



Grado de Estudios Ingleses

**Trabajo de Final de Grado**

**Curso 2015-2019**

**TÍTULO: FACTORS INVOLVED IN L2 VOCABULARY  
LEARNING THROUGH READING**

**NOMBRE DEL ESTUDIANTE: DAVID RUIZ BARRAYCOA**

**NOMBRE DE LA TUTORA: IMMA MIRALPEIX**





### Declaració d'autoria

Amb aquest escrit declaro que sóc l'autor/autora original d'aquest treball i que no he emprat per a la seva elaboració cap altra font, incloses fonts d'Internet i altres mitjans electrònics, a part de les indicades. En el treball he assenyalat com a tals totes les citacions, literals o de contingut, que procedeixen d'altres obres. Tinc coneixement que d'altra manera, i segons el que s'indica a l'article 18, del capítol 5 de les Normes reguladores de l'avaluació i de la qualificació dels aprenentatges de la UB, l'avaluació comporta la qualificació de "Suspens".

Barcelona, a 11 de Juny de 2019

Signatura:

## **Acknowledgments**

I would like to gratefully mention the people who encouraged and supported me, since the beginning and through the development of this project, and without whose help this project would not have been possible. First of all, I am extremely grateful to my tutor, Imma Miralpeix, for giving shape to the vague thoughts that this project was at its genesis, and also for her continuous support and work when it came to the challenges that it presented us with. I am also grateful to the teachers in the degree of English Studies of the University of Barcelona, specially Isabel Seguro, Cristina Alsina, Marta Tirado, Gemma Lopez, Enric Monforte, Bill Phillips, Elisabet Comelles, Isabel Verdaguer, Elsa Tragant and Emilia Castaño for their dynamic and constructive lessons that, not only were productive in terms of content, but also helped me grow as a reader and writer. I would also like to thank my family and specially my parents, Fernando Ruiz and Esther Barrycoa, for always believing in my potential and encouraging me to keep moving forward when I was feeling down or unmotivated. Finally, I want to thank my friends Telmo Zanguitu, Markel Lazaro, Eneko Orueta, Aian Cantabrana, Alex Miner and Iñaki Garrido, and my friends at university, Maripau Gonzalez, Oriol Parcet, Ane Barricarte, Eloy Villalón, Helena Benacloche and Eddy Franco for their helpful suggestions and encouragement.

## **Agradecimientos**

Me gustaría expresar mi profundo agradecimiento a la gente que me ha ayudado y animado desde el principio y a lo largo de este proyecto y que, sin cuya ayuda, no me habría sido posible llevar a cabo. Para empezar, estoy enormemente agradecido a mi tutora, Imma Miralpeix, por dar forma a la confusa idea que este proyecto constituía en su génesis, y también por su continuo apoyo y arduo trabajo a través de los desafíos que iban apareciendo. También estoy agradecido a los profesores del grado de Estudios Ingleses de la Universidad de Barcelona, especialmente a Isabel Seguro, Cristina Alsina, Marta Tirado, Gemma Lopez, Enric Monforte, Bill Phillips, Elisabet Comelles, Isabel Verdager, Elsa Tragant y Emilia Castaño por sus clases dinámicas y constructivas, que además de ser productivas, también me han ayudado a crecer como lector y escritor. También me gustaría dar las gracias a mi familia y especialmente a mis padres, Fernando Ruiz y Esther Barraycoa, por siempre apostar por mi potencial y animarme a seguir cuando me sentía desanimado o desmotivado. Para acabar, quiero dar las gracias a mis amigos Telmo Zanguitu, Markel Lazaro, Eneko Orueta, Aian Cantabrana, Alex Miner and Iñaki Garrido, y a mis amigos y amigas de la Universidad Maripau Gonzalez, Oriol Parcet, Ane Barricarte, Eloy Villalón, Helena Benacloche y Eddy Franco por sus sugerencias constructivas y apoyo continuo.

## **Abstract**

This work focuses on the factors influencing the learning of new vocabulary when reading in a second language. Apart from identifying these factors, the project summarizes relevant findings in the literature in relation to each of them, which can be related to the reader, the text or to both. The present study first introduces the difference between intensive and extensive reading. Then, it deals with aspects related to the reader, such as age, vocabulary size and proficiency level; to the text, such as amount of the type of words present in the text or number of different encounters with new vocabulary; and to both, such as coverage, which is the amount of words learners understand from a particular text. The present study also takes into account how linguistic and extralinguistic support can enhance L2 vocabulary learning when reading (e.g. through glossing or visual representations).

## **Key words:**

L2 reading, vocabulary learning, vocabulary size, coverage, graded readers, word encounters, glossing.

## **Resumen**

Este trabajo se centra en los factores que influyen en el aprendizaje de nuevo vocabulario a partir de la lectura de textos en una lengua no materna. Además de identificar estos factores, el proyecto resume descubrimientos relevantes en la literatura relacionados con cada uno de ellos, que a su vez se pueden relacionar con el lector, el texto o ambos. El estudio presente introduce primero la diferencia entre la lectura intensiva y lectura extensiva de textos. A continuación, trata los aspectos relacionados con el lector, como lo son la edad, el tamaño de vocabulario y el nivel de competencia; el texto, como lo son la cantidad de tipos de palabras en un texto o el número de diferentes encuentros con vocabulario desconocido; y ambos, como lo es la cobertura de textos, la cual se refiere al número de palabras de un texto en particular que los lectores son capaces de comprender. El estudio presente también tiene en cuenta como el soporte lingüístico y extralingüístico pueden ayudar a mejorar la experiencia de aprendizaje de nuevo vocabulario a través de la lectura (p. ej. A través glosarios o representaciones visuales).

