relationship between passive and active vocabularies: effects of language learning context. *Language Learning, 48*(3), 365-393.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1-26.

Laufer, B. & Girsai, N. (2008). Form-focused Instruction in Second Language Vocabulary Learning: A Case for Contrastive Analysis and Translation, *Applied Linguistics*, *29*(4), 694-716.

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*(1), 15-30.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: corpus analysis of learners' English. *Language Learning*, *61*(2), 647-672.

Leiner, D. J. (2019). SoSci Survey (Version 3.1.0 and 3.1.6) [Computer software]. Available at https://www.soscisurvey.de

Leiner, D. J. (2022). SoSci Survey (Version 3.4.12) [Computer software].

Lenneberg, E.H. (1967). *Biological foundations of language*. New York: Wiley.

Lewis, M. (Ed.) (2000). *Teaching collocation: Further developments in the Lexical Approach*. Thomson Heinle.

Li, J. & N. Schmitt (2010). The development of collocation use in academic texts by advanced L2 learners: A multiple case study approach. In D. Wood (ed.), *Perspectives on formulaic language: Acquisition and communication* (pp.22-26). Continuum.

Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics*, *36*(5), 611-634.

REFERENCES

Li, W. & Lu, Q. (2011). A hybrid extraction model for Chinese noun/verb synonym bi-gram. Proceedings of the 25th Pacific Asia Conference on Language, Information and Computation (PACLIC 25), 16-18 Dec, Nanyang Technological University, Singapore.

Lightbown P. & Spada N.(2013). *How languages are learned*. Oxford University Press.

Lin, P. M. S. (2010). The phonology of formulaic sequences: A review. *Perspectives on formulaic language: Acquisition and communication*, 174-193.

Lin, P. M. (2012). Sound evidence: The missing piece of the jigsaw in formulaic language research. *Applied Linguistics*, *33*, 342-347.

Lin, P. M. (2014). Investigating the validity of internet television as a resource for acquiring L2 formulaic sequences. *System*, *42*, 164-176.

Lindgren, E. & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, *10*(1),105-129.

Liontas, J. I. (2003). Killing Two Birds with One Stone: Understanding Spanish VP Idioms in and out of Context. *Hispania*, *86*(2), 289-301. https://doi.org/10.2307/20062862

Liu N., & Nation, I.S.P., (1985). Factors Affecting Guessing Vocabulary in Context. *RELC Journal*, *16*(1), 33-42.

Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology* / Revue canadienne de psychologie, *44*(1), 87-112. https://doi.org/10.1037/h0084237

Long, M. (1991). Focus on Form: A Design Feature in Language Teaching Methodology. In K. De Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign Language Research in Cross-Cultural Perspectives* (pp. 39-52). Amsterdam: John Benjamins.

Lotto, L. & De Groot, M. B. A. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31-69.

Lozano, C. (2016). Pragmatic principles in anaphora resolution at the syntaxdiscourse interface: Advanced English learners of Spanish in the CEDEL2 corpus. In M. A. Ramos (Ed.), *Spanish learner corpus research: Current trends and future perspectives* (pp. 235-265). John Benjamins.

Lubliner, S., & Hiebert, E. H. (2011). An Analysis of English-Spanish Cognates as a Source of General Academic Language. *Bilingual Research Journal*, *34*(1), 76-93.

Macis, M. (2018). Incidental learning of duplex collocations from reading: Three case studies. *Reading in a foreign language 30*(1), 48-75.

Macis, M. & Schmitt, N. (2016). Meaning matters: The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal*, doi: 10.1093/elt/ccw044.

Macis, M., & Schmitt, N. (2017). Not just "Small Potatoes": Knowledge of the idiomatic meanings of collocations. *Language Teaching Research*, *21*(3), 321-340.

Macis, M., Sonbul, S., & Alharbi, R. (2021). The effect of spacing on incidental and deliberate learning of L2 collocations. *System*, 103.

MacWhinney, B. (1999). *The emergence of language*. Mahwah, NJ: Lawrence Erlbaum Associates.

Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2021). Incidental acquisition of multiword expressions through audiovisual materials: The role of repetition and typographic enhancement. *Studies in Second Language Acquisition*, *43*(5), 985-1008.

Malone, J. (2018). Incidental vocabulary learning in SLA: Effects of frequency, aural enhancement, and working memory. *Studies in Second Language Acquisition*, *40*(3), 651-675.

Mäntylä, T. (2001). Incidental versus intentional memory. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 7262-7265). Oxford: Pergamon.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). Language Experience and Proficiency Questionnaire (LEAP. *Journal of Speech Language and Hearing Research*, *50*(4), 940-967.

Martinez, R. & V. Murphy. (2011). Effect of frequency and idiomaticity in second language reading comprehension. *TESOL Quarterly 45*(2), 267-290.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, *33*(3), 299-320.

McCauley SM, & Christiansen MH (2019a). Language learning as language use: A cross-linguistic model of child language development. *Psychological Review*, *126*, 1–51.

McCarthy, D., Keller, B., Carroll, J.(2003). Detecting a Continuum of Compositionality in Phrasal Verbs. In: *Proceedings of the ACL-2003 Workshop on Multi-word Expressions: Analysis, Acquisition and Treatment*.

McClure, E., & M. Mir (1995). Spanish-English codeswitching in the Mexican and Spanish press. *Journal of Linguistic Anthropology*, *5*(1): 33-50.

REFERENCES

McDonald, K., & Asaba, M. (2016). "'I Don't Know'" Use and Guessing on the Bilingual Japanese Vocabulary Size Test: Clarifications and Limitations. *Vocabulary Learning and Instruction*, *5*(1), 34-39.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory, and practice*. Cambridge University Press.

Meara, P. (1996). *The Vocabulary Knowledge Framework. Vocabulary Acquisition Research Group Virtual Library*. Swansea: Swansea University.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.

Meara, P. (2009). *Connected words: Word associations and second language vocabulary acquisition*. Amsterdam: John Benjamins Publishing.

Meara, P., & Wolter, B. (2004). V_Links: Beyond vocabulary depth. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), Angles on the English speaking world 4 (pp. 85-96). Copenhagen: Museum Tusculanum Press.

Melka, F. (1997) Receptive vs. productive aspects of vocabulary. In Schmitt and McCarthy (eds.) *Vocabulary: Description. Acquisition and Pedagogy*. Cambridge University Press, Cambridge : 84 -102.

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics, 32*, 111-129.

Meunier, F. (2020). Resources for learning multiword units. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp.336-350). Routledge.

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: New perspectives on assessment and corpus analysis* (pp. 57-78): EUROSLA monograph 2.

Milton, J., & Fitzpatrick, T. (2014). Introduction: Deconstructing vocabulary knowledge. In J. Milton & T. Fitzpatrick (Eds.), *Dimensions of vocabulary knowledge* (pp. 1-12). Basingstoke: Palgrave Macmillan.

Mitchell, R., Myles, F., & Marsden, E. J. (2013). *Second Language Learning Theories: third edition.* (3rd ed.) Routledge.

REFERENCES

Mondria, J.-A. (2003). The effects of inferring, verifying, and memorizing on the retention of L2 word meanings: An experimental comparison of the 'meaning-inferred method' and the 'meaning-given method'. *Studies in Second Language Acquisition*, *25*, 473-499.

Mondria, J,-A., & Wit-de Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics*, *12*, 249-267.

Mondria, J., & Wiersma, B. (2004). Active, Passive, and Active + Productive L2 Vocabulary Learning: What Difference Does It Make? In P. Bogaards, & B. Laufer (Eds.), *Vocabulary in a Second Language* (Vol. 5, pp. 79-100). Philadelphia, PA: John Benjamins B.V.

Moon, R. (1998a). *Fixed Expressions and Idioms in English: A Corpus-based Approach*. Oxford: Clarendon Press.

Moon, R. (1998b). Frequencies and forms of phrasal lexemes in English. In A. P. Cowie (Ed.), *Phraseology : Theory, Analysis, and Applications* (pp. 79-100). New York: Oxford University Press.

Moon, R. (1998c). Vocabulary connections: multi-word items in English. In M. McCarthy & N. Schmitt (Eds.), *Vocabulary: Description, Acquisition and Pedagogy* (pp. 40-63). Cambridge; New York: Cambridge University Press

Muñoz, C. (2006). *Age and the Rate of Foreign Language Learning.* Clevedon: Multilingual Matters.

Muñoz, C. (2008a) Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, *24*(4), 578-596.

Nagy, W. E. (1995) On the Role of Context in First and Second Language Vocabulary Learning. University of Illinois at Urbana-Champaign, Champaign. http://hdl.handle.net/2142/31277

Nagy, W. E., Herman, P., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, *20*(2), 233-253.

Nagy, W., E. McClure & M. Mir (1997). Linguistic transfer and the use of context by Spanish-English bilinguals. *Applied Psycholinguistics*, *18*, 431-452.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269-284). Erlbaum.

Nation, P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.

Nation, P. (1994). *New Ways in Teaching Vocabulary Alexandria*, Va: TESOL.

Nation, P. (1994). Review of three books on vocabulary teaching. *System, 22*(2),15-19.

REFERENCES

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, P. (2006) *Second language vocabulary*. In K. Brown (ed.) *Encyclopaedia of Language and Linguistics*, 2nd Ed. Oxford: Elsevier. Vol 13: 448-454.

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, *63*(1), 59-82.

Nation, P. (2007). The four strands. *Innovation in Language Learning and Teaching*, *1*(1), 1-12.

Nation, P. (2012). The BNC/COCA word family lists. Retrieved from http://www.victoria.ac.nz/lals/about/staff/paul-nation.

Nation, P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Nation, P. (2016). *Making and Using Word Lists for Language Learning and Testing*. Amsterdam: John Benjamins.

Nation, P. and Wang, K. (1999) Graded readers and vocabulary. *Reading in a Foreign Language 12*, 2: 355-380.

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, *31*(7), 9-13.

Nation, P., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.

Nation, I.P. and Webb, S. (2011) Content-based instruction and vocabulary learning. Eli Hinkel (ed) *Handbook of Research in Second Language Teaching and Learning*. Volume 2, Chapter 38, pages 631-644. New York: Routledge.

Nation, P. & Waring, R. (2020). *Teaching Extensive Reading in Another Language*. New York: Routledge.

Nattinger. J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford University Press.

Nattinger, J. R., & DeCarrico, J.S. (2000). *Lexical Phrases and Language Teaching*. Shanghai: Shanghai Foreign Language Education Press.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223-242.

Nesselhauf, N. (2005). *Collocations in a learner Corpus*. Amsterdam: John Benjamins Publishing Company.

Nguyen, T.M.H., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research, 21*(3), 298-320.

REFERENCES

Northbrook, J., & Conklin, K. (2019). Is what you put in what you get out?—Textbook-derived lexical bundle processing in beginner English learners. *Applied Linguistics, 40*(5), 816-833.

O'Donnell, M.(2008). The UAM CorpusTool: Software for corpus annotation and exploration. In Bretones Callejas, Carmen M. et al. (eds) *Applied Linguistics Now: Understanding Language and Mind / La Lingüística Aplicada Hoy: Comprendiendo el Lenguaje y la Mente*. Almería: Universidad de Almería.

O'Grady, W. (2005). Syntactic carpentry: An emergentist approach to syntax. Mahah, NJ: Erlbaum

O'Grady, W. (2008). The emergentist program. *Lingua*, *118*(4):447-464

Ozubko, J. D., Hourihan, K. L., & MacLeod, C. M. (2012). Production benefits learning: The production effect endures and improves memory for text. *Memory*, *20*(7), 717-727.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press

Papagno C., Valentine, T., & Baddeley, A. D. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*, 331-347.

Papagno, C., & Vallar., G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *Quarterly Journal of Experimental Psychology 44A*, 47-67.

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, *32*, 130-149. doi.org/10.1017/S0267190512000098

Paribakht, T. S., & Wesche, M. (1999). Reading and incidental L2 vocabulary acquisition. *Studies in Second Language Acquisition*, *21*, 195-224.

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*, 745-768. https://doi.org/10.1017/S0272263119000020

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Native-like selection and native-like fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191- 226). Longman.

Pellicer-Sánchez, A. (2017). Learning L2 collocations incidentally from reading. *Language Teaching Research, 21*(3), 381-402.

Pellicer-Sánchez, A. (2020). Learning single words vs. multiword items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp.158-173). Routledge.

REFERENCES

Pellicer-Sánchez, A., Schmitt, N. (2010) Incidental vocabulary acquisition from an authentic novel: Do 'Things Fall Apart'? *Reading in a Foreign Language, 22*(1), 31-55.

Pellicer-Sánchez, A., & Boers, F. (2019). Pedagogical approaches to the teaching and learning of formulaic language. In Siyanova-Chanturia A., Pellicer-Sánchez A. (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 153-173). New York: Routledge.

Penfield, W. & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage. *Journal of English for Academic Purposes*, *14*, 84-94.

Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research, 20*(1), 113-138.

Peters, E. (2019). The effect of imagery and on-screen text on foreign language vocabulary learning from audiovisual input. *TESOL Quarterly*, *53*(4), 1008-1032.

Peters, E. (2020). Factors affecting the learning of single-word items. In S. Webb (Ed.) *The Routledge handbook of vocabulary studies* (pp. 125-142). Routledge.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551-577.

Pienemann, M. (1998). *Language processing and second language development: Processability Theory*. Amsterdam: Benjamins.

Pienemann, M. (2005) *Cross-linguistic Aspects of Processability Theory*. John Benjamins Publishing Company: Amsterdam.

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. Reading in a Foreign Language, 18, 1-28.

Pinker, S. (1995). *The language instinct: how the mind creates language. London: Allen Lane, the Penguin Press.*

Puimège, E., & Peters, E. (2019). Learning L2 vocabulary from audiovisual input: an exploratory study into incidental learning of single words and formulaic sequences. *The Language Learning Journal*, *47*(4), 424-438.

REFERENCES

Puimège, E., & Peters, E. (2020). Learning formulaic sequences through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *42*(3), 525-549.

Pulido, D. (2009). How involved are American L2 learners of Spanish in lexical input processing tasks during reading. *Studies in Second Language Acquisition*, *31*, 31-58.

Pulido, M. F. & Dussias, P. (2020). Desirable difficulties while learning collocations in a second language: Conditions that induce L1 interference improve learning. *Bilingualism: Language & Cognition, 23*(3), 652-667.

Qian, D.D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, *56*, 282-308.

Qian, D.D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning*, *52*, 513-36.

Qian, D. D., & Schedl, M. (2004). Evaluation of an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing, 21*(1), 28-52.

R Core Team (2020). R: A language and environment for statistical computing (Version 4.0.1) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Ramírez, G., Chen, X., & Pasquarella, A. (2013). Cross-linguistic transfer of morphological awareness in Spanish-speaking English language learners: The facilitating effect of cognate knowledge. *Topics in Language Disorders*, *33*(1), 73-92.

Read, J. (2000). *Assessing Vocabulary* (Cambridge Language Assessment). Cambridge: Cambridge University Press.

Redwood, S., & Schmitt, N. (2011). Learner knowledge of phrasal verbs: A corpus-informed study. In F. Meunier, S. De Cock, G. Gilquin, & M. Paquot (Eds.), *A Taste for Corpora*. Amsterdam: John Benjamins.

Richards, J. C. (1976). The Role of Vocabulary Teaching. *TESOL Quarterly*, *10*, 77-89. https://doi.org/10.2307/3585941

Richards, J. C., & Schmidt, R. W. (2013). Longman dictionary of language teaching and applied linguistics. (4th ed). New York: Routledge.

Rieder, A. (2003). Implicit and explicit learning incidental vocabulary acquisition. Vienna English Working Papers.

Rodgers, M. P. (2018). The images in television programs and the potential for learning unknown words: The relationship between on-screen imagery and vocabulary. *ITL-International Journal of Applied Linguistics*, *169*(1), 191-211.

Rodgers, M. P., & Webb, S. (2020). Incidental vocabulary learning through viewing television. *ITL - International Journal of Applied Linguistics*, *171*(2), 191-220.

Rollin, N. (2008).The Oxford Spanish Dictionary : Spanish-English, English-Spanish. 4th ed., Carol Styles Carvajal. Oxford . Oxford University Press.

Rott, S. (1999) The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition through reading. *Studies in Second Language Acquisition, 21*(1), 589-619.

Rychlý, P. (2008). *A Lexicographer-Friendly Association Score*. Proc. 2nd Workshop on Recent Advances in Slavonic Natural Languages Processing.

Sadoski, M. (2005). A dual coding view of vocabulary learning. *Reading & Writing Quarterly*, *21*, 221-238. doi:10.1080/10573560590949359.

Sadoski, M., & Paivio, A. (2001). *Imagery and text: a dual coding theory of reading and writing*. Routledge.

Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences*, *6*(9), 382-386.

Schatz, E. K., & Baldwin, R. S. (1986). Context clues are unreliable predictors of word meanings. *Reading Research Quarterly*, *21*(4), 439-453.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129-158.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, *11*, 11-26.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu: Second Language Teaching and Curriculum Center, University of Hawai'i at Manoa.

Schmidt, R. (2001). *Attention*. In P. Robinson (Ed.), Cognition and second language instruction (pp. 3-32). Cambridge University Press.

REFERENCES

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of CLaSIC 2010, Singapore, December 2*-4 (pp. 721-737). National University of Singapore, Centre for Language Studies.

Schmitt, N. (1995). A fresh approach to vocabulary using a word knowledge framework. *RELC Journal* 26(1), 86-94.

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: Longitudinal study. *Language Learning*, 48(2), 281-317.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329-363.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. London: Palgrave.

Schmitt, N. (2012). Formulaic language and collocation. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-10). New York: Blackwell. doi:10.1002/9781405198431.wbeal0433.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. Language Learning, 64, 4, 913-951.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: word associations and verbal suffixes. *Studies in Second Language Acquisition*, 19, 17-36.

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, *18*(1), 55-88.

Schmitt, N. & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching. 47*(4), 484-503.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, *95*(1), 26-43.

Schmitt, N., & Carter, R. (2004). *Formulaic sequences in action: An introduction*. In N. Schmitt (Ed.), Formulaic Sequences: Acquisition, processing and use. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Schmitt, N, Sonbul, S., Vilkaité-Lozdiené, L., & Macis, M. (2019). Formulaic Language and Collocation. In *The Encyclopedia of Applied Linguistics.* Edited by Carol A. Chapelle. New York: John Wiley and Sons.

REFERENCES

Schmitt, N., & Schmitt, D. (2020). *Vocabulary in Language Teaching* (2nd ed.). Cambridge: Cambridge University Press. doi:10.1017/9781108569057

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10, 209-241.