## **Palabras clave:**

Lectura en una lengua no materna, aprendizaje de vocabulario, tamaño de vocabulario, cobertura de textos, lecturas adaptadas, encuentros de la misma palabra, glosarios.

## Table of contents

Introduction.....	1
Justification.....	3
1. Main types of reading.....	4
1.1 Extensive reading and incidental learning.....	4
1.2 Intensive reading and explicit learning.....	5
2. Factors that influence vocabulary learning when reading.....	7
2.1 Factors related to the learner.....	7
2.1.1 Age.....	8
2.1.2 Vocabulary size.....	9
2.1.3 Proficiency level.....	10
2.2 Factors related to the text.....	10
2.2.1 Words' features and word families.....	11
2.2.2 Word frequency in the language.....	12
2.2.3 Word encounters in the text.....	12
2.2.4 Context and register.....	13
2.3 Factors related to both the learner and the text.....	14
2.3.1 Coverage.....	14
3. Linguistic and extralinguistic support when reading.....	15
3.1 Simplified and unsimplified texts.....	15
3.2 Glosses.....	16
3.3 Multimodal input.....	17
4. Conclusions.....	18
5. My own experience on learning L2 vocabulary from reading.....	20
6. Works cited.....	21



## **Introduction**

Learning a new language is by no means an easy challenge. Languages are very complex tools of communication composed of several features (morphology, phonetics and phonology, syntax, semantics and pragmatics) that are difficult to master, so learning them is a very time-consuming activity. There are individual factors, such as motivation, learner proficiency or relation between the L1 and the L2, among others, that may help acquisition, but the sheer volume of what needs to be learned, for a successful understanding and use of a second language (L2) still remains a hefty difficulty.

In fact, when considering the relations that languages may have, it is important to observe that, even if languages share similarities due to them having a common root, such as in the case of Romance or Germanic languages, the languages that we use nowadays are the by-products of many historical changes. When confronting an L2, vocabulary is considered to be crucial for comprehension. There are also other elements particular to each language, such as grammar and syntax, that need to be properly learned in order to attain an adequate understanding of a language. However, I would rank vocabulary a high priority since I consider it the core element that defines a language's identity, indispensable for both understanding and using of an L2. One may place words in the wrong order, fail to pronounce them perfectly or use them without marking them with the proper grammatical morphemes and still be able to communicate, but if they do not know the words to express an idea, communication can easily break down (Lightbown & Spada 2013). Therefore, vocabulary is a potentially significant element and an ambitious learning goal when learning a L2. It is ambitious, I believe, since learners cannot hope to learn all the words of the L2, not even native speakers master all words in their own languages (Nation, 2006). Thus, vocabulary learning will depend as well on the learner's ability to make the most from activities such as listening or reading a novel or a newspaper, watching a movie or having a conversation (Nation 2006).

Vocabulary learning can take place either implicitly, when they pick up words without their attention being directly put on them by the teacher, or explicitly, when they are asked to make a conscious effort to learn. Likewise, there is more than one kind of activity meant to perform such learning, but one that has been subject to thorough research, and which will be the focus of this work, is the learning of vocabulary through reading. Reading offers a continuous exposure to words in context and can become a good source of vocabulary encounters with new words that may promote learning (Nation,

2001). That said, there are different ways of approaching reading, which is an activity that can be performed either inside the classroom, like intensive reading, or outside the classroom, for example with extensive reading or reading for pleasure, and that have different effects on learning. In the case of extensive reading, Nation (2001) points out that, if conditions like vocabulary size, coverage or repetitions are favourable, a considerable amount of vocabulary learning can happen through reading. However, there is a lot of research on to what degree different conditions, such as types of words encountered, memory, or the adequacy of support material, such as glosses in graded readers can help the learning of new words.

The present work does not attempt to make any new discoveries on the topic of L2 vocabulary learning through reading, nor does it try to come up with any new hypothesis. Its aim is to put together what many researchers and studies have found on the topic and organize it in a way that clearly depicts the factors related to the different actors of the process: learners and texts (or both). It also considers the support materials that may be used to facilitate the reading process (such as glosses, graded readers or readings complemented with other forms of input such as images). Lastly, as a closing section, we will reflect on the findings of the studies presented here in relation to the author's own learning experience when learning English as a L2 through reading.