Serrano, R., Stengers, H., & Housen, A. (2014). Acquisition of formulaic sequences in intensive and regular EFL programmes. *Language Teaching Research*, *19*(1), 89-106.

Shin, D., & Nation, P. (2008). Beyond single words: The most frequent collocations in spoken English. *ELT Journal*, *62*(4), 339-348.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, *31*(4), 487-512.

Sinclair, J. (1987). The Nature of the Evidence in J. Sinclair (ed.) *Looking Up: An Account of the COBUILD Project in Lexical Computing (pp.* 150-159)*. Collins ELT.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Sinclair, J. (2004). *Trust the text: Lexis, corpus, discourse*. London: Routledge.

Siyanova-Chanturia, A. (2015). On the 'holistic' nature of formulaic language : Corpus linguistics and Linguistic theory. *Corpus Linguistics and Linguistic Theory*, *11*(2), 285-301.

Siyanova-Chanturia, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review*, *64*(3), 429-458.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, *27*(2), 251-272.

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. (2011). Seeing a phrase 'time and again' matters: The role of phrasal frequency in the processing of multiword sequences*. Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 776-784.

Siyanova-Chanturia, A., & Webb, S. (2016). Teaching vocabulary in the EFL context. In W. A. Renandya & H. P. Widodo (Eds.), *English language teaching today: Linking theory and practice* (pp. 227-239). English Language Education, Vol. 5. Switzerland: Springer International Publishing.

Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (Eds.). (2019). Understanding formulaic language: A second language acquisition perspective. Routledge.

REFERENCES

Siyanova-Chanturia, A., & Omidian, T. (2020). Key Issues in Researching Multiword Items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 529-544). Routledge.

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Skinner, B. F. (1957). *Verbal behaviour*. New York: Appleton Century Crofts.

Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition*, *18*(3), 419-437.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, *63*(1), 121-159.

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, *36*, 139- 152. http://dx.doi.org/10.1080/09571730802389975

Stanovich, K. (1991). Changing models of reading and reading acquisition. In L. Rieberi & C. Perfetti (Eds.), Learning to read: Basic research and its implications (pp.19-32). Lawrence Erlbaum Associates.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360-407.

Sternberg, R.J. (1987). Most vocabulary is learned from context. In M. McKeown and M. Curtis (eds.) *The Nature of Vocabulary Acquisition Lawrence Erlbaum Associates*, Mahwah, N.J.: 89-105.

Stubbs, M. (2009). Memorial Article: John Sinclair (1933-2007) The Search for Units of Meaning: Sinclair on Empirical Semantics. *Applied Linguistics*, *30*(1), 115-137.

Stubbs, M. & Barth. I. (2003). Using recurrent phrases as text-type discriminators. A quantitative method and some findings. *Functions of Language*, *10*(1), 61-104.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognit. Sci,* 12, 257-285.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learn. Instr.* 4, 295-312.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*, 123-138.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185-233.

Sweller, J., Van Merrienboer J.J., & Paas, F.G. (1998). Cognitive architecture and instructional design. *Educational Psychology Review 10*(3): 251-296.

REFERENCES

Szudarski, P. (2023). *Collocations, Corpora and Language Learning (Elements in Corpus Linguistics).* Cambridge: Cambridge University Press.

Szudarski, P. (2012). Effects of meaning-and form-focused instruction on the acquisition of verb-noun collocations in L2 English. *Journal of Second Language Teaching & Research*, *1*(2), 3-37.

Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics, 26*(2), 245-265.

Taylor J.R. (2004). Why construction grammar is radical. *Annual Review of Cognitive Linguistics*, *2*, 321-348.

Teng, L.S., & Zhang, L.J. (2022). Can self-regulation be transferred to second/foreign language learning and teaching? Current status, controversies, and future directions. *Applied Linguistics*, *43*, 587-595.

Teng, F. (2019) The effects of context and word exposure frequency on incidental vocabulary acquisition and retention through reading. *The Language Learning Journal*, 47(2), 145-158.

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics, 11*(1/2), 61-82.

Tomasello, M. (2005). *Constructing a Language: A Usage-Based Theory of Language Acquisition.* Cambridge, Mass.: Harvard University Press

Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013*). Language Learning, 69*(2),405-439.

Tsai, K.-J. (2015). Profiling the collocation use in ELT textbooks and learner writing. *Language Teaching Research*, *19*(6), 723-740. https://doi.org/10.1177/1362168814559801

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, *69*(3), 559-599.

Van Lancker-Sidtis, D., & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language & Communication*, *24*(3), 207-240

van Zeeland, H. (2014). Lexical Inferencing in first and second language listening. *The Modern Language Journal*, *98*(4), 1006-1021.

van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*(3), 609-624.

REFERENCES

VanPatten, B. (1996). *Input Processing and Grammar Instruction: Theory and Research*. Norwood, NJ: Ablex.

VanPatten, B. & Williams, J. (2015). *Theories in Second Language Acquisition: An Introduction.* Routledge: New York.

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning, 61*(1), 219-258.

Vilkaitė, L. (2017). Incidental acquisition of collocations in L2: Effects of adjacency and prior vocabulary knowledge. *ITL-International Journal of Applied Linguistics, 168*(2), 248-277.

Vu, D. V., & Peters, E. (2020). Learning vocabulary from reading-only, reading-while-listening, and reading with textual input enhancement: Insights from Vietnamese EFL learners. *RELC Journal*, *53*(1), 85-100.

Vu, D. V., & Michel, M. (2021). An exploratory study on the aspects of vocabulary knowledge addressed in EAP textbooks. *Dutch Journal of Applied Linguistics*, *10*, 1-15.

Vu, D. V., & Michel, M. (in press). An exploratory study on the aspects of vocabulary knowledge addressed in EAP textbooks. *Dutch Journal of Applied Linguistics*, *10*, 1-15.

Vu, D. V., & Peters, E. (2022a). Incidental learning of collocations from meaningful input: a longitudinal study into three reading modes and factors that affect learning. Studies in Second Language Acquisition, 44(3), 685-707.

Vu, D. V., & Peters, E. (2022b). A longitudinal study on the effect of mode of reading on incidental collocation learning and predictors of learning gains. *TESOL Quarterly*. Advance online publication. https://doi.org/10.1002/tesq.3111

Vygotsky, L. S. (1986). *Thought and language*. A. Kozulin (Ed.), Cambridge, Massachusetts: The MIT Press.

Waring, R. & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*(2), 130-163.

Webb, S. (2007a). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*(1), 46-65.

Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, *20*(1), 232-245.

REFERENCES

Webb, S. (2019). Helping students become autonomous learners of vocabulary. *TESOL Connections*, 1-4.

Webb, S. (2020). Incidental vocabulary learning. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 225-239). Routledge.

Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly*, *43*(1), 55-77.

Webb, S. & Chang, A, C-S. (2012). Vocabulary learning through assisted and unassisted repeated reading. *Canadian Modern Language Review*, *68*(3), 267-290.

Webb, S., Newton, J., & Chang, A. C.-S. (2013). Incidental learning of collocation. *Language Learning*, *63*, 91-120.

Webb, S., & Chang, A. C. S. (2015). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning?. *Language Teaching Research*, *19*(6), 667-686.

Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford, UK: Oxford University Press.

Webb, S., & Chang, A. C. S. (2020). How does mode of input affect the incidental learning of collocations?. *Studies in Second Language Acquisition*, *44*(1), 35-56.

Webb, S., Uchihara, T., & Yanagisawa, A. (2023). How effective is second language incidental vocabulary learning? A meta-analysis. *Language Teaching*, 1-20.

Weinstein, Y. Sumeracki, M.; Caviglioli, O. (2019). *Understanding How We Learn*. Abingdon, Oxon; New York, NY : Routledge.

Wesche, M., & Paribakht, T. S. (1996). Assessing Second Language Vocabulary Knowledge: Depth vs. Breadth. *Canadian Modern Language Review*, *53*, 13-39.

Wilks, C., & Meara, P. (2007). Graph theory and word association networks. In H Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 167-181). Cambridge: Cambridge University Press.

Wolter, B. (2020). Key issues in teaching multiword items. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp.493-510). Routledge.

Wolter, B., & Gyllstad, H. (2011). Collocational links in the L2 mental lexicon and the influence of L1 intralexical knowledge. *Applied Linguistics*, *32*(4), 430-449.

REFERENCES

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition*, *35*(3), 451-482.

Wolter, B., & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, *36*(5), 1193-1221.

Wolter, B., & Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: What accounts for l2 performance? *Studies in Second Language Acquisition, 40*(2), 395-416.

Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence, and classroom applications*. London/New York: Continuum.

Wood, D. (2020). Classifying and identifying formulaic language. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 30-45). Routledge.

Wray, A. (2000). Formulaic sequences in second language teaching: Principles and practice. *Applied Linguistics 21*(4), 463-489.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Wray, A. (2004). 'Here's one I prepared earlier': Formulaic language learning on television. In N. Schmitt (Ed.), *Formulaic sequences: acquisition, processing and use* (pp. 249-268). Amsterdam: John Benjamins.

Wray, A. (2006). Formulaic language. In Brown, K., editor, *Encyclopedia of Language and Linguistics*, pages 590-597. Elsevier Press, Cambridge.

Wray, A. (2008). *Formulaic language: pushing the boundaries*. Oxford; New York: Oxford University Press.

Wulff, S. (2018). Acquisition of formulaic language from a usage-based perspective. In Siyanova-Chanturia, A., & Pellicer-Sánchez, A. (Eds.), *Understanding multi-word expressions: A second language acquisition perspective*. London, New York: Routledge.

Yamashita, J. (2018). Possibility of semantic involvement in the L1-L2 congruency effect in the processing of L2 collocations. *Journal of Second Language Studies*, *1*(1), 60-78.

Yamashita, J. & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly 44*(4), 647-668.

REFERENCES

Yanagisawa, A., & Webb, S. (2021a). To what extent does the involvement load hypothesis predict incidental L2 vocabulary learning? A meta-analysis. *Language Learning*, *71*(2), 487-536.

Yanagisawa, A., & Webb, S. (2021b). Involvement Load Hypothesis Plus: Creating an improved predictive model of incidental vocabulary learning. *Studies in Second Language Acquisition*. Advance online publication.

Zahar, R., Cobb, T. & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review*, *57*(4), 541-572.

Zhang, X. (2017). Effects of receptive-productive integration tasks and prior knowledge of component words on L2 collocation development. *System*, *66*, 156-167.

Appendices

_____

Appendix 1. Experiment 1

_____

Appendix 1A. Instruments

1. Participant's information sheet and consent form(*)

Participant's information sheet and consent form

Information sheet

As part of my PhD in the School of Cultures and Languages, I am carrying out a study of language acquisition by learners of English as a foreign language.

Because of the nature of the project, you will have to complete different sections at three different times. Each section will be given to you one by one, over the course of five weeks. At the beginning of each part, you will have written instructions for completion, as well as some examples.

In compensation for your participation, you have the opportunity to win one of five Amazon Vouchers worth 100€. Please, bear in mind that to enter this Prize Draw you must complete all sections.
Your participation in this study is voluntary. You may refuse to participate or withdraw from the study at any time. At every stage, your name will remain confidential. The data will be anonymized before the analysis and will be kept securely and used for academic purposes only.

Should you have any further questions about the study, please feel free to reach me by e-mail (i.vina@kent.ac.uk) or any of my PhD supervisors, Dr Gloria Chamorro (gchamorro@flog.uned.es) and Dr Christina Kim (c.s.kim@kent.ac.uk).

Thank you!

Inés de la Viña
i.vina@kent.ac.uk - you can also contact me on Twitter @inesdelavina

Assistant Lecturer in Linguistics
University of Kent
School of Cultures and Languages
Modern Language and Linguistics
CT2 7NF
Canterbury

CONSENT FORM

I have read and understand this consent form, and I volunteer to participate in this research project.

Please sign below:

(*)A similar consent form was also used in Experiments 2 and 3.

2. Language background questionnaire(*)

In order to help us better understand, interpret, and classify your answers, we would like to ask you some questions concerning your language experience, proficiency, and use. This questionnaire will take less than 10 minutes to complete. Please answer every question and give your answers sincerely.

PART 1: Biographical information

1. Age:

2. Gender:

        a) female
        b) male
        c) non-binary
        c) other
        d) prefer not to say

3. Country of birth:

4. Country of residence:

5. Highest level of formal education (degree obtained or school level attended):

- High school
- Vocational training and training/college
- Bachelors degree
- Masters degree
- Doctorate (PhD)
- Other

PART 2: Language history

In this section, we would like you to answer some factual questions about your language experience and proficiency

1. What is your native language? (If you grew up with more than one language, please specify)

2. Do you speak other languages?

__ YES, please specify:

__ NO (if you answered NO, you don't need to continue this form)

3. In which order did you learn the languages you know?

4. If you speak English, how many years have you been studying English in school/language centres?

5. Do you hold a recent (no older than 2 years) official English language certificate (i.e., IELTS, TOEFL, Cambridge, Pearson, etc.)?

__ YES

__ NO

5b. If you answered yes to the previous question, which certificate do you have? Please specify and provide your level, e.g., IELTS (Academic) B2 level, Cambridge Advanced C1

6. How did you learn English up to this point? (check all that apply)

Mainly Mostly Occasionally through formal classroom instruction

Mainly Mostly Occasionally through interacting with people

A mixture of both, but More classroom More interaction Equally both

Other (specify: _____)

7. Have you spent a long period (3 months or more) in an English-speaking country?

__ YES
__ NO

If you answer is yes, how long? _____

8. Which is your level of English?

    a) Beginner
    b) Intermediate
    c) Advanced

9. Rate your ability on the following English skills. Please rate according to the following scale (write down the number in the table):

10-point scale:
1 (very poor) to 10 (proficient)

	Reading skills	Writing skills	Speaking skills	Listening skills
English				

10. Please select which of the following factors have contributed to your learning:
You can select several factors.

- o Formal instruction at school
- o Private tuition (language centre, tutoring, etc.)
- o Interacting with friends
- o Interacting with family
- o Listening to radio/music
- o Watching TV and other media providers (Netflix, HBO, etc.)
- o Reading books and magazines
- o Visiting English websites
- o Playing computer and/or Internet games
- o Others (please specify) _____

Part 3: Language use

In this section, we would like you to answer some questions about your language use.

How much time per week do you spend..

1. reading books, magazines, newspapers in English, or visiting English websites?

   a) 0 hours
   b) An hour or less
   c) Between 1or 2 hours
   d) 2 hours or more

2. watching films, videos, or TV in English?

   a) 0 hours
   b) An hour or less
   c) Between 1or 2 hours
   d) 2 hours or more

3. listening to music in English?

   a) 0 hours
   b) An hour or less
   c) Between 1or 2 hours
   d) 2 hours or more

4. using English to keep in contact with people (using your phone, TikTok, WhatsApp, Facebook, Twitter, Instagram, Zoom, Skype, e-mail, etc.)

   a) 0 hours
   b) An hour or less
   c) Between 1or 2 hours
   d) 2 hours or more

5. playing computer/Internet games in English?

   a) 0 hours
   b) An hour or less
   c) Between 1or 2 hours
   d) 2 hours or more

Thank you for completing this questionnaire! I really appreciate it :-)

If you have any questions or comments, please write them below:

(*)The background questionnaire was also used in Experiments 2 and 3.

3. Reading task

In this part, we are going to ask you to read a short text titled "Sugar and Other Sweeteners"

- The text is split into different screens.
- Please, read the passages <u>only once.</u>
- When you are done reading one screen, move on to the next one by clicking "Next".
- Please, bear in mind that you cannot go back to a previously read screen.
- It takes approx. 15 minutes to read the whole text.

Read the passage <u>carefully</u>, take your time, and enjoy the reading!

4. Reading text

Target collocations are bolded below for ease of illustration. They were not bolded in the text seen by participants.

Sugar and other sweeteners

The sweetness of a substance results from physical contact between that substance and the many thousand **taste buds** of the tongue. The **taste buds** are clustered around several hundred small, fleshy protrusions called taste papilla which provide a large surface area for the **taste buds** and ensure maximum contact with a substance.

Although there are many millions of olfactory cells in the nose, taste is a more intense experience than smell; food technologists believe this is because of the strong pleasure relationship between the brain and food. And it is universally acknowledged that sweetness is the ultimate pleasurable taste sensation. For example, the French writer Marcel Proust is famous for using this idea in his work: eating a particular cake by chance one day brings back extremely **fond memories** of childhood for the narrator of his epic In Search of Past Time. The words 'sugar', 'honey' and 'sweetie' are used by lovers as terms of endearment. Pregnant women can often ward off **morning sickness** by eating something sweet. In Tudor times, to have teeth blackened by decay from eating too much sugar was seen as a desirable characteristic open only to the rich and aristocratic upper class. In fact, for the nobility, a traditional **hearty meal** almost always included sweet foods such as preserved fruit, gingerbread, sugared almonds, or jelly. Even recently, with the harm sugar can do much more widely known, advertisers have managed to create demand for sweet-tasting foods with the catch-phrase 'naughty but nice'. Makers of sugary and energetic drinks marketed toward children and teens are the worst offenders; because they drag them into developing long-term unhealthy eating habits. Too many people eat too much **saturated fat**, added salt, added sugars, and alcohol. The typical Western (American) diet is high in junk food, protein, and fat but low in fruits, vegetables, and fiber.

Our attraction to sweet-tasting foods is completely natural. We have evolved from times when food was scarce, and high calorie food was rewarding both to our **taste buds**, and for our survival. However, some people feel 'addicted' to sugar, and report **sugar cravings** that are often satisfied by eating a **chocolate bar** or a slice of cake. Sweet food certainly acts on the reward systems in our brains, and, for most of us, sweet food has a positive impact on mood, at least in the short-term. However, sugar can also exert a powerful influence over behaviour, making cutting it out of our diets very difficult. In fact, quitting eating a high sugar diet can even lead to **withdrawal symptoms**. The length of unpleasant **withdrawal symptoms** following a sugar detox varies. Some people quickly adjust to functioning without sugar, while others may experience severe cravings and find it very difficult to resist sugary foods.

Despite the attraction of all things sugary, however, no-one is sure what exactly makes a substance sweet. Nature is abundant with sweet foodstuffs, the most common naturally occurring substance being fructose, which is found in almost all fruits and berries. Of course, once eaten, all foods provide one or more of the three basic food components - protein, fat and carbohydrate - which eventually break down to supply the body with the essential sugar glucose. Nature also supplies us with sucrose, a naturally occurring sugar within the **sugar cane** plant, which was discovered and exploited many centuries BC. Sucrose, or table sugar, breaks down into glucose within the body. Nowadays, this white sugar is the food industry standard taste for sugar - the benchmark against which all other sweet tastes are measured. In the U.S.A. a number of foods, and especially **soft drinks**, are commonly sweetened with High Fructose Corn Syrup (HFCS), derived from corn starch by a process developed in the late 1960s. And man has further added to nature's repertoire by developing **artificial sweeteners** that are considered harmless, non-active chemicals with the additional property of sweetness, to cater for his sweet tooth.

While there is currently no scientific evidence that **artificial sweeteners** pose a risk to human health, excessive consumption of **artificial sweeteners** can cause undesirable **side effects** such as diarrhea and headaches. Additionally, even though **artificial sweeteners** are calorie-free, some research studies show that **artificial sweeteners** are associated with increased body weight, which suggests that people may be replacing the lost calories through other sources, possibly offsetting weight loss or health benefits. Evidence is inconclusive at this time, however, and the use of non-nutritive sweeteners is supported in moderation and as part of an overall balanced diet by numerous reputable organizations.

Sugar is, in its various forms, the **gold standard** for sweetness. And there is, indeed, an innate desire in humans (and some animals) to seek out and enjoy sugary food. Since sweet substances provide energy and sustain life, they have always been highly prized. All food manufacturers capitalise on this craving for sweetness by flavouring most processed foods with carefully measured amounts of sugar in one form or another. The maximum level of sweetness that can be attained before the intrinsic

taste of the original foodstuff is lost or unacceptably diminished is, in each case, determined by trial and error.

Furthermore, the most acceptable level of sweetness for every product -that which produces the optimum amount of pleasure for most people - is surprisingly constant, even across completely different cultures. This probably goes a long way towards explaining the almost universal appeal of Coca-Cola, although some people point to their flawless **marketing ploy.**

Artificial sweeteners cannot match the luxurious smoothness and mouth-feel of white sugar. Even corn syrup has a slightly lingering after-taste. The reason why food technologists have not yet been able to create a perfect alternative to sucrose (presumably a non-kilojoule-producing substitute) is simple. There is no molecular structure yet known that predisposes towards sweetness. In fact, there is no way to know for certain if a substance will taste sweet or even taste of anything at all. Our currently available **artificial sweeteners** were all discovered to be sweet purely by accident.

Sample screenshot as it appeared on the online platform:

Sweet food certainly acts on the reward systems in our brains, and, for most of us, sweet food has a positive impact on mood, at least in the short-term. However, sugar can also exert a powerful influence over behaviour, making cutting it out of our diets very difficult. In fact, quitting eating a high sugar diet can even lead to withdrawal symptoms. The length of unpleasant withdrawal symptoms following a sugar detox varies. Some people quickly adjust to functioning without sugar, while others may experience severe cravings and find it very difficult to resist sugary foods.

5. Reading comprehension test

*According to the reading, please state whether the following statements are **TRUE** or **FALSE**:*

1. Food tasting is a less intense experience than smelling.

2. Food technologists have created a synthetic alternative to sucrose.

3. In Tudor England, having black teeth from eating too much sugar was fashionable.

4. There is no evidence that humans have an innate desire to seek out and consume sweet-tasting foods.

5. Food manufacturers flavour processed foods with carefully measured amounts of sugar.

6. The ideal level of sweetness for every product is surprisingly consistent across different cultures.

7. Sweet food consumption can have a long-term negative impact on your mood.

8. Some studies suggest that artificial sweeteners can cause diarrhea and headaches.

6. Form recall

In this first test, you will find a word or a term in Spanish. Write the appropriate translation in English in the spaces provided. You will need two words for each term (one word in each blank). The first letter of each word is shown to help you.

Please <u>do not guess</u>. If you do not know the answer, leave it blank.

Here is an example:	Here is the answer:
tarta de cumpleaños:	**tarta de cumpleaños:**
b_____	**b**irthday
c_____	**c**ake

1) estómago vacío: e_____ s_____

2) caña de azúcar: s_____ c_____

3) papilas gustativas: t_____ b_____

4) tarta de zanahoria: c_____ c_____

5) mantequilla de cacahuete: p_____ b_____

6) buenos recuerdos: f_____ m_____

7) edulcorantes artificiales: a_____ s_____

8) pérdida de peso: w_____ l_____

9) efectos secudarios: s_____ e_____

10) antojo de azúcar: s_____ c_____

11) comida picante: s_____ f_____

12) fruta tropical: t_____ f_____

13) síndrome de abstinencia: w_____ s_____

14) estrategia de mercado: m_____ p_____

15) comida abundante: h_____ m_____

16) refrescos: s_____ d_____

17) comida rápida: f_____ f_____

18) dieta equilibrada: b_____ d_____

19) récord de ventas: r_____ s_____

20) grasas saturadas: s_____ f_____

21) comida procesada: p_____ f_____

22) aperitivo ligero: l_____ s_____

23) té helado: i_____ t_____

24) chocolatina: c_____ b_____

25) criterio de referencia: g_____ s_____

26) naúseas matutinas: m_____ s_____

7. Form recognition

In this task, you have to match *one word* from the *left-side* columns to a word from the *right-side* column to form two-word combinations in English. You can move the words on the left-hand side of the screen by dragging them (or double clicking) to the right-side column.

There are 26 two-word possible combinations in total. But don't worry if you are unable to find them all, just try to identify as many as you can.

You can also use the "I don't know" option card for the ones you are unsure/unable to match.

Screenshot of the matching task (as it appeared on the experimental platform)

Drag and drop (or double click) from A to B

Column A

morning	record	sugar
balanced	taste	fast
withdrawal	weight	fond
spicy	marketing	tropical
chocolate	empty	artificial
iced	sugar	light
hearty	carrot	side
peanut	gold	soft
saturated	processed	

Column B

butter	bar
fat	food
cake	ploy
drinks	sickness
standard	loss
snack	food
effects	symptoms
tea	buds
meal	diet
stomach	food
craving	cane
fruit	sales
sweeteners	sickness
	I don't know

Whenever possible, provide the Spanish translations of the two-word combinations you have identified. If you don't know the exact translation, you can provide a synonym or an explanation of the term.

Write the English term followed by the Spanish translation as shown in the example below:

e.g., apple pie: tarta de manzana

(*) Note: The translation task appeared in the pretest only.

8. Meaning recognition

Please select the appropriate choice out of the five options provided that best matches the meaning of the two-word combination. Only one is the correct option.
If you don't know the meaning of the two-word combination or you are unsure, please select option "e" (I don't know).

1. taste buds:
   a) sabor
   b) papilas gustativas
   c) degustación
   d) sentido del gusto
   e) I don't know

2. record sales:
   a) segundas rebajas
   b) récord de ventas
   c) artículos de liquidación
   d) récord de rebajas
   e) I don't know

3. fond memories:
   a) memorias agradables
   b) buenos recuerdos
   c) memorias de la infancia
   d) cálidos recuerdos
   e) I don't know

4. morning sickness:
   a) malestar de la mañana
   b) náuseas de la mañana
   c) malestar matutino
   d) náuseas matutinas
   e) I don't know

5. balanced diet:
   a) dieta blanda
   b) dieta saludable
   c) dieta equilibrada
   d) dieta balanceada
   e) I don't know

6. sugar cravings:
   a) caña de azúcar
   b) hambriento de azúcar
   c) antojo de dulce
   d) capricho de azúcar
   e) I don't know

7. withdrawal symptoms:
   a) tratamiento de desintoxicación
   b) síntomas de desintoxicación
   c) signos de abstinencia
   d) síndrome de abstinencia
   e) I don't know

8. sugar cane
   a) algodón de azúcar
   b) azúcar refinado
   c) terrón de azúcar
   d) caña de azúcar
   e) I don't know

9. side effects
   a) efectos secundarios
   b) efectos nocivos
   c) efectos colaterales
   d) efectos adversos
   e) I don't know

10. artificial sweeteners
    a) colorantes artificiales
    b) edulcorantes artificiales
    c) endulzantes artificiales
    d) dulcificantes artificiales
    e) I don't know

11. weight loss
    a) desnutrición
    b) pérdida de peso
    c) dieta
    d) régimen
    e) I don't know

12. gold standard
    a) estándar dorado
    b) estándar de calidad
    c) criterio de referencia
    d) patrón oro
    e) I don't know

13. marketing ploy
    a) estratagema empresarial
    b) estrategias de mercado
    c) campaña de promoción
    d) técnicas de marketing
    e) I don't know

14. chocolate bar
    a) chocolatina
    b) barra de chocolate
    c) bombón
    d) cacao
    e) I don't know

15. carrot cake
 a) tarta de manzana
 b) tarta de zanahoria
 c) bollo de manzana
 d) bollo de zanahoria
 e) I don't know

16. fast food:
 a) comida ligera
 b) comida grasa
 c) comida rápida
 d) comida ultraprocesada
 e) I don't know

17. body weight
 a) peso muscular
 b) masa corporal
 c) envergadura
 d) peso corporal
 e) I don't know

18. soft drinks
 a) gaseosas
 b) refrescos
 c) bebidas azucaradas
 d) bebidas carbonatadas
 e) I don't know

19. light snack
 a) comida liviana
 b) merienda ligera
 c) aperitivo ligero
 d) refrigerio
 e) I don't know

20. iced tea
 a) té verde
 b) té fresco
 c) té helado
 d) sorbete
 e) I don't know

21. peanut butter
 a) manteca de cacao
 b) manteca de cerdo
 c) mantequilla de cacahuete
 d) mantequilla de anarcados
 e) I don't know

22. spicy food
 a) comida caliente
 b) comida especiada
 c) comida sabrosa
 d) comida picante
 e) I don't know

23. tropical fruit
 a) fruta del trópico
 b) fruta de temporada
 c) fruta tropical
 d) fruta exótica
 e) I don't know

24. hearty meal
 a) comilona
 b) comida abundante
 c) comida casera
 d) comida del corazón
 e) I don't know

25. empty stomach
 a) estómago lleno
 b) estómago vacío
 c) hambriento
 d) malnutrido
 e) I don't know

26. saturated fat
 a) grasas saturadas
 b) grasa animal
 c) comida saturada
 d) grasa vegetal
 e) I don't know

9. Retrospective protocol analysis

Finally, we would like to ask you a few more questions about your participation in this experiment.
Please, answer the questions <u>sincerely</u>.
Your feedback is really important to us!

1. Were you familiar with the topic of the reading ("Sugar and Other Sweeteners")?
- o yes
- o no

1b. If you answered YES to the previous question, please rate your familiarity with the topic

Please feel free to use the input field below your rating to add more details

○	○	○
a little familiar	somewhat familliar	very familiar

2. Please rate the reading "Sugar and Other Sweeteners" in terms of your interest/engagement

○	○	○	○	○
not interesting at all	not very interesting	somewhat interesting	very interesting	extremely interesting

3. Please rate the level of difficulty of the reading

Please consider how difficult the reading was for you to understand

○	○	○	○	○
not difficult at all	not very difficult	somewhat difficult	very difficult	extremely difficult

4. Did you learn anything new about the topic of the text?

○	○	○
I didn't learn anything new	I learned something, but the topic was mostly familiar to me	I learned a lot about the topic

5. Did you learn any new words or expressions that you didn't know before reading the text?

○	○	○
I didn't learn any new expressions	I learned one or two new expressions	I learned a lot of new expressions that I didn't know before

Please feel free to add more feedback here:

6. If you were asked to listen to the audio while reading the text, please rate how much easier or harder the audio made it for you to read the text

○	○	○	○
made it much harder	made it somewhat harder	made it somewhat easier	made it much easier
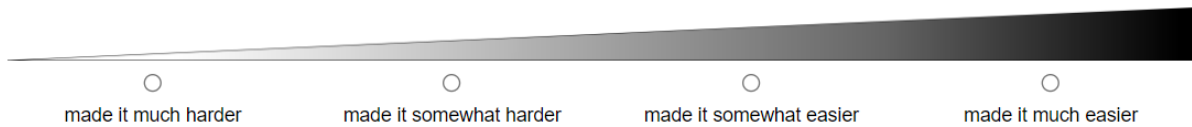
6b. Please elaborate on your answer to the previous question. In your experience, why do you think the audio made reading the text easier or more difficult?

[   ]

10. Semantic transparency task(*)

Semantic transparency norms were collected from 25 native speakers of Spanish with an advanced level of English and a university degree, who did not participate in the main experiment. The task was adapted from Macis and Schmitt (2017).

Some words co-occur in a language more common than others. These are called collocations.

The difficulty in learning collocations lies on a continuum of semantic transparency. Semantic transparency means the meaning of a collocation can be more easy or difficult to guess based on the literal meaning of its component parts.

For example, if a learner knows the meaning of the two words making up a transparent collocation, then that collocation can be understood through decoding the constituents in their literal sense ("food allergy" = alergia alimentaria). However, low transparent collocations meanings cannot be guessed even if a learner knows the meaning of both components ("melting pot"= crisol).

Please rate each collocation's meaning on a scale from 1 to 4: 1 being very difficult to guess (very opaque/non-transparent), and 4 being very easy to guess (very transparent). The Spanish equivalents are provided in brackets in case you are not familiar with the word combination.

1. gum disease (= enfermedad de las encías)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

2. withdrawal symptoms (= síndrome de abstinencia)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

3. sugar cane (= caña de azúcar)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

4. artificial sweeteners (= edulcorantes artificiales)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

5. marketing ploy (= estrategia de marketing)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

6. side effects (= efectos secundarios)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

7. head lice (= piojos)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

8. cerebral palsy (= parálisis cerebral)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

9. heat rash (= sarpullido)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

10. acute pain (= dolor agudo)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

11. stem cells (= células madre)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

12 . kidney failure (= fallo renal)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

13. yeast infection (= infección de hongos)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

14. chief complaint (= motivo de la consulta)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

15. wasp sting (= picadura de avispa)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

16. rib cage (= caja torácica)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

17. stretch marks (= estrías)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

18. stomach cramps (= retortijones de barriga)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

19. taste buds (= papilas gustativas)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

20. fond memories (= buenos recuerdos)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

21. morning sickness (= náuseas)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

22. hot flushes (= sofocos)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

23. bedside manner (= trato al paciente)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

24. hay fever (= alergia al polen)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

25. cold sore (= herpes labial)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

26. hearty meal (= comida copiosa)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

27. gold standard (= punto de referencia)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

28. soft drinks (= refrescos)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

29. chocolate bar (= tableta de chocolate)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

30. bone marrow (= médula espinal)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

31. saturated fat (= grasas saturadas)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

32. sugar craving (= antojo de dulce)

	1	2	3	4	
very difficult (very opaque)	○	○	○	○	very easy (very transparent)

(*)This task was used in Experiments 1 and 2, it contains the target items of both experiments.

Appendix 1B. Regression models

This appendix provides the final regression models for Experiment 1, described in Chapter 4.

Form recall models

- **Immediate posttest regression model:**

Model:	
Response ~	ReadingMode + VocabSize + Congruency + Repetition
	+ VocabSize:Congruency + Congruency:ReadingMode
	+ (1+VocabSize+ReadingMode|ParticipantID) + (1+Congruency|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.	
ID	(Intercept)	0.13	0.36		
	VocabSize	0.27	0.52	0.27	
	Reading	0.34	0.59	-0.60	0.61
ITEM	(Intercept)	0.33	0.57		
	Congruency	0.23	0.47	-0.63	

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	-0.25	0.23	-1.067	0.287
ReadingMode	-0.096	0.15	-0.68	0.51
VocabSize	0.57	0.17	3.47	0.000***
Congruency	0.56	0.21	2.67	0.005***
Repetition	0.28	0.16	1.74	0.082.
VocabSize:Congruency	-0.10	0.085	-1.19	0.23
ReadingMode:Congruency	-0.002	0.080	-0.037	0.97

Correlation of Fixed effects:						
	(Intr)	Vocab.	Congr.	Reading	Repetition	Vocab:Con
Reading	-0.089	0.39	-0.006			
Vocab.	-0.031					
Congr.	0.40	0.022				
Repetition	-0.008	0.014	-0.106	0.000		
Vocab:Con	-0.042	-0.039	-0.026	-0.027	0.023	
Read:Con	-0.007	-0.022	-0.008	0.012	-0.003	0. 27

- **Delayed posttest regression model:**

Model:

Response ~ ReadingMode + VocabSize + Congruency + Repetitions
+ VocabSize:Congruency + Congruency:ReadingMode
+ (1+VocabSize+ReadingMode|ParticipantID) + (1+Congruency|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	0.90	0.95	
	Repetition	1.043	1.021	0.69
ITEM	(Intercept)	1.89	1.38	

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	2.30	0.45	5.18	2.92e-07***
ReadingMode	-0.046	0.15	-0.31	0.76
VocabSize	0.80	0.16	4.98	6.47e-07***
Congruency	1.31	0.42	3.14	0.001**
Repetition	0.69	0.53	1.31	0.19
VocabSize:Congruency	0.26	0.12	2.15	0.031*
ReadingMode:Congruency	0.21	0.12	1.82	0.069.