## **Justification**

The reason I decided to put together this work lies in my learning experience of English. Just as the average educated person in Spain, I received a somewhat poor education in English during my school years. By attending year-long courses in language academies for 4 years, my level, regarding grammatical, syntactic and lexical knowledge, improved quite significantly. However, even if I knew the knowledge was there, it still proved difficult to access it in a fluent way, so I would often find myself struggling for words or structures when writing or speaking in English. One day I became interested in reading some novels that were of Japanese origin but that had only been translated to English, so I had no other choice but to read them in English. At first, the amount of unknown words, structures and even pragmatic or metaphoric use of words that I knew made it hard to progress through the novels. But since I was interested and invested time on them, I struggled and used online dictionaries and other resources. The unknown words and structures kept appearing and repeating themselves the more I read, so at some point I began to learn and remember them, I would not need the help of external resources to understand them. The fact that I had some context to work with also helped a great deal. But the real surprise came when, at the time of writing or speaking, words and structures would come to my mind with much less effort and more fluently than before. Not only had I learned new vocabulary and grammar, but what I already knew before I started reading seemed to be more available for me when producing, as I had seen countless examples in the books. Thus, I became very interested in knowing how we learn L2 vocabulary when reading and wanted to explore factors that can facilitate acquisition while doing this activity. Reading turned out to be such a turning point in my learning experience that I wanted to make a state-of-the-art project, one that put together the findings of several studies and depicted clearly the factors involved in the process.

## **1. Main types of reading**

According to the Cambridge Dictionary, the term 'reading', by itself, means the act of reading any kind of text and of any kind of length, so it can be too broad for research and explanatory purposes and warrants a specific division into sub-terms, which can account for the different features or purposes that texts can present. Although there may be no end to the means followed to make such categorizations, like according to genre or type of text, or length, when it comes to language learning through reading, the sub-terms 'extensive reading' and 'intensive reading' show a rather widespread use in academia. There seems to be a consensus regarding their use to refer to types of reading for language learning. Both extensive and intensive have their own purposes and offer different kinds of benefits, so it is worth considering what research has to say about them and the advantages and disadvantages they may present.

### **1.1. Extensive reading and incidental learning**

Nation (2001) defines extensive reading as the reading of considerable amounts of texts which are chosen by the reader and which are not difficult to read due to having few unknown words (less than 5% of the running words of the text). Its most significant feature is that readers regard it as a pleasurable activity they would spend their free time on, and not as language learning activity they need to actively engage on. Among its advantages, Nation considers the availability of plenty of material for different proficiency levels, the wide variety of texts that can appeal to readers and motivate them and, most importantly, the chance of having contact with the L2 outside the classroom environment.

It is under the favourable conditions mentioned thus far that incidental vocabulary learning is expected to happen, but exhaustive research has shown that the equation is not that simple. For instance, while an average percentage of less than 2% of unknown words in a text can help guessing and inferring the meaning of new words (Nation, 2001), that does not result in immediate word acquisition and could even lead to learners avoiding unknown words, if they consider that the context suffices for the understanding of the text (Pigada and Schmitt, 2006). In addition, although there are texts for many proficiency levels, even the lowest proficiency level text would require a minimum vocabulary knowledge for any learning to take place. Horst et al. (1998) carried out a study with L2 learners to see the extent to which extensive reading was helpful for vocabulary learning, and they found that low-intermediate or early learners learned very few words and that

they would benefit more from direct instruction in their early stages of learning. This has been a recurrent finding. That said, the study does agree with Nation (2001) in the fact that extensive reading can also account for enhancing already known frequent vocabulary if it was pre-taught explicitly.

On a similar note, regarding other benefits that extensive reading can provide, the study by Pigada and Schmitt (2006) showed that aspects of word knowledge other than meaning could be learned by extensive reading, such as spelling and grammar. This can be related to the way in which Nation (2001) considers extensive reading as an activity that has also a positive effect on other areas of language use, such as writing. As a final remark, the factors we have mentioned seem to point towards the fact that extensive reading does not provide immediate results on language learning in general and they require time and repeated encounters with vocabulary to happen, which is in line with Nation's (2001: 258) claim that "the benefits of extensive reading do not come in the short term, and the substantial long-term benefits justify the high degree of commitment needed."

### 1.2. Intensive Reading and explicit learning

Different from its counterpart 'extensive reading', intensive reading consists in the reading of smaller quantities of text that allow for the focus on certain language forms or reading aspects, which can be worked on with ease as it is possible to read the text multiple times. The clearest example for intensive reading would be the reading exercises that you can find in an English textbook, which tend to be located in different units to focus on specific language aspects as well as to practice reading-comprehension. The length of this kind of texts usually ranges from 300 to 500 words and, although they are usually made to focus on set language or reading features, they also offer a good chance to notice and learn vocabulary and grammar (Nation, 2001).

At first sight, the main differences that set intensive reading apart from extensive reading seem to be the length of the text, that seems to be shorter, and the focus on language learning rather than reading for pleasure. This then means that readers are exposed to less input and, by extension, less repetitions of words and grammatical structures. It could be argued that the possibility of re-reading the text multiple times can make up for that lack, but since the focus is on learning, the burden on the reader is heavier than if s/he was reading for pleasure. Therefore, this limits significantly the amount of

times they may re-read it before getting tired or bored of that particular text. However, intensive reading also offers advantages that extensive reading cannot account for, as it is the case with the direct teaching of high frequency words in context. It is worth devoting time to high frequency words, because they are a limited set of words that can enable readers to access more texts. By working on them in class, awareness of certain words can be raised so that readers are more conscious of them and notice them while reading, which would lead better comprehension (Nation, 2001). In addition, as Nation (2001) further develops, unknown words in intensive reading can also be used to practice strategies that may come in handy for readers when they encounter unknown words in longer readings, such as guessing from context, recognizing word parts or relating the unknown word to other known words. Regarding the relation between words, a study by Nation and Ming-Tzu (2004) mentions that intensive reading offers a good chance to look into polysemic words so that readers can see several senses and uses of a word, learn how their meanings are related, and make better connections when encountering that word in the future.