Correlation of Fixed effects:

	(Intr)	Vocab.	Congr.	Reading	Repetition	Vocab:Con
Reading	0.015	0.20	0.17			
Vocab.	0.13					
Congr.	0. 25	0.140				
Repetition	0.167	-0.078	-0.12	0.003		
Vocab:Con	0.14	0.43	0.16	0.090	-0.038	
Read:Con	0.043	0.11	0.026	0.44	0.014	0.20

Form recognition models

- **Immediate posttest regression model:**

Model:

Response ~   ReadingMode + VocabSize + Congruency + Congruency:ReadingMode
+ (1+Congruency|ParticipantID) + (1+Congruency|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	0.89	0.94	
	Congruency	0.16	0.41	-0.50
ITEM	(Intercept)	0.38	0.62	
	Congruency	0.36	0.60	-0.50

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	1.92	0.29	6.65	2.89e-11***
ReadingMode	0.068	0.16	0.43	0.667
VocabSize	0.97	0.17	5.82	5.78e-09***
Congruency	0.267	0.26	1.016	0.31
ReadingMode:Congruency	0.003	0.11	0.034	0.97

Correlation of Fixed effects:

	(Intr)	VocabSize	Congruency	ReadingMode
VocabSize	0.18			
Congruency	-0.24	0.10		
ReadingMode	0.017	0. 24	0.017	
Reading:Congruency	-0.014	-0.019	-0.041	-0.044

- **Delayed posttest regression model:**

Model:

Response ~ ReadingMode + VocabSize + Congruency + Congruency:ReadingMode
+ (1+Congruency|ParticipantID) + (1+Congruency|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	0.76	0.90	
	Congruency	0.18	0.35	-0.40
ITEM	(Intercept)	0.48	0.56	
	Congruency	0.35	0.57	-0.34

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	6.59	1.38	4.77	1.81e-06***
ReadingMode	0.44	0.58	0.760	0.45
VocabSize	0.88	0.55	1.634	0.10
Congruency	0.72	0.25	2.90	0.003**
ReadingMode:Congruency	-0.573	0.24	-2.340	0.019*

Correlation of Fixed effects:

	(Intr)	VocabSize	Congruency	ReadingMode
ReadingMode	0.082	0.34	-0.18	
VocabSize	0.13			
Congruency	0.17	0.014		
ReadingMode:Congruency	-0.11	-0.010	0. 27	0. 26

Meaning recognition models

- ▪ **Immediate posttest meaning recognition model:**

Model:	
Response ~	ReadingMode + VocabSize + Congruency + Repetition
	+ (1+VocabSize+Repetition|ParticipantID)+(1+VocabSize+ Repetition|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.	
ID	(Intercept)	0.60	0.78		
	VocabSize	0.21	0.46	0.77	
	Repetition	0.96	0.98	-0.92	-0.96
ITEM	(Intercept)	1.80	1.34		
	VocabSize	0.01	0.11	-0.95	
	Repetition	0.96	0.43	-1.00	0.96

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	-0.53	0.89	-0.57	0.57
ReadingMode	-0.19	0.28	-0.69	0.49
VocabSize	0.079	0.17	0.46	0.65
Congruency	2.79	0.70	3.97	7.03e-05***
Repetition	1.69	0.69	2.45	0.014*

Correlation of Fixed effects:

	(Intr)	VocabSize	Congruency	Repetition
ReadingMode	-0.15	0. 23	-0.024	-0.005
VocabSize	-0.033			
Congruency	-0.43	-0.010		
Repetition	-0.90	-0.10	0.24	

- **Delayed posttest meaning recognition model:**

Model:

Response ~   ReadingMode + VocabSize + Congruency + Repetition

+ (1+VocabSize+Repetition|ParticipantID)+(1+VocabSize+RepetitionItemID)

Random effects:

Groups Name		Variance	Std. Dev.	Corr.
ID	(Intercept)	1.82	1.35	
	VocabSize	0.032	0.18	0.16
	Repetition	2.68	1.64	-0.93 -0.50
ITEM	(Intercept)	0.18	0.43	
	VocabSize	0.0073	0.086	1.00
	Repetition	0.020	0.14	-0.99 -1.00

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	-0.64	1.98	-0.32	0.75
ReadingMode	-0.31	0.39	-0.79	0.43
VocabSize	-0.14	0.25	-0.71	0.48
Congruency	4.29	1.15	3.73	0.000***
Repetition	2.82	1.86	1.52	0.3

Correlation of Fixed effects:

	(Intr)	VocabSize	Congruency	Repetition
ReadingMode	-0.20	0.29	-0.010	0.082
VocabSize	0.13			
Congruency	-0.15	-0.065		
Repetition	-0.98	-0.17	0.14	

Appendix 2. Experiment 2

Appendix 2A. Instruments

1. Reading task

In this part, we are going to ask you to read a few passages on different topics:

- The passages are split into different screens.
- Please, read the passages only <u>once</u>.
- When you are done reading one passage, move on to the next one by clicking "Next".
- Please, bear in mind that you cannot go back to a previously read screen.
- It takes approx. 20-30 minutes approx. to read all passages.

Take your time and enjoy the reading!

2. Reading materials per condition (1, 3, and 5 repetitions)

Target collocations are bolded below for ease of illustration. They were not bolded in the text seen by participants. 1rep=1 repetition, 3rep=3 repetitions, 5rep=5 repetitions.

1 repetition

(1)
1rep: Advanced **gum disease** can lead to the loss of tissue and bone that support the teeth and it may become more severe over time. Warning signs include red, swollen or tender gums. Some people may also experience persistent bad breath or a bad taste in their mouth. Good dental care at home is essential to help keep this condition from becoming more serious or recurring. Brushing and flossing teeth daily will help keep your teeth and gums healthy. Dental professionals also recommend fluoride toothpaste and oral irrigation as a great way to reduce plaque.

3rep: Advanced **gum disease** can lead to the loss of tissue and bone that support the teeth and it may become more severe over time. Warning signs include red, swollen or tender gums. Some people may also experience persistent bad breath or a bad taste in their mouth. Good dental care at home is essential to help keep **gum disease** from becoming more serious or recurring. Brushing and flossing teeth daily will help keep your teeth and gums healthy. Dental professionals also recommend fluoride toothpaste and oral irrigation as a great way to reduce plaque and prevent **gum disease**.

5rep: Advanced **gum disease** can lead to the loss of tissue and bone that support the teeth and it may become more severe over time. Warning signs of **gum disease** include red, swollen or

209

tender gums. Some people with **gum disease** may also experience persistent bad breath or a bad taste in their mouth. Good dental care at home is essential to help keep **gum disease** from becoming more serious or recurring. Brushing and flossing teeth daily will help keep your teeth and gums healthy. Dental professionals also recommend fluoride toothpaste and oral irrigation as a great way to reduce plaque and prevent **gum disease**.

(2)

1rep: Stretch marks can appear on your stomach, or sometimes on your upper thighs and breasts, as your pregnancy progresses and your bump starts to grow. When they appear will be different from woman to woman. You'll start to notice them around the end of the second trimester into the beginning of the third trimester, when you're between 6 and 7 months pregnant. That said, they do sometimes appear sooner. Up to 90 percent of all expecting women get these pink, red, brown or sometimes purplish streaks.

3rep: Stretch marks can appear on your stomach, or sometimes on your upper thighs and breasts, as your pregnancy progresses and your bump starts to grow. When they appear will be different from woman to woman. You'll start to notice **stretch marks** around the end of the second trimester into the beginning of the third trimester, when you're between 6 and 7 months pregnant. That said, **stretch marks** do sometimes appear sooner. Up to 90 percent of all expecting women get these pink, red, brown or sometimes purplish streaks.

5rep: Stretch marks can appear on your stomach, or sometimes on your upper thighs and breasts, as your pregnancy progresses and your bump starts to grow. When **stretch marks** appear will be different from woman to woman. You'll start to notice **stretch marks** around the end of the second trimester into the beginning of the third trimester, when you're between 6 and 7 months pregnant. That said, **stretch marks** do sometimes appear sooner. Up to 90 percent of all expecting women get these pink streaks known as **stretch marks**.

(3)

1rep: The goal of a **bone marrow** transplant is to cure many serious blood diseases and types of cancer including leukaemia, lymphoma and myeloma. This transplant takes healthy blood-forming cells from a donor and infuses those cells into the patient's bloodstream, where the cells may begin to grow and produce healthy red blood cells and white blood cells. The preparations for a transplant like this vary depending on the type of transplant, the disease needing transplant, and your tolerance for certain medicines. Before and after the transplant, blood samples and transfusions are common.

3rep: The goal of a **bone marrow** transplant is to cure many serious blood diseases and types of cancer including leukaemia, lymphoma and myeloma. A **bone marrow** transplant takes healthy blood-forming cells from a donor and infuses those cells into the patient's bloodstream, where the cells may begin to grow and produce healthy red blood cells and white blood cells. The preparations for a transplant like this vary depending on the type of transplant, the disease

needing transplant, and your tolerance for certain medicines. Before and after the **bone marrow** transplant, blood samples and transfusions are common.

5rep: The goal of a **bone marrow** transplant is to cure many serious blood diseases and types of cancer including leukaemia, lymphoma and myeloma. A **bone marrow** transplant takes healthy blood-forming cells from a **bone marrow** donor and infuses those cells into the patient's bloodstream, where the cells may begin to grow and produce healthy red blood cells and white blood cells. The preparations for a **bone marrow** transplant vary depending on the type of transplant, the disease needing transplant, and your tolerance for certain medicines. Before and after the **bone marrow** transplant, blood samples and transfusions are common.

(4)

1rep: Menopause can cause symptoms like anxiety, mood swings, and irregular periods, although **hot flushes** are considered the hallmark symptom of menopause. Sometimes called 'the change of life', going through the menopause is a natural biological process. Most women experience these for 6 months to 2 years, although some reports suggest that they can last considerably longer. In the US, up to 75 percent of women report experiencing these transient sensations of body heat that can occur during the day or night. Women of Afro-Caribbean origin have been shown to have them last for more years than those of white women.

3rep: Menopause can cause symptoms like anxiety, mood swings, and irregular periods, although **hot flushes** are considered the hallmark symptom of menopause. Sometimes called 'the change of life', going through the menopause is a natural biological process. Most women experience **hot flushes** for 6 months to 2 years, although some reports suggest that they can last considerably longer. In the US, up to 75 percent of women report experiencing these transient periods of heat that can occur during the day or night. Women of Afro-Caribbean origin have been shown to have **hot flushes** which last for more years than those of white women.

5rep: Menopause can cause symptoms like anxiety, mood swings, and irregular periods, although **hot flushes** are considered the hallmark symptom of menopause. Sometimes called 'the change of life', going through the menopause is a natural biological process. Most women experience **hot flushes** for 6 months to 2 years, although some reports suggest that **hot flushes** can last considerably longer. In the US, up to 75 percent of women report experiencing **hot flushes** that can occur during the day or night. Women of Afro-Caribbean origin have been shown to have **hot flushes** which last for more years than those of white women.

(5)

1rep: Patients want physicians who are knowledgeable about their specific medical issue, but also are more likely to trust doctors with good **bedside manners**. Every physician needs to know how to optimize this skill to be an effective medical professional. Fundamentally, the best way to develop it is to imagine your patient as a family member or even as yourself. Learning good doctor-patient relationships brings the focus back to the human side of medicine and results in

the best patient care. Whether you're still in school or just beginning your career in medicine, developing this skill is critical for your success as a physician.

3rep: Patients want physicians who are knowledgeable about their specific medical issue, but also are more likely to trust doctors with good **bedside manners**. Every physician needs to know how to optimize this skill to be an effective medical professional. Fundamentally, the best way to develop **bedside manners** is to imagine your patient as a family member or even as yourself. Learning good doctor-patient relationships brings the focus back to the human side of medicine and results in the best patient care. Whether you're still in school or just beginning your career in medicine, developing good **beside manners** is critical for your success as a physician.

5rep: Patients want physicians who are knowledgeable about their specific medical issue, but also are more likely to trust doctors with good **bedside manners**. Every physician needs to know how to optimize their **bedside manners** to be an effective medical professional. Fundamentally, the best way to develop a good **bedside manner** is to imagine your patient as a family member or even as yourself. Learning good **bedside manners** brings the focus back to the human side of medicine and results in the best patient care. Whether you're still in school or just beginning your career in medicine, developing good **beside manners** is critical for your success as a physician.

(6)

1rep: The **rib cage** is an excellent example of the human body's multi-faceted and multi-functional design. This basketlike enclosure has three essential functions: protection, support, and respiration. It is strong enough to support the skeleton and protect the vital organs in the chest cavity. It surrounds the lungs and the heart, which are two of our most important organs. The small joints between the ribs and the vertebrae permit a gliding motion of the ribs on the vertebrae during breathing and other activities. It is also flexible enough to expand and contract.

3rep: The **rib cage** is an excellent example of the human body's multi-faceted and multi-functional design. This basketlike enclosure has three essential functions: protection, support, and respiration. **The rib cage** is strong enough to support the skeleton and protect the vital organs in the chest cavity. It surrounds the lungs and the heart, which are two of our most important organs. The small joints between the ribs and the vertebrae permit a gliding motion of the ribs on the vertebrae during breathing and other activities. The **rib cage** is also flexible enough to expand and contract.

5rep: The **rib cage** is an excellent example of the human body's multi-faceted and multi-functional design. The **rib cage** has three essential functions: protection, support, and respiration. **The rib cage** is strong enough to support the skeleton and protect the vital organs in the chest cavity. The **rib cage** surrounds the lungs and the heart, which are two of our most important organs. The small joints between the ribs and the vertebrae permit a gliding motion of the ribs on the

vertebrae during breathing and other activities. The **rib cage** is also flexible enough to expand and contract.

(7)

1rep: If you're someone who regularly gets **hay fever** at particular times of the year, you might be able to spot your typical symptoms when they start. Having a runny or blocked nose is a very common symptom of this unpleasant allergic reaction. It is important to note that, unlike a cold, this condition isn't caused by a virus, is caused by an allergic response to outdoor or indoor allergens. There's currently no cure for this condition, but you can do things to ease your symptoms when the pollen count is high.

3rep: If you're someone who regularly gets **hay fever** at particular times of the year, you might be able to spot your typical symptoms when they start. Having a runny or blocked nose is a very common symptom of **hay fever**. It is important to note that, unlike a cold, this condition isn't caused by a virus, is caused by an allergic response to outdoor or indoor allergens. There's currently no cure for **hay fever**, but you can do things to ease your symptoms when the pollen count is high.

5rep: If you're someone who regularly gets **hay fever** at particular times of the year, you might be able to spot the typical **hay fever** symptoms when they start. Having a runny or blocked nose is a very common symptom of **hay fever**. It is important to note that, unlike a cold, **hay fever** isn't caused by a virus, is caused by an allergic response to outdoor or indoor allergens. There's currently no cure for **hay fever**, but you can do things to ease your symptoms when the pollen count is high.

(8)

1rep: When the virus is active, medications are used to help reduce and prevent **cold sore** symptoms, but there is no cure for it. In some people, they tend to erupt following a trigger event such as a cold. Other triggers include stress, sunlight, feverish illnesses, like influenza or chest infections. These small blisters tend to form on or near the lips and inside the mouth and they usually recur in the same place. If you regularly get them, use antiviral creams as soon as you recognise the early tingling feeling.

3rep: When the virus is active, medications are used to help reduce and prevent **cold sore** symptoms, but there is no cure for it. In some people, **cold sores** tend to erupt following a trigger event such as a cold. Other triggers include stress, sunlight, feverish illnesses, like influenza or chest infections. These small blisters tend to form on or near the lips and inside the mouth and they usually recur in the same place. If you regularly get **cold sores**, use antiviral creams as soon as you recognise the early tingling feeling.

5rep: When the virus is active, medications are used to help reduce and prevent **cold sore** symptoms, but there is no cure for it. In some people, **cold sores** tend to erupt following a trigger event such as a cold. Other triggers of **cold sores** include stress, sunlight, feverish illnesses, like influenza or chest infections. These small blisters tend to form on or near the lips

and inside the mouth and they usually recur in the same place. If you regularly get **cold sores**, use antiviral creams as soon as you recognise the early tingling feeling of a **cold sore**.

(9)

1rep: It is important to note that **head lice** are not a health hazard or a sign of poor hygiene and are not responsible for the spread of any disease. Though they may be a nuisance, they don't cause serious illness or carry any diseases. Despite this knowledge, there is significant stigma resulting from these insects infestations in many developed countries, resulting in children being ostracized from their schools, friends, and other social events. All socioeconomic groups are affected, and infestations are seen throughout the world. Such infestation is not significantly influenced by hair length or by frequent brushing or shampooing.

3rep: It is important to note that **head lice** are not a health hazard or a sign of poor hygiene and are not responsible for the spread of any disease. Though they may be a nuisance, they don't cause serious illness or carry any diseases. Despite this knowledge, there is significant stigma resulting from these insects infestations in many developed countries, resulting in children with **head lice** being ostracized from their schools, friends, and other social events. All socioeconomic groups are affected, and **head lice** infestations are seen throughout the world. Such infestation is not significantly influenced by hair length or by frequent brushing or shampooing.

5rep: It is important to note that **head lice** are not a health hazard or a sign of poor hygiene and are not responsible for the spread of any disease. Though they may be a nuisance, they don't cause serious illness or carry any diseases. Despite this knowledge, there is significant stigma resulting from **head lice** infestations in many developed countries, resulting in children with **head lice** being ostracized from their schools, friends, and other social events. All socioeconomic groups are affected, and **head lice** infestations are seen throughout the world. **Head lice** infestation is not significantly influenced by hair length or by frequent brushing or shampooing.

(10)

1rep: The handicap known as **cerebral palsy** affects around one in every 500 births. The abnormalities of muscle control that define this handicap are often accompanied by other neurological and physical abnormalities. The symptoms and severity are quite variable. Those with it may have only minor difficulty with fine motor skills, such as grasping and manipulating items with their hands. A severe form of this handicap could involve significant muscle problems in all four limbs, mental retardation, seizures, and difficulties with vision, speech, and hearing.

3rep: The handicap known as **cerebral palsy** affects around one in every 500 births. The abnormalities of muscle control that define this handicap are often accompanied by other neurological and physical abnormalities. The symptoms of **cerebral palsy** and their severity are quite variable. Those with it may have only minor difficulty with fine motor skills, such as

grasping and manipulating items with their hands. A severe form of **cerebral palsy** could involve significant muscle problems in all four limbs, mental retardation, seizures, and difficulties with vision, speech, and hearing.

5rep: The handicap known as **cerebral palsy** affects around one in every 500 births. The abnormalities of muscle control that define **cerebral palsy** are often accompanied by other neurological and physical abnormalities. The symptoms of **cerebral palsy** and their severity are quite variable. Those with **cerebral palsy** may have only minor difficulty with fine motor skills, such as grasping and manipulating items with their hands. A severe form of **cerebral palsy** could involve significant muscle problems in all four limbs, mental retardation, seizures, and difficulties with vision, speech, and hearing.

(11)

1rep: If you've ever had a sudden, uncontrolled, tight feeling in the muscles of your stomach, then you've probably had **stomach cramps**. They're uncomfortable and sometimes hurt. Most of the time, these aches aren't serious and don't need to be diagnosed. If they happen often, are severe, or last for more than a day, this could be a sign of a more harmful medical problem, and you should see a doctor. This discomfort in your belly may be caused by intestinal gas, food poisoning, constipation or ulcers. Some people also experience an nausea and vomiting.

3rep: If you've ever had a sudden, uncontrolled, tight feeling in the muscles of your stomach, then you've probably had **stomach cramps**. They're uncomfortable and sometimes hurt. Most of the time, **stomach cramps** aren't serious and don't need to be diagnosed. If they happen often, are severe, or last for more than a day, this could be a sign of a more harmful medical problem, and you should see a doctor. **Stomach cramps** result in discomfort in your belly, which may be caused by intestinal gas, food poisoning, constipation or ulcers. Some people also experience nausea and vomiting.