So far we have talked about extensive and intensive reading and their advantages and shortcomings, but it is important to clarify that neither of them is better than the other, they are not mutually exclusive, since they can perform different functions and, when you study them closely, you can notice that they actually cover for each other's lacks. For instance, while intensive reading proves useful to learn high frequency vocabulary, the same cannot be applied to mid and low frequency vocabulary, as we are talking about a way too ample range of words that is not worth teaching. However, in extensive reading, meetings with unknown mid-low frequency words can happen more often, as texts are longer, and, by using the aforementioned strategies learned through intensive reading, readers can become acquainted with them and learn them over time. Nation (2001) also offers a clear example of this symbiotic relationship, referring to unsimplified texts and how they are hard to tackle for L2 readers in extensive reading due to high vocabulary burden. Nevertheless, it can become a good exercise in intensive reading, as more time can be devoted on the parts they struggle with.

Now that we have made clear what extensive and intensive reading are and what they imply, the next part will focus on the main topic of this work, that is to say, the main factors that take part in the process of learning new vocabulary through reading.

## **2. Factors that influence vocabulary learning when reading**

So far we have looked into two of the main ways of approaching reading and how they contribute to learning according to academia, and now, bearing this in mind, we can concentrate on the main topic, which is the factors intervening in the learning of new vocabulary when reading. Again, this topic has already been object to many studies and research has shed light on the factors that intervene in the process. Among them we can find the role of coverage for text comprehension and inference from context (Nation, 2001; Nation, 2006), the significance of word repetitions (Nation, 2001; Waring & Takaki, 2003; Webb, 2007) in how words are learned and retained (Rott, 2007), how syntactic complexity can affect incidental learning (Bordag et. al, 2015), or the role of age and proficiency level (Weber-Fox & Neville, 1996; Horst et al, 1998; Meschyan & Hernandez, 2006), among others. By paying close attention to them, one can discern that these factors, which influence vocabulary learning when reading, can be related to either the reader, such as age and proficiency level, the text, such as word repetitions or syntactic complexity, or both the text and the reader, as is the case of coverage. Therefore, it should be possible to classify the aforementioned main factors under these three groups, ‘factors related to the reader’, ‘factors related to the text’ and ‘factors related to both the reader and the text’, to make for a clear and organized representation of the process of learning new words through reading. It should be added, however, that most of them refer to extensive reading more than intensive reading.

### **2.1. Factors related to the learner**

When it comes to vocabulary learning, or any kind of learning, it is important to keep in mind that not all learners take on the task in front of them under the same conditions. Different factors and experiences acquired before, or the lack of them, may place them in a position where the task can be taken on with more ease or difficulty. Needless to say that ‘factors’ and ‘experiences’ are way too broad terms that could account for an unlimited number of possibilities, so it is almost a given fact that we cannot consider each and every one of them for every single person, although that would be the ideal approach to come close to the most complete conclusion on the way new words are learned. However, that is not to say that all of them are particular to each certain individual, there are some that are common to learners regardless of origin and growth or development. In the case of learning new words through reading, when thinking about the factors that would condition the approach and engagement of the reader with the text, what first

comes to mind are factors related to the comprehension of the text or the speed at which the text is read. That is because of the assumption that, if a text takes too much time or effort to read, then it can become a burden for the reader and the learning of new words will probably diminish. With this in mind, we could say that the age, proficiency level and vocabulary size of the reader are important variables that can affect comprehension or reading speed and, therefore, the learning of new vocabulary.

### 2.1.1. Age

For the purpose of this work, we can consider age in at least 2 different ways. The first one would refer to the current real chronological age of the learner when approaching a certain type of text. Regardless of the proficiency level of the reader on the language, some texts and genres, such as philosophical texts or may be harder to read and understand because they demand a high level of thinking and reflecting upon the reader. This would translate in younger readers, with a good grasp of the language, be it their L1 or L2, having trouble with such complex texts since they lack the mental maturity to understand them, which on average is gradually acquired by people when they grow and are exposed to many experiences they can then relate to.

Chandler (1997) analyzes research by Buckingham (1993) on how children understand television genres and mentions how his findings point toward children developing a discourse of genre, as they mature, and the way that helps them identify details and make richer connections between programs. Following this idea, he then explains that genres are not just characteristics that define or classify texts, but, more importantly, a way for texts to mediate with reader. Therefore, they influence how the text is read and the reader's interpretation of it, enhancing the comprehension of it if they are familiar with the genre. But, as mentioned before, the knowledge of genres is developed over time, as readers mature in age and become familiar with them, so early or young readers are, conversely, hindered by their lack of this knowledge when approaching certain texts.

The second way to consider age would refer to the time when the exposure to the language began (this is often referred to in the literature as Age of Onset -AoO-). Depending on the age in which the reader started to be exposed to the language, more or less time of exposure and learning will have taken place by the time they engage with a text, and that is likely to have an effect on language processing when reading. Although



there is a plethora of factors other than age that can affect language processing, the most logical conclusion one can come up with would be that, the earlier the exposure to the language, the higher the chances for a better language processing are. However, Weber-Fox and Neville (1996) came across some discrepancies about that, as not all language features seemed to follow the same developmental pattern. In their study they found that, while age affected syntactic processing, causing delays starting from L2 exposure at the 1-3 years of age, semantic processing did not yield the same results, showing that only those who started to be exposed to the L2 at the 11-13 or after 16 years of age presented slower processing, although not that significant. This seems to confirm that the age at which one started to be exposed to a L2 can influence, to a certain degree, the reading and comprehension of texts, at least regarding their syntactic complexity. Even so, considering that the delay in syntactic processing is not that significant, we could say that AoO is just one of the factors that can affect readers when approaching a certain text.