5rep: If you've ever had a sudden, uncontrolled, tight feeling in the muscles of your stomach, then you've probably had **stomach cramps**. **Stomach cramps** are uncomfortable and sometimes hurt. Most of the time, **stomach cramps** aren't serious and don't need to be diagnosed. If they happen often, are severe, or last for more than a day, this could be a sign of a more harmful medical problem, and you should see a doctor. **Stomach cramps** result in discomfort in your belly, which may be caused by intestinal gas, food poisoning, constipation or ulcers. Some people with **stomach cramps** also experience nausea and vomiting.

(12)

1rep: Heat rash is common in people from cooler climates who travel to warmer climates. The most common trigger for this irritation is exposure to heat for a long time. This may be especially true in very humid areas where the sweat has a harder time evaporating off the skin. Infants and obese people are also more susceptible to it. Any drugs that raise the body temperature or alter the function of the sweat glands can also increase the risk of developing it.

3rep: Heat rash is common in people from cooler climates who travel to warmer climates. The most common trigger for **heat rash** is exposure to heat for a long time. This may be especially true in very humid areas where the sweat has a harder time evaporating off the skin. Infants and obese people are also more susceptible to **heat rash**. Any drugs that raise the body temperature or alter the function of the sweat glands can also increase the risk of developing it.

5rep: Heat rash is common in people from cooler climates who travel to warmer climates. The most common trigger for **heat rash** is exposure to heat for a long time. This may be especially true in very humid areas where the sweat has a harder time evaporating off the skin. Infants and obese people are also more susceptible to **heat rash**. Any drugs that raise the body temperature or alter the function of the sweat glands can also increase the risk of **heat rash**. **Heat rash** will usually go away on its own, but it may have bothersome symptoms.

(13)

1rep: Severe untreated **acute pain** may predispose to the development of chronic pain. Treatment is thus essential to facilitate recovery from surgery or trauma by enabling early mobilisation and avoiding complications. The key decisional dilemma involves the selection of interventions to provide pain relief. Opioids are very effective in treating this sort of pain and are best used as part of a multimodal analgesic approach in combination with paracetamol and local anaesthetics. An opioid pain medication prescription for 7 or fewer days appears sufficient for most patients in primary care.

3rep: Severe untreated **acute pain** may predispose to the development of chronic pain. Treatment is thus essential to facilitate recovery from surgery or trauma by enabling early mobilisation and avoiding complications. The key decisional dilemma for treating **acute pain** involves the selection of interventions to provide pain relief. Opioids are very effective in treating this sort of pain and are best used as part of a multimodal analgesic approach in combination with paracetamol and local anaesthetics. An opioid pain medication prescription for 7 or fewer days appears sufficient for most patients with **acute pain** in primary care.

5rep: Severe untreated **acute pain** may predispose to the development of chronic pain. The treatment of **acute pain** is thus essential to facilitate recovery from surgery or trauma by enabling early mobilisation and avoiding complications. The key decisional dilemma for treating **acute pain** involves the selection of interventions to provide pain relief. Opioids are very effective in treating **acute pain** and are best used as part of a multimodal analgesic approach in combination with paracetamol and local anaesthetics. An opioid pain medication prescription for 7 or fewer days appears sufficient for most patients with **acute pain** in primary care.

(14)

1rep: For most people a single bee or **wasp sting** is one too many. Insect bites and stings will usually cause a red, swollen lump to develop on the skin. This may be painful and in some cases can be very itchy. The symptoms will normally improve within a few hours or days, although sometimes they can last a little longer. Most insect bites are not serious, but occasionally they

can become infected, cause a severe allergic reaction (anaphylaxis) or spread serious illnesses such as malaria.

3rep: For most people a single bee or **wasp sting** is one too many. Insect bites and stings will usually cause a red, swollen lump to develop on the skin. **Wasp stings** may be painful and in some cases can be very itchy. The symptoms will normally improve within a few hours or days, although sometimes they can last a little longer. Most **wasp stings** are not serious, but occasionally they can become infected, cause a severe allergic reaction (anaphylaxis) or spread serious illnesses such as malaria.

5rep: For most people a single bee or **wasp sting** is one too many. Insect bites and stings will usually cause a red, swollen lump to develop on the skin. **Wasp stings** may be painful and in some cases can be very itchy. **Wasp stings** symptoms will normally improve within a few hours or days, although sometimes they can last a little longer. Most **wasp stings** are not serious, but occasionally they can become infected, cause a severe allergic reaction (anaphylaxis). **Wasp stings** can also spread serious illnesses such as malaria.

(15)

1rep: Laboratory studies of **stem cells** enable scientists to learn about the cells' essential properties and what makes them different from specialised cell types. Today, doctors routinely use these cells to treat several conditions affecting the blood cells, such as leukaemia and lymphoma. Researchers think that these cells will be soon used to help create new tissue. Some even believe that they will give us life-changing therapies for multiple sclerosis, type 1 diabetes, and Parkinson's disease, amongst others. The development of treatment methods has evoked great expectations.

3rep: Laboratory studies of **stem cells** enable scientists to learn about the cells' essential properties and what makes them different from specialised cell types. Today, doctors routinely use **stem cells** to treat several conditions affecting the blood cells, such as leukaemia and lymphoma. Researchers think that these cells will be soon used to help create new tissue. Some even believe that they will give us life-changing therapies for multiple sclerosis, type 1 diabetes, and Parkinson's disease, amongst others. The development of **stem cells** therapies has evoked great expectations.

5rep: Laboratory studies of **stem cells** enable scientists to learn about the cells' essential properties and what makes them different from specialised cell types. Today, doctors routinely use **stem cells** to treat several conditions affecting the blood cells, such as leukaemia and lymphoma. Researchers think that **stem cells** will be soon used to help create new tissue. Some even believe that **stem cells** will give us life-changing therapies for multiple sclerosis, type 1 diabetes, and Parkinson's disease, amongst others. The development of **stem cells** therapies has evoked great expectations.

(16)

1rep: The chances of acquiring acute **kidney failure** are greater if you're an older person or if you have any of the following long-term health problems: kidney disease, liver disease, high blood pressure or diabetes. Symptoms of the disease may begin so slowly that you don't notice them right away. You need dialysis when you develop end stage kidney damage, usually by the time you lose about 85 to 90 percent of your kidney function. Dialysis usually makes you feel better because it helps many of the problems caused by this disease. Some people may be able to receive a kidney transplant.

3rep: The chances of acquiring acute **kidney failure** are greater if you're an older person or if you have any of the following long-term health problems: kidney disease, liver disease, high blood pressure or diabetes. Symptoms of the **kidney failure** disease may begin so slowly that you don't notice them right away. You need dialysis when you develop end stage kidney damage, usually by the time you lose about 85 to 90 percent of your kidney function. Dialysis usually makes you feel better because it helps many of the problems caused by **kidney failure**. Some people may be able to receive a kidney transplant.

5rep: The chances of acquiring acute **kidney failure** are greater if you're an older person or if you have any of the following long-term health problems: kidney disease, liver disease, high blood pressure or diabetes. Symptoms of **kidney failure** disease may begin so slowly that you don't notice them right away. You need dialysis when you develop end stage **kidney failure**, usually by the time you lose about 85 to 90 percent of your kidney function. Dialysis usually makes you feel better because it helps many of the problems caused by **kidney failure**. Some people with **kidney failure** may be able to receive a kidney transplant.

(17)

1rep: Common symptoms of a **yeast infection** are a rash, irritation or itching. It is usually harmless, but it can be uncomfortable, depending on where it happens on your body. Occasionally, a mild one may clear up by itself, but most need medical treatment, either from a healthcare provider or at-home care before they resolve. Your healthcare provider will consider your age, overall health, how widespread it is and other factors to determine the best treatment. Treatment may include over-the-counter creams, prescription medications, and home remedies.

3rep: Common symptoms of a **yeast infection** are a rash, irritation or itching. It is usually harmless, but it can be uncomfortable, depending on where it happens on your body. Occasionally, a mild **yeast infection** may clear up by itself, but most need medical treatment, either from a healthcare provider or at-home care before they resolve. Your healthcare provider will consider your age, overall health, how widespread it is and other factors to determine the best treatment for your **yeast infection.** Treatment may include over-the-counter creams, prescription medications, and home remedies.

5rep: Common symptoms of a **yeast infection** are a rash, irritation or itching. A **yeast infection** is usually harmless, but it can be uncomfortable, depending on where it happens on your body. Occasionally, a mild **yeast infection** may clear up by itself, but most need medical treatment, either from a healthcare provider or at-home care before they resolve. Your healthcare provider

will consider your age, overall health, how widespread the **yeast infection** is and other factors to determine the best treatment for your **yeast infection.** Treatment may include over-the-counter creams, prescription medications, and home remedies.

(18)

1rep: The **chief complaint** is an important element of the medical history. It represents the primary problem the patient is presenting with, as initially assessed by a clinician. As a result, it is a key piece of information that guides decision-making on behalf of the patient based on symptom-based information and diagnostics. This information is normally obtained by the doctor in the initial part of the visit when the medical history is being taken. All data is then recorded by the nursing team as structured data from a list of problems.

3rep: The **chief complaint** is an important element of the medical history. It represents the primary problem the patient is presenting with, as initially assessed by a clinician. As a result, the **chief complaint** is a key piece of information that guides decision-making on behalf of the patient based on symptom-based information and diagnostics. The **chief complaint** is normally obtained by the doctor in the initial part of the visit when the medical history is being taken. All data is then recorded by the nursing team as structured data from a list of problems.

5rep: The **chief complaint** is an important element of the medical history. In fact, the patient's **chief complaint** is a key piece of information that guides decision-making on behalf of the patient based on symptom-based information and diagnostics. The **chief complaint** is normally obtained by the physician in the initial part of the visit when the medical history is being taken. All **chief complaints** are then recorded by the nursing team as structured data from a list of problems. Patients can have more than one recorded **chief complaint**.

Screenshot sample as it appeared on the online platform:

Menopause can cause symptoms like anxiety, mood swings, and irregular periods, although hot flushes are considered the hallmark symptom of menopause. Sometimes called 'the change of life', going through the menopause is a natural biological process. Most women experience hot flushes for 6 months to 2 years, although some reports suggest that they can last considerably longer. In the US, up to 75 percent of women report experiencing these transient periods of heat that can occur during the day or night. Women of Afro-Caribbean origin have been shown to have hot flushes which last for more years than those of white women.

3. Reading comprehension

First, we are going to ask you a couple of comprehension questions in relation to the passages you have read. Please, answer the questions sincerely.

1. Please use the scale below to select how easy or difficult you found the texts:

○	○	○	○	○
Very difficult	Somewhat difficult	Neither easy nor difficult	Somewhat easy	Very easy

2. Please use the scale below to select how well you think you understood the texts:

○	○	○	○	○
I struggled to understand	I only understood the main ideas	I understood the main ideas and a few details	I understood the main ideas and most details	I had no trouble understanding
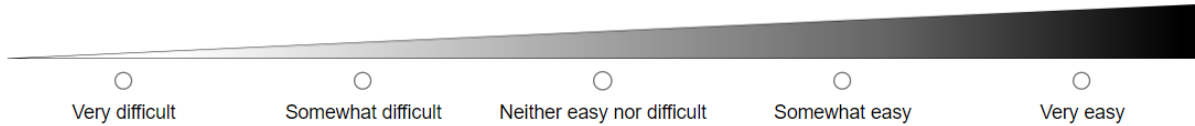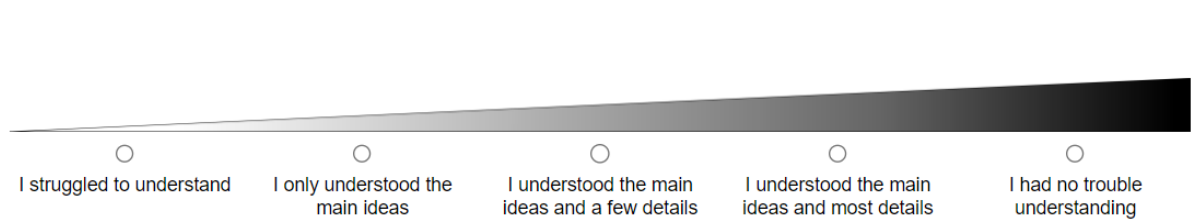
3. Please tick the topics you think you had encountered in the texts.

Only five options are correct.

Topic	
Causes for the increasing cancer cases	
Dental problems and oral health	
Treating infections and bites	
Recent developments in health technology	
Eating disorders	
Good medical practice	
Fitness and exercise	
Looking after mental health	
Organs transplantation and donation	
Symptoms	

4. Form recall

In this first test, you will find a word or a term in Spanish. Write the appropriate translation in English in the spaces provided. You will need two words for each term (one word in each blank). The first letter of each word is shown to help you.

Please <u>do not guess</u>. If you do not know the answer, leave it blank.

1) enfermedad de las encías: g_____ d_____

2) estrías: s_____ m_____

3) médula ósea: b_____ m_____

4) sofocos: h_____ f_____

5) trato al paciente: b_____ m_____

6) alergia al polen: h_____ f_____

7) contracciones: l_____ p_____

8) infarto: h_____ a_____

9) herpes: c_____ s_____

10) piojos: h_____ l_____

11) inmunidad de grupo: h_____ i_____

12) parálisis cerebral: c_____ p_____

13) retortijones: s_____ c_____

14) bulto en el pecho: b_____ l_____

15) sarpullido: p_____ h_____

16) dolor agudo: a_____ p_____

17) enfermedad terminal: t_____ i_____

18) cuerdas vocales: v_____ c_____

19) células madre: s_____ c_____

20) dolor de regla: p_____ p_____

21) coágulo (sanguíneo): b_____ c_____

22) fallo renal: k_____ f_____

23) infección de hongos: y_____ i_____

24) seguro médico: h_____ i_____

25) vena: b_____ v_____

26) reloj biológico: b_____ c_____

27) motivo de la consulta: c_____ c_____

28) médico de cabecera: g_____ p_____

29) picadura de avispa: w_____ s_____

30) salud física: p_____ h_____

31) dientes de leche: m_____ t_____

32) mareos: m_____ s_____

33) esguince de tobillo: s_____ a_____

34) pastillas para dormir (o somníferos): s_____ p_____

5. Form recognition

*Some words occur together in language more common than others. In this test, you will need to select the correct word from five choices in order to form two-word combinations in English. Some options are possible to a greater or lesser degree, but there is only **one** best answer. You can only choose one.*

Please do not guess. *If you do not know the answer, select the "I don't know" option.*

1. gum _____

    a. disorder
    b. soreness
    c. illness
    d. disease
    e. I don't know

2. stretch _____

    a. lines
    b. marks
    c. streaks
    d. stripes
    e. I don't know

3. bone _____

    a. back
    b. marrow
    c. collar
    d. hip
    e. I don't know

4. hot _____

    a. spasms
    b. flushes
    c. blushes
    d. burns
    e. I don't know

5. bedside _____

    a. attitude
    b. care
    c. treatment
    d. manner
    e. I don't know

6. hay _____

    a. allergy
    b. sensitivity
    c. fever
    d. febricity
    e. I don't know

7. labour _____

    a. discomfort
    b. pains
    c. aches
    d. efforts
    e. I don't know

8. sleeping _____

    a. dose
    b. drops
    c. capsules
    d. pills
    e. I don't know

9. cold_____

    a. sore
    b. cut
    c. ulcer
    d. bruise
    e. I don't know

10. head _____

    a. nits
    b. lice
    c. insects
    d. worms
    e. I don't know

11. period_____

    a. discomfort
    b. hurt
    c. pain
    d. ache
    e. I don't know

12. cerebral_____

    a. palsy
    b. paralysis
    c. impairment
    d. disorder
    e. I don't know

13. stomach_____

    a. stitches
    b. cramps
    c. stones
    d. poisoning
    e. I don't know

14. breast_____

    a. nodule
    b. knob
    c. lump
    d. globule
    e. I don't know

15. terminal _____

    a. health
    b. virus
    c. illness
    d. disorder
    e. I don't know

16. acute_____

    a. soreness
    b. ache
    c. pain
    d. hurt
    e. I don't know

17. vaginal_____

    a. discharge
    b. leak
    c. fluid
    d. secretion
    e. I don't know

18. physical_____

    a. health
    b. fitness
    c. status
    d. body
    e. I don't know

19. stem_____

    a. nerves
    b. cells
    c. germs
    d. organisms
    e. I don't know

20. physical _____

    a. health
    b. constitution
    c. complex
    d. healthy
    e. I don't know

21. blood_____

    a. clot
    b. knot
    c. clump
    d. gland
    e. I don't know

22. kidney_____

    a. failure
    b. loss
    c. seeds
    d. rocks
    e. I don't know

23. yeast_____

    a. illness
    b. virus
    c. disease
    d. infection
    e. I don't know

24. health_____

    a. security
    b. insurance
    c. assurance
    d. safety
    e. I don't know

25. blood_____

    a. vase
    b. tube
    c. vessel
    d. tissue
    e. I don't know

26. heart_____

    a. stroke
    b. attack
    c. strike
    d. spasm
    e. I don't know

27. chief_____

    a. consultation
    b. motive
    c. complaint
    d. problem
    e. I don't know

28. general_____

    a. doctor
    b. practitioner
    c. surgeon
    d. medical
    e. I don't know

29. wasp_____

    a. bite
    b. sting
    c. wound
    d. pinch
    e. I don't know

30. rib _____

    a. cage
    b. bone
    c. fold
    d. break
    e. I don't know

31. milk_____

    a. molars
    b. teeth
    c. jaw
    d. formula
    e. I don't know

32. motion_____

    a. nausea
    b. condition
    c. sickness
    d. illness
    e. I don't know

33. sprained_____

    a. bone
    b. ankle
    c. body
    d. tissue
    e. I don't know

34. pain_____

    a. prescriptions
    b. tablets
    c. relievers
    d. killers
    e. I don't know

35. biological_____

    a. clock
    b. time
    c. moment
    d. alarm
    e. I don't know

36. herd_____

    a. impunity
    b. protection
    c. resistance
    d. immunity
    e. I don't know

6. Retrospective protocol

Finally, we are going to ask you a few general questions about your participation in the study. Please, answer the questions <u>sincerely</u>.

Your feedback is really important to us!

1. Did you look up any of the words/terms on a dictionary or the Internet in the course of this study?

    o   yes
    o   no

1b) If you answered yes to the previous question, in which of the three parts did you check the words?

You can select more than one options here

    o   Part 1 (when you were asked to fill out your biographical data, language background, etc.)
    o   Part 2 (when you were asked to read the texts and took some vocabulary tests)
    o   Part 3 (this is Part 3)

1c) Did you look up the words during or after the tests?