It should be noted though, that age is often confounded with proficiency: in terms of real chronological age, young learners tend to be less proficient and sometimes they have still reading problems in the L1. Older learners tend to be more proficient and have better reading skills. In addition, an early AoO is often related to better proficiency, although it might not be the case if learners have not received much exposure (Muñoz, 2006). However, more proficient learners are more readily able to make the most from the text when compared to less proficient learners. In relation to this, one of the variables that has been shown to be related to proficiency in the L2 is vocabulary size; i.e. the amount of words that learners know at a particular stage of their learning (Miralpeix and Muñoz, 2018).

### 2.1.2. Vocabulary Size

One of the essential factors to access a wide variety of L2 texts, and also essential for overall language use, is the vocabulary size of the readers. It is considered a measure with considerable weight since it can severely condition language comprehension or use, depending on whether the amount of vocabulary is high or low. Nation (2001) elaborates on the possible existence of a language knowledge threshold that would tell whether adequate comprehension can take place or not. This implies that readers who have not crossed the threshold have lower chances, or cannot, attain an adequate comprehension when reading a text, whereas those who have crossed the threshold have better chances to attain it. For instance, readers with a low vocabulary size might find themselves

struggling with the core parts of a text that are necessary for its understanding, even if the text is not a complex one. Although small vocabularies may imply learning more words, the burden of not being able to have an overall picture of the text would probably jeopardize any significant progress in the learning of new words. Conversely, readers with a sufficient vocabulary size are probably more liable to learn new words from texts, as they do not need to place as many cognitive resources on trying to grasp the overall meaning of the text (Horst et. al, 1998). However, in the results of their study, Horst et. al (1998) found that the relationship between vocabulary size and incidental learning of new words, although existent, was not highly significant. All things considered, vocabulary size appears to be an important factor regarding text comprehension, and one worth considering in the process of learning of new vocabulary. Learners with small vocabulary sizes will have difficulties to learn new words just from reading and may need explicit teaching to cope with the vocabulary demands.

### 2.1.3. Proficiency level

Proficiency level in the L2 of the readers is expected to play a significant role in the reading and comprehension of texts. It is also related to the previous factor, vocabulary size, as large amounts of vocabulary are usually attained over time, and that would mean that other aspects of language might have also developed along that time. In this respect, higher proficiency level would lead to easier recognition of words and structures and, therefore, it would facilitate reading and comprehension, whereas low proficiency level would likely result in slower reading (it would take more time to identify structures) and poorer comprehension. Meschyan and Hernandez (2006) conducted research on the effect of language proficiency on neural activity and found that, when the proficiency in the language is low, the toll on the articulatory motor processing is greater and so, it takes more time to read or produce in that language. As mentioned before, this implies that learning new words through reading gets more difficult when the proficiency level is low, as the reading task becomes harder than with higher proficiency levels.

## 2.2. Factors Related to the text

If the previous part showed how research has dealt with factors related to readers that may affect the acquisition of vocabulary when reading in an L2, this section will deal with factors associated to texts. From the moment they are written and released for the

public to read, texts will also condition what the reader can learn from them. Primarily, texts are made of words and not all of them are easy to acquire. However, difficulty or easiness do not necessarily have to do with the type of word we are talking about (nouns, verbs, concrete, abstract) but also on how often words are repeated for example. Certain genres repeat words more often than others and this seems to be positive for acquisition of new vocabulary. Contexts in which words are found can also determine if these words will be learned easily from that text or not.

### 2.2.1. Words' features and word families

Words, in any language, are overwhelmingly numerous and cannot be considered equally. They each have their own semantic and morphological properties, among others, as well as other possible difficulties, which can influence their learning and acquisition through reading. In a study concerned with the rate at which learners retained new vocabulary from reading graded readers, Waring and Takaki (2003) noticed that some words seemed to be more easily learned than others. This led them to think that depending on the text in which they were found, some words could be easier to acquire than others (e.g. graded readers focused on meaning and it might be easier to learn from there than from a list of words in isolation).

It is possible to look into a more specific feature that some words possess and that can affect their ease or difficulty to be learned, and this is the case of words that present *polysemy* or *homography*. *Polysemy* involves having two or more related meanings in the same form and, while somewhat similar, *homography* is about the two or more meanings of the same form being unrelated (Ming-Tzu & Nation, 2004). With this in mind, polysemous words would seem easier to learn since their different meanings are related and so, the more of them you learn the better you might understand them. Words that present *homography*, however, could be either more difficult to learn, due to the confusion they can cause to the reader, or easier, as the possibility of having more than one meaning might make them more conspicuous. Whatever the case, if at least one of the meanings has been learned and the form of the word has already been acquired by the reader, adding a new meaning should be less of an effort than learning a whole new form with a meaning. Finally, to conclude this part, we can say that if the new L2 word bears any similarity with the same word in the L1 of the reader, learning will probably be easier, as in the case of cognates (Webb, 2007).

In the case of word families, as they are composed of word-forms that are similar or related (Ming-Tzu & Nation, 2004), if readers already know one or more of the forms of a family, and have some knowledge on the morphology of the language, they should be able recognize new words of the same family with ease (Nation, 2001; Nation, 2006). At the moment of encounter readers may not know the exact meaning, but being familiar with the form, and the possibility of the meaning being similar to the other word-forms, will probably facilitate the learning.

### 2.2.2. Word frequency in the language

As we have mentioned before, words in texts are not in isolation, so they also need to be considered as part of the whole language system they belong to. Some words can appear more often than others and this will also determine learning. Turning now to the frequency of use of words in a language, one can easily notice that some of the words in a language see a much larger use than others. They can be classified into high frequency words and low frequency words, the former constituting a considerable smaller group than the latter (Nation, 2001). Although Horst et. al (1998) considered that, if their input level of English was not enough, readers would not benefit from the repetitions of high frequency words, that also could mean that, provided they have an appropriate level of input, readers should be able to learn a few unknown high frequency words thanks to the several possible encounters in a text. Low frequency words, however, seem to pose a higher difficulty to be learned, not only because of the sheer number of them, but also because repetitions in a text may be few and far between, unless the specific word sees a higher use rate due to the particular type, genre or content of the text. Taft (2013) comments on the frequency effect and how words seem to be stored in the mind according to frequency, meaning that high frequency words are accessed more easily than low frequency words, which probably helps in the steps towards their full learning.

### 2.2.3. Word encounters in the text

It is also worth noting that, as we have mentioned earlier, word repetitions also affect vocabulary learning through reading. Repeated encounters of the same word in a text are believed to play a major role in learning. The idea is that, for new words to be strongly encoded and connected in the mental lexicon, repeated encounters of those words in a text are needed, so learning them incidentally becomes a cumulative process (Rott, 2007) in which each encounter is seen as coming a step closer of accomplishing it. But if it is a

fact that repeated encounters through reading help in the learning of new words, then the question that unavoidably arises is ‘How many encounters are needed for learning or acquisition to happen?’. Different studies have come up with different answers: some studies claim there will be consistent learning at about 8 encounters or more (Horst et. el, 1998), between 8 and 12 (Nation, 2001), 10 to 15 or more depending on the word and other possible factors (Waring & Takaki, 2003; Webb, 2007), and more or less encounters depending on the proficiency level of the learners (Pigada & Schmitt, 2006). Therefore, even if these studies show that a certain number of encounters is necessary, there does not appear to be an exact or approximate number that ensures that learning will take place (Webb, 2007). This is probably due to the many other factors involved in the process of learning new words from reading that we are disclosing in this project.

#### 2.2.4. Context and register

Aside from the words and language, it is also important to take into account the characteristics of texts themselves, as they also influence word learning. Texts can differ in genre and this can cause certain words to be repeated more times than others. Also, the contexts they offer around unknown words influence in their understanding and possible learning. Starting with context, we can say that as words in a text are always found in context with other words, it becomes unavoidable to see them in relation to one another. However, contexts can become a double-edged sword when it comes to vocabulary learning, and that happens because while some do provide hints as to the meaning of an unknown word, others may provide confusing or very little information (Webb, 2007). So, it is not always a good idea to place all the bets on context for understanding unknown words in a text.

Regarding text types and register, because each type has its own features and is meant for a certain audience, the vocabulary used can turn out to be the same across different texts, as long as they belong to the same type or have the same register (Nation, 2001). Therefore, if the unknown word happens to be one that is used often across a certain type of text or register, even if it happens to be a low frequency one, readers will have more chances to encounter it and learn it.

### 2.3- Factors related to both the learner and the text.

The factors mentioned in the previous two sections concerned either the reader or the text individually, but it is important to remember that the relationship between reader and text is an interactive one. This means that there may be factors in the learning of new vocabulary that can only be noticed when looking into that interaction. The most important factor to be considered is coverage.

#### 2.3.1 Coverage

While vocabulary size refers to the total of words in a particular language, text coverage is concerned with the total of words that the learner knows from a particular text. In addition, text coverage is usually measured in percentages and seen as the percentage of the total of words that readers know in a text (Nation, 2006). Vocabulary size cannot be measured in percentages, as the overwhelming amount of words that a language possesses would make it impossible to reach accurate results. The two notions are closely related, because readers must have certain vocabulary sizes to be able to meet different text coverage percentages. Therefore, it becomes quite clear that text coverage is related to comprehension: below a certain percentage, the amount of unknown words would make it too hard for readers to follow the text (Nation, 2006). Nation (2001) establishes an easy to understand depiction of coverage percentage in texts, where 80% of text coverage results in 1 word in every 5 being unknown, 1 in every 10 for 90% and 1 in every 20 for 95%. Also, Hu and Nation (2000) analyzed these text coverage percentages in relation to text comprehension and found that, at 80% comprehension was impossible, at 95% comprehension was still hard, and only from 98% onwards was comprehension reasonable for readers. This is relevant for vocabulary learning through reading since, if text coverage happens to be too low, readers will feel overwhelmed and probably not learn anything, whereas if coverage is high enough, the few unknown words will be surrounded by known words that will facilitate their learning (Laufer, 1989). Furthermore, regarding extensive reading, as readers will be probably reading a text of their choice for their own enjoyment, text coverage is expected to be rather high and, therefore, there are higher chances for them to learn new words. Even if coverage is a bit low for a reader, as long as comprehension is not hindered, the learning of new words might still take place. If words are correctly inferred, this means that the coverage of a particular text for a particular reader will be higher.