You can select more than one options here

    o   during
    o   after

1d) Could you tell us why you looked up the words/terms?

    o   I was curious about the terms I didn't know
    o   I wanted to doublecheck my responses
    o   Other (please specify):_____

2. Did you learn any words or terms that you didn't know before participating in this study?

| I didn't learn any new words or terms | I learned one or two new words or terms | I learned a few new words or terms I didn't know before | I learned a lot of new words or terms I didn't know before |

3. Did you find the topic of the readings engaging?

| Not engaging at all | Not very engaging | Somewhat engaging | Very engaging | Extremely engaging |

4. Do you use any vocabulary strategies when you read in English?

    o   yes
    o   no

4b. [If you answered yes to the previous question] Please tell us which of the following strategies you use when learning new vocabulary:

    a) Guessing from context

    b) Bilingual dictionary use (Spanish-English)

    c) Monolingual dictionary use (only English)

    d) Morphemic analysis, i.e., inferring the meaning of new words by examining their meaningful parts (prefixes, suffixes, etc.)

    e) Others: please specify _____

5. Do you think learning vocabulary is important when learning English?

○	○	○	○	○
Not important at all	Not very important	Somewhat important	Very important	Extremely important
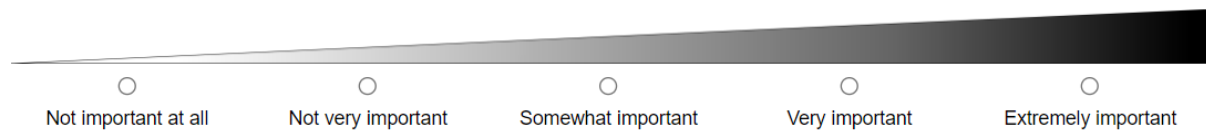
General feedback

Please leave us any comments or your feedback in the box below:

---

7. Contextual support task

The contextual support task was distributed to 5 native English speakers with a TEFL (teaching EFL) background. They were asked to read the passages and rate the level of contextual support on a scale from 1 (contextually unsupportive) to 3 (contextually supportive). The experiment only included texts that received a rating of 3. *This task is a modified version of Webb's (2007a) context specification scale. This task was used to rate the level of contextual support of the stimulus materials.*

You will read some paragraphs that contain expressions in English.

We would like you to tell us how easy or difficult it is to infer the meaning of the expressions from a given text, i.e., if somebody read these passages, would they be able to guess the meaning of the expressions?

Please click "Next" to continue.

_____

This is an example

Please rate (from 1 to 3) how easy or difficult it is to infer the meaning of the expressions from a passage using the scale below, i.e., if somebody read these passages, would they be able to guess the meaning of the expressions?

HOT FLUSHES

Menopause can cause symptoms like anxiety, mood swings, and irregular periods, although **hot flushes** are considered the hallmark symptom of menopause. Sometimes called 'the change of life', going through the menopause is a natural biological process. Most women experience these for 6 months to 2 years, although some reports suggest that they can last considerably longer. In the US, up to 75 percent of women report experiencing these transient sensations of body heat that can occur during the day or night. Women of Afro-Caribbean origin have been shown to have them last for more years than those of white women.

	1	2	3	
It is unlikely that readers can guess the meaning of the expression. The text contains very few or no context cues at all.	○	○	○	It is likely that readers can infer the meaning of the expression. The text contains supportive context cues.

Appendix 2B. Regression models

This appendix provides the final regression models for Experiment 2, described in Chapter 5.

Form recall models

- **Immediate posttest regression model:**

Model:	
Response ~	VocabSize + Repetition + Congruency + Association
	+ Association:Repetition + VocabSize:Repetition
	+ (1+VocabSize|ParticipantID) + (1+VocabSize|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	0.059	0.24	
	VocabSize	0.18	0.43	1.00
ITEM	(Intercept)	0.13	0.36	
	VocabSize	0.0074	0.086	-0.90

Fixed effects	Estimate	SE	z value	Pr (>|z|)
Intercept	-2.79	0.50	-5.60	2.09e-08***
VocabSize	0.75	0.16	4.63	3.72e-06***
Repetition	0.49	0.12	4.23	2.33e-05***
Congruency	0.54	0.21	2.59	0.00**
Association	0.14	0.054	2.52	0.01*
Association:Repetition	-0.022	0.013	1.75	0.08.
VocabSize:Repetition	-0.076	0.040	1.90	0.05.

Correlation of Fixed effects:

	(Intr)	Vocab.	Repetition	Congruency	Assoc	Vocab:Rep
VocabSize	-0.11					
Repetition	-0.75	0.13				
Congruency	-0.22	0.047	0.061			
Association	-0.91	0.039	0.71	-0.030		
Assoc:Rep	0.70	-0.070	-0.94	-0.027	-0.74	
Vocab:Rep	0.099	-0.82	-0.12	-0.033	-0.038	0.067

- **Delayed posttest regression model:**

Model:

Response ~   VocabSize + Repetition + Congruency

            + Association:Repetition + VocabSize:Repetition

            + (1+VocabSize|ParticipantID) + (1+VocabSize|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	8.29e-07	0.0009	
	VocabSize	2.059e-01	0.45	0.06
ITEM	(Intercept)	3.80e-01	0.62	
	VocabSize	3.55e-02	0.19	-0.85

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	-3.65	0.62	-5.99	3.19e-09***
VocabSize	0.63	0.17	3.62	0.000***
Repetition	0.57	0.13	4.85	1.24e-06***
Congruency	0.48	0.29	1.62	0.10
Association	0.18	0.067	2.79	0.005**
Association:Repetition	-0.028	0.013	2.18	0.029*
VocabSize:Repetition	-0.086	0.041	2.10	0.035*

Correlation of Fixed effects:

	(Intr)	Vocab.	Repetition	Congruency	Assoc	Rep:Vocab
VocabSize	-0.14					
Repetition	-0.61	0.097				
Congruency	-0. 29	0.034	0.009			
Association	-0.90	0.025	0.59	0.031		
Repetition	0.56	-0-034	-0.94	0.013	-0.620=	
Rep:Vocab	0.069	-0.80	-0.11	-0.018	-0.016	0.038

Form recognition models

- **Immediate posttest regression model:**

Model:	
Response ~	VocabSize + Repetition + Congruency
	+ (1+VocabSize|ParticipantID) + (1+VocabSize|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	0.37	0.61	
	VocabSize	0.19	0.44	0.77
ITEM	(Intercept)	0.13	0.36	
	VocabSize	0.0016	0.041	1.00

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	0.64	0.37	1.74	0.082.
VocabSize	0.47	0.11	4.43	9.32e-06***
Repetition	0.16	0.042	3.87	0.000***
Congruency	0.18	0.22	0.83	0.41
Association	0.039	0.39	1.014	0.31

Correlation of fixed effects

	(Intr)	VocabSize	Repetition	Congruency
VocabSize	0.12			
Repetition	-0.34	0.026		
Congruency	-0.23	0.022	0.039	
Association	-0.82	0.002	0.020	-0.091

- **Delayed posttest regression model:**

Model:
Response ~   VocabSize + Repetition + Congruency + Association
+ (1+VocabSize+Repetition|ParticipantID)
+ (1+VocabSize+Repetition|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.	
ID	(Intercept)	0.048	0.22		
	VocabSize	0.051	0.23	1.00	
	Repetition	0.0040	0.063	1.00	1.00
ITEM	(Intercept)	0.14	0.37		
	VocabSize	0.013	0.11	1.00	
	Repetition	0.00	0.0078	0.97	0.98

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	0.070	0.36	0.19	0.85
Vocab.Size	0.45	0.080	5.66	1.53e-08***
Repetition	0.17	0.044	3.73	0.000***
Congruency	0.46	0.21	2.19	0.028*
Association	0.072	0.038	1.90	0.057.

Correlation of fixed effects

	(Intr)	VocabSize	Repetition	Association
VocabSize	0.15			
Repetition	-0.36	0.12		
Association	-0.84	-0.004	0.063	
Congruency	-0.24	-0.020	0.037	-0.055

Appendix 3. Experiment 3

_____

Appendix 3A. Instruments

1. Reading task

In this part, we are going to ask you to read a few short passages on different topics:

- The passages are split into different screens.
- Please, read the passages <u>only once.</u>
- When you are done reading one passage, move on to the next one by clicking "Next".
- Please, bear in mind that you cannot go back to a previously read screen.

It takes 20-30 minutes to read all passages.

Take your time and enjoy the reading!

2. Reading materials by level of contextual support (high, mid, low)

Collocations are presented in bold only for the reader of the dissertation. In the actual experiment, the collocations were presented without any numbering or enhancement. H=high informativity, M=mid informativity, L=low informativity.

(1)

H: The goal of this transplant is to cure many serious blood diseases and types of cancer including leukaemia, lymphoma and myeloma. A **bone marrow** transplant takes healthy blood-forming cells from a donor and infuses those cells into the patient's bloodstream, where the cells may begin to grow and produce healthy red blood cells and white blood cells. The preparations for a transplant like this vary depending on the type of transplant, the disease needing transplant, and your tolerance for certain medicines.

M: This year, more than 130,000 Americans will be diagnosed with it. In 2020, 4,864 unrelated and 4,160 related **bone marrow** transplants were performed in the United States. Although millions of people are already included in the global donor registry, it still lacks ethnic and racial diversity and is of limited benefit to patients of non-Northwestern European descent. This makes it harder for some patients to get the treatments they need. The global registry hopes to increase its diversity by recruiting through social media and community outreach.

L: These procedures aren't as rare as you might think. In fact, surgeons perform nearly 30,000 of these procedures each year in the United States. The majority are kidney interventions, followed by **bone marrow**, liver, heart, and lung surgical procedures. However, recent data shows that the most frequent procedures performed in the US are now blood transfusions, respiratory intubation

and ventilation. While blood transfusions are a very common medical procedure, the increased demand of invasive and non-invasive ventilation was triggered by the COVID-19 outbreak.

(2)

H: Sometimes called 'the change of life', going through the menopause is a natural biological process. Menopause can cause symptoms like anxiety, mood swings, brain fog, and irregular periods, although **hot flushes** are considered the hallmark symptom of menopause. Most women experience these for 6 months to 2 years, although some reports suggest that they can last considerably longer. In the US, up to 75 percent of women report experiencing these transient sensations of body heat that can occur during the day or night.

M: Cultural differences can explain variations in symptom reporting and people's attitudes toward aging. A recent study suggests that cognitive factors, particularly beliefs about other people's reactions to their **hot flushes**, might increase distress, causing embarrassment and behavioural avoidance. Even so, the cultural differences highlighted by the survey responses underscore how differences in diet, physical activity, and cultural expectations might influence how people experience symptoms. The social context plays a role too as it might influence gender roles and socioeconomic status.

L: Over two-thirds of the patients who completed the study reported an improvement in their symptoms. Behavioural therapy and exercise appear to be a promising way of reducing **hot flushes**, forgetfulness, depression, sleeping disturbances, among other common problems. Psychological therapy has thus been proposed as a low-risk treatment for a number of symptoms. Nevertheless, other alternative medicine methods have been also suggested to alleviate and manage some troublesome symptoms. Complementary and alternative medical practices suggest administration of vitamin E, soy, black cohosh and red clover.

(3)

H: They appear on your stomach, or sometimes on your upper thighs and breasts, as your pregnancy progresses and your bump starts to grow. When **stretch marks** appear will be different from woman to woman. You'll start to notice them around the end of the second trimester into the beginning of the third trimester, when you're between 6 and 7 months pregnant. That said, they do sometimes appear sooner. Up to 90 percent of all expecting women get these pink, red, brown or sometimes purplish streaks.

M: Anyone can develop them, men and women at any age. Estimates suggest up to 90 percent of people have them. While there is no way to guarantee that a person will not develop **stretch marks**, there are some ways to prevent them. One of the most helpful things you can do is to remain at a healthy weight. Researchers say many remedies such as almond oil, cocoa butter, olive oil and vitamin E do not help. Using products containing Centella asiatica herb and hyaluronic acid may help, but only to a certain extent.

L: Corticosteroids are a type of medication that contains steroids. Prolonged use of corticosteroids is associated with mild side effects, such as the formation of **stretch marks**, whilst long-term use can

result in several severe adverse effects. Among their many functions, corticosteroids regulate the production of certain skin cells. Corticosteroid overproduction can lead to high blood pressure and to those symptoms associated with low levels of potassium, such as weakness, muscle aches, spasms, and disorders of the adrenal glands, including diabetes.

(4)

H: These pitted holes that form in your teeth are among the world's most common health problems. They're especially common in children, teenagers and older adults. But **tooth decay** can also affect infants. To prevent cavities, it is important to follow good oral hygiene by brushing at least twice a day, and flossing at least once a day. Brushing after every meal would be ideal. Also, try to avoid sugary foods and drinks as they will increase the amount of acid in the mouth and increase your chances of getting cavities.

M: Mild illnesses are part of growing up. The good news for parents is that some of the most common child health diseases are preventable with vaccines. Other minor but widespread issues such as **tooth decay** can also be prevented. Parents play an important role in instilling good healthy habits. Good personal hygiene - e.g., washing hands before and after meals, good oral care, bathing - can keep germs and other health-related issues at bay. When parents practice good health habits, children have great examples to follow.

L: Areas with more income deprivation are more likely to have a range of health conditions including mental illness, diabetes, and learning disabilities. According to NHS data, **tooth decay** has the worst rate in the North West of England while the North East has the highest prevalence of obesity. Deprivation gaps also vary in different regions. In the East of England, the obesity gap between the most and least deprived areas is 18%. But in London, it is 78%. Environmental factors and the labour market can also help to explain regional variations.

(5)

H: When the virus is active, medications are used to help reduce and prevent symptoms. But there is no cure for it. In some people, **cold sores** erupt following a trigger event such as a cold. Other triggers include stress, sunlight, feverish illnesses, like influenza or chest infections. These small blisters tend to form on or near the lips and inside the mouth and they usually recur in the same place. If you regularly get them, use antiviral creams as soon as you recognise the early tingling feeling.

M: Symptoms can occur a bit differently in each child. Sometimes symptoms are so mild that parents may not even notice any of them. In the vast majority of children, **cold sores** are annoying and painful but don't cause complications or serious consequences. Once a child is infected with the virus, it is more likely to return during times when the body's immune system is run down. Children with more than five or six outbreaks a year might benefit from antiviral medications.

L: This medicine-chest pain reliever may speed the healing process. A 125-milligram (mg) dose of the pain reliever, taken at the start of an outbreak, was found to cut **cold sore** healing time nearly in half in a study of 42 volunteers at Semmelweis Medical University. These results need to be

verified in subsequent studies, but, in the meantime, over-the-counter pain medications such as ibuprofen can't hurt, and may well help. Other treatments that can help to healing go more quickly must have drying agents, such as alcohol.

(6)

H: The most common trigger for this irritation is exposure to heat for a long time. This may be especially true in very humid areas where the sweat has a harder time evaporating off the skin. **Heat rash** is common in people from cooler climates who travel to warmer climates. Infants and obese people are also more susceptible to it. Any drugs that raise the body temperature or alter the function of the sweat glands can also increase the risk of developing it.

M: Overweight people and people who are physically active may need to take special precautions. Age, general lifestyle, and diet will also make people more susceptible to **heat rash**. Environmental conditions can also increase how susceptible the body is to specific illnesses. For example, exposure to some types of chemicals can cause breathing problems, headaches, and nausea. Certain chemicals in the workplace may cause sterility in men or fertility problems in women. Symptoms of an environmental illness depend on what is causing it.

L: U.S. prescriptions for anti-anxiety medications jumped 10.2% from 8.8 million in March 2019 to 9.7 million in March 2020. A fact many people ignore is that certain medicines may lead to insomnia, **heat rash**, decreased libido and even worsened anxiety. Anxiety medications might not get as much media attention as opioids, but they could be fuelling the next big drug crisis. Today, about 40 million American adults have an anxiety disorder. Given the far-reaching impact of anxiety disorders, medicines used to treat anxiety have grown in popularity.

(7)

H: Although there is still a long way to go, future prospects look very promising. Decades of research have allowed us to glimpse the potential of **stem cells** to treat a wide range of diseases and conditions. It is possible they will give us life-changing therapies for people with spinal cord injuries, multiple sclerosis, and Parkinson's disease, amongst others. With the right stimulation, they can take on the role of any type of cell, and they can regenerate damaged tissue, under the right conditions.

M: In recent years, this industry has been on the rise around the globe, putting patients at great risk of potential harm. One 2020 study counted 280 clinics providing unproven **stem cell** therapies in the United States alone. These clinics appear to offer various treatments for conditions ranging from sports injuries to autism. But a majority of US clinics are problematic because they do not provide enough information on the risk of adverse effects and portray their therapies on their websites in ethically questionable ways.

L: In the last 20 years, the development of these medical advances has evoked great expectations. But there are also alternatives to **stem cell** treatments. Researchers are working on a wide range of innovative medtech from robotics to augmented reality. Across multiple industries, artificial intelligence (AI) has made great waves as a useful technology. For those being treated for the

effects of COVID-19, AI is helping analyse CT scans to detect pneumonia. AI can also be applied to reveal chemical changes in our brain, which can lead to a number of mental symptoms.

(8)

H: Unlike a cold, this condition isn't caused by a virus, is caused by an allergic response to outdoor or indoor allergens. If you're someone who regularly gets **hay fever** at particular times of the year, you might be able to spot your typical symptoms when they start. Having a runny or blocked nose is a very common symptom of this unpleasant allergic reaction. There's currently no cure for this condition, but you can do things to ease your symptoms when the pollen count is high.

M: Some health concerns are more than just a nuisance. One study found that medical costs add up to more than £13 billion per year. NHS England has changed its guidance to say that people with mild-to-moderate **hay fever**, must pay for their medicines instead of getting them on prescription. Prescription medications are expensive. A box of 30 tablets can cost as little as £2, especially if buying non branded products while a prescription for the same medication costs the NHS considerably more.

L: The results of the study found that adults aged 45-54 were most likely to be affected by mental health problems. These statistics also suggest that UK young residents aged between 17-28 are more likely to complain of **hay fever** than of almost any other ongoing condition, including migraines and sleeping problems. The only categories that come close are mental illnesses like anxiety and depression. 1 in 6 people report experiencing a common mental health problem in any given week in England.

(9)

H: Back pain is one of the most common medical conditions, impacting 540 million people around the globe at any one time. Nearly 65 million American residents can report a recent episode of back pain. A **dull ache** in the lower back that never seems to go away remains one of the Western world's most common health problems. This pain is usually used to describe chronic or persistent pain. This is a deep ache felt in an area that typically doesn't stop you from daily activities.

M: The illness has a peak incidence at around 30 years of age, with the highest incidence rates being in the 30 to 34 age group. Important associated symptoms include fever, dysuria, and a **dull ache** in the lower abdomen or groin. Warning signs may also include breast tenderness, growth of breast tissue or significant weight loss. Men are at a higher risk if they do heavy physical work or if they play full-contact sports. Other identified risk factors include race and ethnicity and family history.

L: Concentrates is a broad term referring to all products that have been extracted from the plant. New research has found that concentrates and extracts may help with the short-term improvement of a **dull ache**, anxiety, and other conditions. The problem is that some of these products are manufactured using solvents. Solvent-based products tend to be especially potent, which increase the risks of physical dependence and addiction. When solvents are used to produce concentrates, the preparation process itself can be dangerous.

(10)

H: Physicians can be divided into those who are primary caregivers to their patients, and those who are specialists. Family medicine is the major field for primary care physicians. **General practitioners** are able to treat a wide range of medical conditions and health issues in cases of illness or basic injuries. They are skilled in managing uncertainty, undifferentiated illness and complexity, able to utilise best practice evidence in the light of individual circumstances, and engage patients and families in understanding and managing their health.

M: The growing elderly population means that the medical community must give thought to the management of their care. The most important person to supervise the care of the elderly is the **general practitioner.** He keeps an eye on the health of his elderly patients, knows the score, and can conduct the whole orchestra of welfare on their behalf when they need it. They play an essential role in successful care settings for older people as they can promote and facilitate healthy ageing.

L: The number of UK workers in professional-level employment rose by 647,200 in 2020. The most in-demand jobs include software developers and **general practitioners**. The UK has always attracted highly skilled workers from overseas. Several industries have played a role in attracting these highly qualified professionals, particularly finance and professional/scientific activities. The demands of overseas workers in the UK is growing constantly. Due to Brexit, the country is facing labour shortages in many areas and hence it will rely on UK immigration for highly skilled foreign nationals

(11)

H: Children may get minor cuts and wounds to the face while playing, climbing, or during sports activities. Most of these injuries can be handled at home with simple first-aid treatment. Parents can treat **flesh wounds**, grazes and minor cuts by stopping any bleeding, cleaning the wound, and covering it with a plaster or dressing. Some may not need a bandage. Most of these wounds do not go deep into the skin, they are superficial and usually heal in a few days.

M: Each year, it is estimated that around 2 million children under the age of 15 are taken to accident and emergency (A&E) after being involved in home accidents. Falls and minor injuries relating to objects such as **flesh wounds** and cuts are the top causes of these accidents. While supervision is the best way to prevent injuries, improved home safety is also needed. The best way to reduce the risk of injury in the home is either to remove a potentially dangerous item or add a suitable safety product.

L: Ancient Roman medicine was primarily based on using herbs, foods and diet as therapeutic tools. The Romans made a great progress in understanding anatomy. They were also renowned for their knowledge of **flesh wounds**. Doctors used all kinds of natural substances for their medicine. They extracted juices from crushes herbs, plants and powdered spices for their medicine. The use of herbal medicine was very popular during the Roman Empire. Medicinal plants include cinnamon, mint, and oregano. Some Romans also believed some stones had medicinal qualities.

(12)

H: Around 90% of people who spend more than 3 hours per day using a digital device, including children, will experience some degree of digital **eye strain**. Researchers have found that increased use of digital devices leads to dry eyes because people blink less. Improper distance from the screen and poor lighting can also cause discomfort. Eye and vision-related problems that result from prolonged use of a digital device have risen but experts note that little reliable evidence exists of longer-term damage.

M: Medical issues are a common, even inevitable, part of the aging process. Some can lead to chronic diseases such as diabetes, arthritis or heart failure. The elderly are also prone to cataracts, **eye strain** and glaucoma. While not everyone will experience the same issues, many seniors may also suffer from sleep and behavioural disorders and cognitive deterioration. Mental health problems are also frequent among seniors and may include isolation, affective and anxiety disorders. The most common issues are depression, dementia, and anxiety.

L: The exact causes of different types of headaches are difficult to identify, though there is a clear connection between headaches and stress. Anxiety and worry can also cause stress-induced headaches. **Eye strain** and hormonal changes are also identified triggers for some types of headaches. Because it is usually hard to identify the exact cause of the headache, the symptoms are targeted. Prescription and over-the-counter medications are usually effective for treating headache pain.

(13)

H: In women with heart problems, the excessive pain can cause changes in blood circulation that overwhelms the sick heart. **Labour pain** can also cause psychological trauma that leads to postpartum depression, anxiety disorder and post-traumatic stress disorder (PTSD). These psychological effects of severe pain should not be overlooked. Some women may also have physical complications from a traumatic birth. Unfortunately, symptoms associated with childbirth have been considered normal, and many women have thought they just have to live with them.

M: Pain relief is complex and often challenging without regional analgesia, but there are also a number of non-medical options for dealing with **labour pain**. These include natural techniques that can help you cope with such severe pain without medication. A great coping technique is to have and look at a focal point or deep breathing. Moving around can also help your body lower pain. Alternative treatments include acupuncture, aromatherapy, and reflexology. However, most of these techniques are not proven to provide effective pain relief.

L: Pain is a tricky subject - as it is often considered subjective and dependent on individual pain threshold. But experts agree that certain types of pain are especially excruciating. **Labour pain**, migraine headaches, heart attack and arthritis are among the most painful physical conditions, according to both patients and doctors. Yet, chronic pain is the most debilitating as it also affects a person's ability to perform activities of daily living. It can lead to depression, anxiety and trouble sleeping, which can make your pain worse.

(14)

H: Incorrect posture and computer overuse can cause debilitating physical problems. Using phones and laptops can lead to poor neck and shoulder posture and thus can cause pain and tightness in the shoulders and **shoulder blades**. According to a 2018 research, a forward head posture over stretches the muscles, reduces balance, and leads to chronic neck and shoulder pain. This often comes with headaches if the computer screen isn't set up correctly, wrist pain due to typing and mouse use, as well as pains in the mid-back from poor posture.

M: People are spending most of their time with mobile devices and laptops. Desktops that are poorly set up will cause upper back pain as this will put extra stress on the neck muscles of the **shoulder blade**. Without proper computer set-up and use, there are many injuries that may result. Tendonitis is the most common problem, involving tendon inflammation and localized pain in the forearm, wrist or hand. Bad posture can cause fatigue, muscle strain, and, in later stages, pain.

L: Pain can be much more subtle for women as some symptoms may go unnoticed until a significant blockage causes problems. Women are more likely to experience light-headedness, vomiting and pain the **shoulder blade** region, but not all cases come with clear warning signs. Because more women are likely to brush off their symptoms or delay seeking treatment, they can become their own worst enemies. Women are also less likely to survive than men. This may be because the symptoms differ between the sexes.

(15)

H: This virus can infect your stomach, small intestine, and large intestine. The biggest danger is dehydration from loss of fluid due to diarrhea and vomiting. Children may also have a high temperature. **Stomach bugs** are very common in young children. Children who are vomiting should keep taking small sips of clear fluid, such as water or clear broth. Keep children away from milk. It can make stomach problems worse. Drinks that have a lot of acid (orange juice) can also cause belly problems and discomfort.

M: The UK Health Security Agency publishes surveillance outputs to monitor and predict national trends for England, Scotland and Wales. The latest report shows that a quarter of people in the UK will develop a **stomach bug** whereas adults will experience an average of 2-3 colds in the colder months, between November 2022 and April 2023. If you have cold-like symptoms, it is advisable to take a Covid test as there are some crossover symptoms between the two (headache, coughing, etc.).

L: Hospital admissions for non-COVID-19 disease groups decreased during the UK's COVID lockdown compared with the pre-pandemic figures. Cases of respiratory-related diseases and **stomach bugs** were substantially reduced during the first six months of the pandemic. The lockdown appears to have had a massive impact on the reduction in admissions for non-critical clinical situations. When the lockdown restrictions began to ease between August and September 2020, hospital admission rates were lower as compared to those reported at the same time in pre-pandemic years.

(16)

H: Hospital treatment is free if you are ordinarily resident in the UK. This will include many private and NHS hospitals that provide services to the NHS. In England, **outpatient care** may be provided in a hospital, as well as a walk-in clinic, and even your doctor's office. You may be offered a face-to-face appointment in a hospital or clinic, or a telephone or video appointment. These appointments involve services like consultations, routine physical exams, lab tests such as bloodwork or same-day emergent care.

M: The volume of health care services has climbed for decades, as more Britons need assistance for a variety of concerns. As the population of the United Kingdom ages, the demand for **outpatient care** will continue to rise. Without a considerable increase in the professional workforce, the traditional model cannot provide the capacity required to keep pace with demand. Clinicians are increasingly frustrated with, and fatigued by, growing pressures from waiting lists and overbooked clinics. It is also harder to manage increases in demand for services that are understaffed.

L: Addressing NHS workforce shortages is fundamental to developing a modern, flexible system, based on patient location and needs. Our colleagues will not thank us if we collectively fail to bring **outpatient care** into the 21st century. There is a major staffing crisis across disciplines. The NHS has too few healthcare professionals with at least 105,000 vacancies. Staff are leaving the service due to low pay, work-related stress and reduced job satisfaction, whilst recruitment and retention continue to be a growing problem.

(17)

H: Many people have hair or scalp problems. The common culprit of painful bumps on the scalp is inflammation from eczema or acne which can cause pimples. **Sore spots**, blisters or bumps that develop on the scalp could be the result of anything from dandruff to allergy to skin conditions like psoriasis. An itchy scalp can be upsetting, but it is usually not caused by serious medical problems. It can often be treated at home - there are multiple home remedies, but sometimes requires medical treatment.

M: To spot skin cancers early it helps to know how your skin normally looks. That way, you'll notice any changes more easily. An unusual mole, skin growth or a **sore spot** that doesn't go away may be the first indication of a non-melanoma skin cancer. Other skin cancer symptoms may include: a red- or skin-coloured shiny bump on the top of the skin or a wart-like growth. Skin cancer symptoms can vary depending on the type of skin cancer and its location on the skin.

L: Some patients report certain foods aggravate the symptoms. Studies have also shown that food allergies can be behind these outbreaks. In such cases, a **sore spot** will be visible for a longer period of time when it's not entirely scared. Research has also uncovered that common nutritional deficiencies - lack of B12, iron, and zinc - can increase the likelihood of developing these sores. These nutrients can all be tested for by your doctor to determine exactly what supplements are necessary.

(18)

H: For most people, symptoms usually don't last long. Some people get used to motion on a longer journey, like a cruise, after several days. Although **motion sickness** does not cause long-term problems, it can make you feel miserable and very uncomfortable, and for some people, it can make travelling by car, boat or plane very unpleasant. Symptoms can include mild nausea, dizziness, sweating, an increase in saliva production, pale skin, and vomiting. In addition, some people get headaches, feel very tired, or have shallow breathing.

M: Sickness caused by some chemotherapy drugs is for many people the most difficult side effect to cope with. Clinical observations have found that susceptibility to **motion sickness** is a contributing factor of the type and magnitude of the side effects resulting from cancer chemotherapy. The severity of side effects varies greatly from person to person and will depend on the location of the treatment, chemo drug involved and general health. The most common side effects of chemotherapy include: nausea and vomiting, fatigue, and diarrhea.

L: Our mental state can be crucial in determining our experience of symptoms such as stress, pain, and fatigue, as well as more complex syndromes. For example, people who think they will get **motion sickness** are more likely to get it. Researchers suspect that people who are more positive may be better protected against across a spectrum of conditions. A 2020 study found that people with a family history of heart disease who also had a positive outlook were one-third less likely to have a heart attack than those with a more negative outlook.

(19)

H: People can develop physical or psychological addiction, or both. Psychological addiction occurs when a person believes they need the addictive substance to function. Physical dependence means that **withdrawal symptoms** appear if they stop the addictive substance or behaviour. Symptoms can include insomnia, irritability, changing moods, depression, anxiety, aches and pains, cravings, fatigue, hallucinations and nausea. Severe symptoms, especially for drugs and alcohol, can include paranoia, confusion, tremors and disorientation. Symptoms can last for a few days or weeks, but they will eventually stop.

M: Some people who are chronic, heavy, or long-term users of certain sleep aids (especially controlled medications or heavy sedatives) can receive medications that relieve **withdrawal symptoms**. Sometimes medication is essential. These drugs can be prescribed as part of medical detox or medication-assisted treatment — the latter of which is often combined with counselling. Most studies support the idea that more people stop taking sedatives when they have psychological support. The most successful psychotherapeutic approach appears to be cognitive behavioral therapy (CBT).

L: Research shows that over time, people can recover from mental illness or have remission from their illness or mental disorders. Around 75 % of people who experience severe **withdrawal symptoms** achieve their goal though it may take them some time to achieve full remission. Although recovery rates vary, the average number of attempts before success is five, though the

median number is just two, meaning that some people need five-plus attempts. Although there isn't a cure for mental illnesses, recovery is certainly possible.

(20)

H: It is strong enough to support the skeleton and protect the vital organs in the chest cavity. For this reason, the **rib cage** is an excellent example of the human body's multi-faceted and multi-functional design. This basketlike enclosure has three essential functions: protection, support, and respiration. It surrounds the lungs and the heart, which are two of our most important organs. The small joints between the ribs and the vertebrae permit a gliding motion of the ribs on the vertebrae during breathing and other activities. It is also flexible enough to expand and contract.

M: Various conditions can cause pain (it can be either sharp or dull) and tenderness in your chest. Costochondritis is by far the most common cause of **rib cage** pain. It may result from heavy lifting or hard exercise, or sustained sneezing. Your doctor may order an electrocardiogram to rule out any cardiac issues. Depending on whether you have any other symptoms, your doctor may also order additional tests. If you're diagnosed with costochondritis, though, it usually goes away on its own in a few days to a few weeks.

L: This may not be obvious in early childhood but becomes more noticeable during growth spurts in adolescence. We are not sure what causes **rib cage** problems but there may be a genetic link to it as it seems to run in families. In some children and young people, both sides are affected but in others it is asymmetrical with one side more pronounced than the other. This condition is more common in boys than girls and affects around one in every 1,500 children.

(21)

H: The knee relies on 12 muscles at least to perform its function. It also has the most well-developed cartilage in the body. The knee is a complex joint between the thigh bone and the **shin bone**, which is the weight bearing bone of the leg and it's an important part of your ability to stand and move as it supports lots of important muscles, tendons, nerves and ligaments. What happens after injuring your knee depends on the type of injury, diagnosis and what treatment is possible.

M: Arthritis in the elderly is painful, uncomfortable and can be debilitating. It can cause inflammation, making it difficult to stay active. Arthritis of the knee is common and occurs when the surfaces of the **shin bone** become rough due to wear and tear. Different kinds of arthritis can also affect the hands and wrists. There's no cure for arthritis as yet, but symptoms can usually be managed. Staying active, maintaining a healthy weight and receiving certain treatments might slow progression of the disease and help improve pain.

L: The UK has one of the highest rates of fracture in Europe. Fractures remain a considerable part of the current health and social care expenditure and have a devastating impact on sufferers. **Shin bone** fractures have a massive impact on healthcare budgets. Current estimates suggest that this amounts to £5 million per day, or £2 billion per year; a figure that is only set to rise as the aging population increases. This will create a significant burden on the economy.

3. Form recall

Now, we are going to ask you to complete some tests.

In this first test, you will find a word or a term in Spanish. Write the appropriate translation in English in the spaces provided. You will need two words for each term (one word in each blank).

The first letter of each word is shown to help you.

Please <u>do not guess</u>. If you do not know the answer, leave it blank.

Here is an example:	Here is the answer:
efectos secundarios:	**efectos secundarios:**
s_____	side
e_____	effects

1) llaga: s_____ s_____

2) coágulo: b_____ c_____

3) omóplato: s_____ b_____

4) caja torácica: r_____ c_____

5) médula ósea: b_____ m_____

6) atención médica: h_____ c_____

7) mareo por movimiento: m_____ s_____

8) células madre: s_____ c_____

9) síndrome de abstinencia: w_____ s_____

10) contracciones: l_____ p_____

11) quirófano: o_____ r_____

12) arañazo: f_____ w_____

13) virus estomacal: s_____ b_____

14) fatiga visual: e_____ s_____

15) moqueo: r_____ n_____

16) tibia: s_____ b_____

17) atención ambulatoria: o_____ c_____

18) herpes: c_____ s_____

19) sofocos: h_____ f_____

20) dolor sordo: d_____ a_____

21) sala de urgencias: e_____ r_____

22) médico de cabecera: g_____ p_____

23) alergia al polen: h_____ f_____

24) sarpullido: h_____ r_____

25) caries: t_____ d_____

26) dientes de leche: m_____ t_____

27) reloj biológico: b_____ c_____

28) pastillas para dormir: s_____ p_____

29) pérdida de óido: h_____ l_____

30) anticonceptivo: b_____ c_____

31) dolor de regla: p_____ p_____

32) seguro médico: h_____ i_____

33) medicina oral: o_____ m_____

34) enfermedad del corazón: h_____ d_____

35) salud mental: m_____ h_____

4. Meaning recall

Please translate the following English expressions into Spanish. You can also provide a definition of the term if you know what it means but are unable to recall the equivalent term in Spanish.
Please do not guess. If you do not know the answer, leave it blank.

Write the appropriate translation in English in the spaces provided:

1) heart disease:

2) blood clot:

3) shoulder blade:

4) oral medicine:

5) bone marrow:

6) health care:

7) motion sickness:

8) stem cells:

9) milk teeth:

10) labour pain:

11) operating room:

12) flesh wound:

13) runny nose:

14) eye strain:

15) stomach bug:

16) shin bone:

17) emergency room:

18) cold sore:

19) hot flushes:

20) dull ache:

21) outpatient care:

22) general practitioner:

23) biological clock:

24) hear rash:

25) tooth decay:

26) withdrawal symptoms:

27) hay fever:

28) sleeping pills:

29) hearing loss:

30) birth control:

31) period pain:

32) health insurance:

33) rib cage:

34) sore spot:

35) mental health:

5. Reading comprehension

Now, we are going to ask you about your reading experience of the texts (content comprehension, level of difficulty, etc.) as well as some more general questions about your reading habits.

Please, answer the questions sincerely.