### **3. Linguistic and extralinguistic support when reading**

In this final section, it would be interesting to shift the focus towards the real and concrete context where the learning of new vocabulary takes place. In other words, having a look at the types of texts readers engage with will probably help in seeing how some of the different factors mentioned previously, from the reader, text and both, fall into place. In addition, apart from the linguistic adaptations that can be made to particular texts, it would also be beneficial to look at some extralinguistic support that can be used to encourage vocabulary learning more actively. Regarding linguistic adaptations and types of texts, they can be easily classified into simplified texts, which constitute texts that have been revised and altered somehow, and original or unsimplified texts (Nation, 2001), which are texts in the original form they were originally written. As for additional extralinguistic support, we focus here on two well-known tools that learners can use, these are, on the one hand, glosses and, on the other, aural support and pictures (e.g. in audiobooks).

#### **3.1 Simplified and unsimplified texts**

As a starting point, just like in the previous case at the beginning with extensive and intensive reading, simplified and unsimplified texts are different types of material that offer each their own benefits to the reader and have as well some lacks and shortcomings due to their nature. Interestingly enough, extensive reading can be related to simplified texts, in the sense that early L2 readers or readers with low vocabulary or proficiency level are better off starting with easier texts since they can develop and learn from them. In a similar fashion, intensive reading can be related to unsimplified texts, as these early L2 readers can benefit from experiencing short extracts from unsimplified material and analyzing them (Nation, 2001). In the case of simplified texts, Nation (2001) notes that they are usually known as graded readers and can be found in different levels, each aimed at a readership with a particular proficiency level or vocabulary size and, therefore, written with a limited vocabulary, but still remaining full books that can be enjoyed. He further claims that their main advantages are the possibility of expanding the reader's vocabulary size by means of the context created by known words, encountering words seen recently and having the chance to reinforce their learning, and practicing reading skills so that readers can read faster and more comfortably. Waring and Takaki (2003) agree with Nation (2001) in that graded readers are good for reinforcing already known vocabulary, but their study on graded readers, together with Horst et. al (1998), also shows that not many new words are learned through their reading, especially if the L2

readers are at a low-intermediate proficiency level. A possible explanation for this could be that if readers already know most of the words, even if new words appear, as long as the context is enough for them to have a good grasp of the meaning of the text, new words could be skipped, as they do not hinder the reader too much (Bordag et. al, 2015).

This is not the case, however, with unsimplified texts that are written for a native speaker audience, as although they place a vocabulary burden, among others, on early L2 learners (Nation, 2001), they possess a syntactic and grammatical complexity that apparently can help readers learn new vocabulary. Bordag et. al (2015) found that when syntactic complexity raises, readers need to pay more attention to words in order to understand the context, and that results in better semantic representations of new words they encounter. However, that also means that the reading process becomes slower, as paying more attention translates into investing more time on words, and that can easily tire readers, which would make any read not as pleasurable as it could be. Bringing all together, it is possible to discern that, in the case of L2 readers, both unsimplified texts, which are more useful for high proficiency readers, and simplified texts, which come in handy for lower level readers, have their positive and negative sides, but that any is better than the other and could be equally fitting for different learning purposes.

### 3.2 Glosses

Generally speaking, glosses (or definitions of target vocabulary in a text) are a way of aiding readers with difficult or challenging words or expressions that they encounter in the text. They usually come with definitions of words or expressions and, depending on the kind of text or other factors like space available on the margins, can be found in different parts of the text (sometimes glosses are found at the end of the book for example). The main advantages they provide are the possibility of providing definitions and examples for low-frequency unknown words if there are not too many, raising the coverage of a text and, therefore, enhancing the comprehension of it, and making readers pay more attention to new words, which can add to their future learning (Nation, 2001). Rott (2006) carried out a study in which different types of glosses were analyzed, and found that words met in glosses left a form-meaning connection that could last up to 6 weeks later. Needless to say, however, glosses are not an all-encompassing solution to access texts of high difficulty, as there is a limit to how many words can be glossed without throwing the text off balance. Therefore, although they are good for readers that



want to try a text that is a bit over their level of proficiency, they cannot compensate for the large amount of unknown words that an unsimplified text can contain.

### 3.3 Multimodal input

Texts can often be found accompanied with images that give a visual input of what is written on a certain page. This can help readers get a better understanding of what is being said in the text, making hard to grasp parts easier to follow. In a similar way, texts can also be read with the aid of oral input, meaning that readers read the text while listening to someone reading it out loud simultaneously. At a glance, it would seem that having input other than written words could cause the reading pace of readers to slow down, as they would be expected to divide their attentional resources. However, this is not necessarily true, as Pellicer-Sánchez et. al. (2018) found that their participants took more time to read a text when it was not supported with oral input than when it was. In the case of images, their results showed that even if participants did notice and pay attention to them, focus was still placed on the text, the time spent on images was not significant enough to hinder the reading process. In addition, this kind of support does provide good benefits to the reader, such as reading fluency, in the sense that the oral input influences the reading pace of readers, or access to texts that they could not read without this kind of support, which may in turn enhance their language proficiency and vocabulary knowledge (Pellicer-Sánchez et. al, 2018). A possible negative effect would be that, although the use of these kinds of inputs would probably greatly enhance learning in the language classroom, in the case of extensive reading, when learners are reading for pleasure, this may not probably be so relevant and most books do not offer this possibility.