1.   Please use the scale below to select how easy or difficult you found the texts:

○	○	○	○	○
Very difficult	Somewhat difficult	Neither difficult not easy	Somewhat easy	Very easy
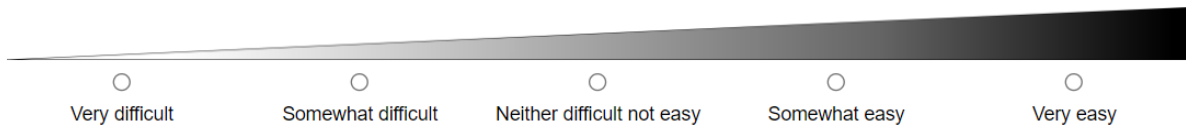
2.   Please use the slider below to select how well you think you understood the texts:

○	○	○	○	○
I struggled to understand	I only understood the main ideas	I understood the main ideas and a few details	I understood the main ideas and most details	I had no trouble understanding

3.   Please tick which of the topics you think you have read before

- euthanasia
- diseases
- medical equipment
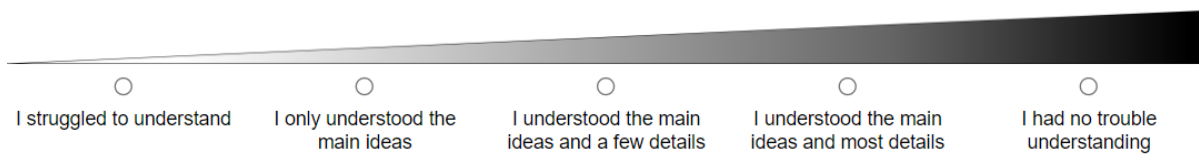- medical conditions
- body parts
- abortion laws

4.   Did you find the topic of the readings engaging?

○	○	○	○	○
Not engaging at all	Not very engaging	Somewhat engaging	Very engaging	Extremely engaging
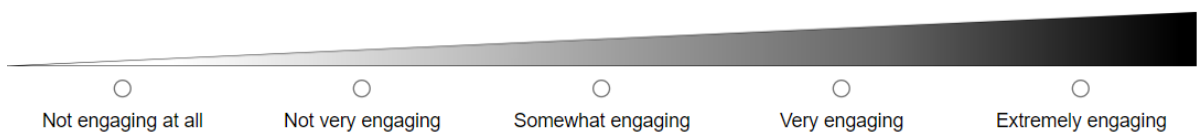
5.   Do you use any vocabulary strategies when you read in English?

a) yes
b) no

5b. If you answered YES to the previous question, please tell us the strategies you normally use when you read in English:

a) Guessing from context
b) Bilingual dictionary use (Spanish-English)   Monolingual dictionary use (English)
c) Morphemic analysis (inferring the meanings of words by examining their meaningful parts: prefixes, suffixes, etc.)
d) Other (please, specify): _____

6. Do you read normally in English?

    a) yes
    b) no

6b. If you answered YES to the previous question, how much time per week do you spend reading in English?

| 1-30 minutes | 31-60 minutes | 1-2 hours | 2-3 hours | 3+ hours |

6c. Do you prefer reading long texts (e.g., a novel) or short texts (e.g., a short story, a piece of news, etc.)?

You can select multiple options

    a) Long texts
    b) Short texts
    c) Other (please specify): _____

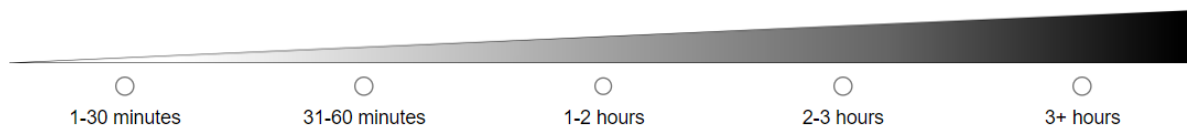Thank you for completing this questionnaire! :-)

6. Distractor task

Some phrases are especially useful for natural conversation. For the questions below, read the text and decide which answer (A, B, C or D) best fits each gap. There is only one correct answer. If you don't know the answer or are unsure, please select option e ("I don't know").

1.- 'What was the concert like?'
-'Pretty .'

a.   good
b.   much
c.   fine
d.   well
e.   I don't know

2.'I'm going to borrow the car even if you say no!'
-' I'll kill you!'

a.   Just you can!
b.   Just you will!
c.   Just you do!
d.   Just you try!
e.   I don't know

3.'What do you think they're doing now?'
- '   .'

a.   How much I know?
b.   How should I know?
c.   How could I tell?
d.   How might I know?
e.   I don't know

4. What's happening about the party next week?'
-'I haven't got   .'

a.   the cloudiest idea.
b.   the foggiest.
c.   a chance to know.
d.   the smallest idea.
e.   I don't know

5. 'Is this the right address?'
- 'Yes,   .'

a)   knowingly
b)   surely
c)   definitely
d)   sincerely
e)   I don't know

6.'My break-up with George seems like the end of the world to me' .
- '   .'

a.   Worse things happen at sea.
b.   There's no use crying over spilt tea.
c.   Never mind.
d.   Bright up!
e.   I don't know

7.'We are going to be late again.!' .
-'I am going as fast as I can.'

a.   Get moving
b.   Get stuffed
c.   Get out
d.   Get real
e.   I don't know

8 .'Sir, it's getting late.'
-'I understand it's been a long day people. But please .'

a.   hold up.
b.   bare with me.
c.   stay along.
d.   listen up.
e.   I don't know

9. 'Do you think they'll have cleaned the house before we get back?'
- '   .'

a.   Some luck.
b.   Some possibility.
c.   Some day.
d.   Some hope.
e.   I don't know

10. -'I don't feel like going out tonight.'
-'Why? '

a.   What's out?
b.   What's for?
c.   What's up?
d.   What's down?
e.   I don't know

7. Retrospective protocol

Finally, we are going to ask you a few general questions about your participation in the study. Please, answer the questions <u>sincerely</u>.

Your feedback is really important to us!

1. Did you look up any of the words/terms on a dictionary or the Internet in the course of this study?

- o   yes
- o   no

1b) If you answered yes to the previous question, in which of the three parts did you check the words?

You can select more than one options here

- o   Part 1 (when you were asked to fill out your biographical data, language background, etc.)
- o   Part 2 (when you were asked to read the texts and took some vocabulary tests)
- o   Part 3 (this is Part 3)

1c) Did you look up the words during or after the tests?
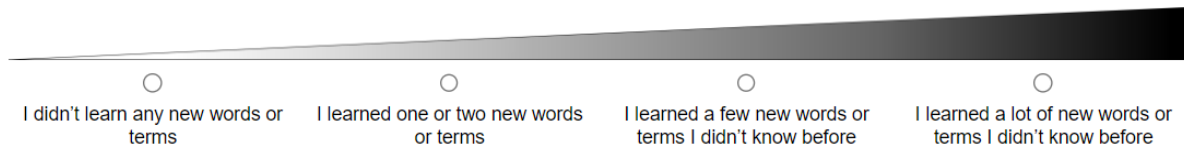
You can select more than one options here

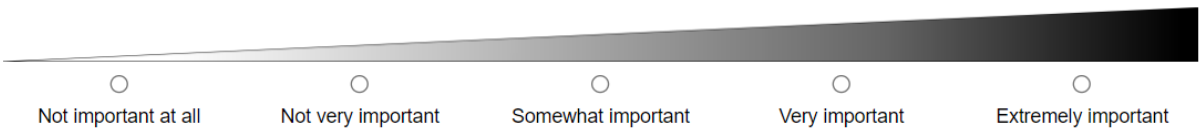- o   during

   o after

1d) Could you tell us why you looked up the words/terms?

   o I was curious about the terms I didn't know
   o I wanted to doublecheck my responses
   o Other (please specify):_____

2. Did you learn any words or terms that you didn't know before participating in this study?

◯	◯	◯	◯
I didn't learn any new words or terms	I learned one or two new words or terms	I learned a few new words or terms I didn't know before	I learned a lot of new words or terms I didn't know before

3. Do you think learning vocabulary is important when learning English?

◯	◯	◯	◯	◯
Not important at all	Not very important	Somewhat important	Very important	Extremely important

General feedback

Please leave us any comments or your feedback in the box below:

---

8. List of Spanish-English cognates

This translation task was completed by native Spanish speakers who did not participate in the main experiment, in order to assess the presence of Spanish-English cognates in the reading materials.

Cognates are words that have a similar spelling, pronunciation, and meaning across two languages. Spanish-English cognates share the same Latin and/or Greek root, are very similar in spelling and have the same or similar meaning. Some examples of Spanish-English cognates are: imparcial and impartial; immortal and inmortal, especial and special, lamentar and lament.

The list below contains Spanish-English cognates. Please translate the English words into Spanish.

abdomen	advisable
acid	ageing
acne	aggravate
acupuncture	alcohol
addiction	allergens
addictive	allergic
admissions	allergies
adolescence	alleviate
adrenal	almond
adverse	analgesia

anatomy

antiviral

aromatherapy

arthritis

asymmetrical

augmented

autism

bandage

biological

cancer

cardiac

cartilage

cataracts

chemotherapy

chronic

circulation

climates

clinic

clinical

clinicians

cocoa

cognitive

corticosteroid

costochondritis

crisis

culprit

debilitating

deficiencies

dehydration

dementia

dependence

deprivation

deprived

descent

desktops

detox

diabetes

diagnosis

diarrhea

donor

eczema

electrocardiogram

episode

fatigue

fertility

fracture

genetic

glands

glaucoma

hallucinations

herb

hormonal

humid

immune

infections

inflammation

insomnia

intestine

intubation

invasive

irritability

leukaemia

libido

ligaments

localized

lymphoma

median

melanoma

menopause

migraine

muscle

nausea

nutrients

obese

olive

opioids

oral

orchestra

oregano

pandemic

paranoia

parkinson

pneumonia

pollen

postpartum

posture

potassium

potent

precautions

prescription

prevalence

progression

psoriasis

psychotherapeutic

racial

reflexology

regenerate

registry

remission

respiratory

robotics

routine

saliva

scans

sclerosis

sedatives

skeleton

solvent

spasms

spectrum

spinal

sterility

steroids

susceptible

symptom

syndromes

tendon

tendonitis

therapeutic

therapy

tolerance

transfusions

transplant

trauma

trimester

vast

ventilation

vertebrae

vertebrate

virus

vital

vitamin

vomiting

zinc

9. Congruency task

Congruency norms were collected from 8 native Spanish speakers who did not participate in the main experiment. They were all English-Spanish translators with university degrees in Translation and at least 5 years of experience as professional translators. These ratings captured the variability in congruency that remained in the test items after highly congruent collocations had been removed.

Some words co-occur in a language more common than others. These are called collocations.

Congruency, which is the presence or absence of a literal first language (L1) translation equivalent, affects the learnability of collocations.

A congruent collocation has the corresponding equivalent in the L1 in terms of the core meanings of the constituent words, in a word-for-word translation. For example, the English collocation "chronic pain" is congruent with its Spanish equivalent "dolor crónico". However, an incongruent collocation has no corresponding equivalent in the L1 in a word-for-word translation. For example, the English collocation "chicken pox" is incongruent with its Spanish equivalent "varicela".
Congruent collocations are therefore easier to learn than incongruent ones.

Please click "Next" to continue.
_____

Instructions:

Please rate each collocation's congruency (overlap between L1 Spanish and L2 English form-meaning connection) on a scale from 1 to 3: 1 being incongruent (no overlap in L1 and L2 form-meaning connection), 2 partially congruent (one word has a high degree of overlap between in L1 and L2 form-meaning connection), and 3 congruent (high degree of L1 and L2 overlap in form-meaning connection).

1. bone marrow

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

2. hot flushes

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

3. stretch marks

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

4. cold sore

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

5. heat rash

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

6. withdrawal symptoms

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

7. hay fever

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

8. dull ache

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

9. general practitioner

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

10. flesh wound

	1	2	3	
incongruent	◯	◯	◯	congruent

How would you translate this expression into Spanish? _____

11. eye strain

	1	2	3	
incongruent	◯	◯	◯	congruent

How would you translate this expression into Spanish? _____

12. labour pain

	1	2	3	
incongruent	◯	◯	◯	congruent

How would you translate this expression into Spanish? _____

13. sore spot

	1	2	3	
incongruent	◯	◯	◯	congruent

How would you translate this expression into Spanish? _____

14. shoulder blade

	1	2	3	
incongruent	◯	◯	◯	congruent

How would you translate this expression into Spanish? _____

15. outpatient care

	1	2	3	
incongruent	◯	◯	◯	congruent

How would you translate this expression into Spanish? _____

16. stomach bug

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

17. tooth decay

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

18. motion sickness

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

19. shin bone

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

20. rib cage

	1	2	3	
incongruent	○	○	○	congruent

How would you translate this expression into Spanish? _____

10. Contextual support task

In order to verify that the reading materials in the three informativity conditions did in fact vary in contextual support according to native speaker judgments, norming data was collected from 24 native speakers of English with backgrounds in (applied) linguistics. Informativity was counterbalanced across three lists distributed to 8 raters each. The rating scale was adapted from Webb (2008) for the present experiment.

You will read some paragraphs that contain expressions in English.

We would like you to tell us how easy or difficult it is to infer the meaning of the expressions from a given text, i.e., if somebody read these passages, would they be able to guess the meaning of the expressions?

Please click "Next" to continue.

_____

This is an example

Please rate (from 1 to 5) how easy or difficult it is to infer the meaning of the expressions from a passage, i.e., if somebody read these passages, would they be able to guess the meaning of the expressions?

SIDE EFFECTS   *

Around 230,000 Australians are admitted to hospital every year because of problems with their medicines, including adverse effects. While most **side effects** can be managed, some can be very serious and may even cause death. Some common examples of mild adverse effects related to drugs include: constipation, skin rash, diarrhea, headache and nausea. Some effects are more likely than others. For example, research shows that a headache occurs in 1 in every 10 patients who use the drug according to the instructions. Heart palpitations may occur in 1 in every 100 patients.

	1	2	3	4	5	
Extremely unlikely that readers can guess the meaning of the expression	○	○	○	○	○	Extremely likely that readers can infer the meaning of the expression

Now you will read the passages.

Please rate (from 1 to 5) how easy or difficult it is to infer the meaning of the expressions from a passage, i.e., if somebody read these passages, would they be able to guess the meaning of the expressions?

All passages are available in the Reading Materials section from page 224.

Appendix 3B. Regression models

This appendix provides the final regression models for Experiment 3, described in Chapter 6.

Form recall models

- **Immediate posttest regression model**

Model:

Response ~ Informativity + VocabSize + Congruency

     + (1+Congruency|ParticipantID) + (1+Congruency|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	-2.57	0.46	
	VocabSize	0.029	0.145	0.58
	Congruency	0.22	0.47	-0.98 -0.40
ITEM	(Intercept)	2.61	1.61	
	Congruency	0.94	0.97	0.88

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	-2.57	0.79	-3.26	0.001**
Inf.Helmert1	0.39	0.15	2.70	0.007**
Inf.Helmert2	0.18	0.12	1.45	0.15
Vocabulary	0.22	0.071	3.034	0.002**
Congruency	0.92	0.56	1.65	0.098

Correlation of fixed effects

	(Intr)	VocabSize	Inf.Helmert1	Inf.Helmert2
VocabSize	-0.001			
Inf.Helmert1	-0.015	0.020		
Inf.Helmert	-0.008	0.051	0.053	
Congruency	-0.95	-0.002	0.008	0.004

- **Delayed posttest regression model**

Model:

Response ~ Informativity + VocabSize + Congruency
+ (1+Congruency|ParticipantID) + (1+Congruency|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	0.24	0.49	
	VocabSize	0.022	0.15	0.75
	Congruency	0.102	0.32	-0.70 -0.04
ITEM	(Intercept)	0.89	0.94	
	Congruency	0.12	0.35	1.00

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	-2.71	1.20	-2.26	0.023*
Inf.Helmert1	0.34	0.15	2.23	0.025*
Inf.Helmert2	0.10	0.13	0.82	0.41
Vocabulary	0.41	0.082	5.009	5.46-07***
Congruency	1.20	0.86	1.40	0.16

Correlation of fixed effects

	(Intr)	VocabSize	Inf.Helmert1	Inf.Helmert2
VocabSize	-0.002			
Inf.Helmert1	-0.011	0.003		
Inf.Helmert	-0.004	0.024	0.024	
Congruency	-0.96	0.008	0.009	0.003

Meaning recall models

- **Immediate posttest regression model**

Model:

Response ~ Informativity + VocabSize + Congruency

+(1+VocabSize+Congruency|ParticipantID)+(1+VocabSize+Congruency |ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	2.7808	1.6676	
	VocabSize	0.0005	0.0231	1.00
	Congruency	1.1417	1.0685	-1.00 -1.00
ITEM	(Intercept)	2.2362	1.4954	
	VocabSize	0.0337	0.1836	0.23
	Congruency	0.6114	0.7819	-0.94 0.13

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	0.002	0.58	0.005	0.99
Inf.Helmert1	0.96	0.13	7.14	9.49e-13***
Inf.Helmert2	0.14	0.12	1.19	0.23
Vocabulary	0.16	0.071	2.26	0.023*
Congruency	0.48	0.38	1.27	0.20

Correlation of fixed effects

	(Intr)	VocabSize	Inf.Helmert1	Inf.Helmert2
VocabSize	0.017			
Inf.Helmert1	0.018	0.013		
Inf.Helmert	0.004	-0.004	-0.044	
Congruency	-0.959	0.109	-0.013	-0.007

- **Delayed posttest regression model**

Model:

Response ~ InformativityHelmert1 + InformativityHelmert2 + VocabSize + Congruency
+ (1+Congruency|ParticipantID) + (1|ItemID)

Random effects:

Groups	Name	Variance	Std. Dev.	Corr.
ID	(Intercept)	2.016	1.42	
	Congruency	0.93	0.96	-1.00
ITEM	(Intercept)	0.77	0.88	

Fixed effects:	Estimate	SE	z value	Pr (>|z|)
Intercept	-0.40	0.64	-0.63	0.53
Inf.Helmert1	1.09	0.13	8.094	5.75e-16***
Inf.Helmert2	0.002	0.11	0.021	0.98
Vocabulary	0.26	0.055	4.75	2.28e-06***
Congruency	0.50	0.42	1.18	0.24

Correlation of fixed effects

	(Intr)	VocabSize	Inf.Helmert1	Inf.Helmert2
VocabSize	-0.009			
Inf.Helmert1	-0.002	0.048		
Inf.Helmert	-0.003	-0.003	-0.015	
Congruency	-0.95	0.013	0.006	-0.002