#### **4. Conclusions**

After having examined some of the factors involved in the process of vocabulary learning through reading, related to the reader, the text and both, the complexity that learning a L2 entails mentioned at the beginning becomes more apparent. Starting with the contrast between intensive and extensive reading, the way in which they differ and complement each other already points toward several paths that the process of learning new words can take. More specifically, intensive reading seems to enhance the reading skills of readers and prepares them for the future extensive reading of texts. Regarding new L2 vocabulary learning, neither intensive nor extensive reading alone can be the key to learn new vocabulary as there are many factors intervening in the process of growing a vocabulary. However, in the case of extensive reading there seems to be a reinforcement of already familiar vocabulary, and intensive reading can offer a good chance to practice ways of dealing with new vocabulary explicitly in the classroom.

Moving onto the factors involved in vocabulary learning through reading, when it comes to the reader, vocabulary size and proficiency seem to be the most important ones since any lacks can provoke problems when reading and more effort on the part of the reader when engaging with texts. Age, however, seems to be relevant only when considering the current age of the reader, due to developmental processes, and not the age when the exposure to the L2 started (as the amount of contact with the language can be more important than AoO). In the case of text, words themselves are one of the most important agents intervening in the process: their intrinsic difficulty, similarity to the L1 of the reader, frequency of appearance or contexts in which they are found are all crucial to help or hinder the learning of new vocabulary. Finally, coverage proved to be of relevance for vocabulary learning, and this is related to both reader and text: depending on the percentage of words from the text that the reader knows, the learning of new words will be easier (if it is high enough) or harder (if it happens to be too low).

We would also like to mention that other factors not contemplated in this study may be of relevance to determine the amount of lexical knowledge that can be gained from reading. For example, other factors that should be taken into consideration can be attention and motivation on the part of the learner or saliency in relation to words in the texts (e.g. words in bold, in italic, etc. may be easily noticed by the reader and more attention could be put on them). Other individual variables such as reading speed and short and long term memory can also be considered.

Finally, the analysis of linguistic and extralinguistic support when reading showed a concrete example of how these factors influence the process of new L2 vocabulary learning. Graded readers, with set vocabulary sizes and aimed at a particular proficiency level, are good for strengthening previously seen vocabulary, whereas unsimplified texts can offer a syntactic complexity that makes readers focus more on contexts and words, making the learning new words more probable. In the case of glosses, they can help readers reach a higher coverage level for texts they are not be able to understand well enough yet. Regarding the use of oral or input together with the written text, it can help readers focus on the task at hand and help them understand it better, opening the chance for learning to happen, also if images that are clear enough (i.e. not confusing for the learner) are present to support learning.

We are aware, though, that this work is just “the tip of the iceberg” regarding a satisfactory classification of the factors involved in the process of vocabulary learning through reading. Future work along the same line could aim to analyze more factors and even come up with new possible categories different from the ones presented in this work.

## **5. My own experience on learning L2 vocabulary from reading**

My first approach to reading in an L2 was through intensive reading of textbook fragments in a classroom context. Even if at the time I was not aware of it, the language activities and stream of readings available made the development of reading skills possible. This became apparent when I started with the extensive reading of graded readers. Although it was quite slowly at first, reading became possible and easier with each new graded reader.

I do remember encountering unknown words, but I do not recall learning many of them, as my focus was rather on words I could just recognize and the overall meaning of the text. Graded readers proved to be good to get into the world of reading in a L2 and reinforce vocabulary and grammar I was already familiar with, but not so good to expand my vocabulary.

It was not until I was older, and my proficiency level was higher, and after having been exposed to more English, both written and spoken, that I started to engage with unsimplified texts. At first, my vocabulary size was not big enough and, therefore, text coverage was quite lower than what would be recommended for pleasure reading of texts. To make up for my vocabulary lacks, I had to draw upon dictionary definitions and that slowed my reading pace. However, with each novel, and even within the same novel, the same unknown words would appear repeatedly, and after looking for them several times I started to learn and retain them. Little by little text coverage for unsimplified texts started to raise and my reading pace became quicker, which made me enjoy L2 reading more and encouraged me to read more texts and encounter more words. Usually I would learn these new words as synonyms of words I already knew, so that I could remember them better later. In my case, once I had access to unsimplified texts, with a coverage that allowed me to grasp the general meaning, “a snowball process” took place by which my vocabulary size saw a considerable and continuous increase. I guess this trajectory can also be found in other L2 learners, but this would be an issue for further research.

## **6. Works cited**

- Bordag, D., Kirschenbaum, A., Tschirner, E., & Opitz, A. (2015). Incidental acquisition of new words during reading in L2: Inference of meaning and its integration in the L2 mental lexicon. *Bilingualism: Language and Cognition*, 18(3), (p. 372-390). Cambridge: Cambridge University Press.
- Buckingham, D. (1993). *Children talking television: The making of television literacy* (Vol. 9). London: The Falmer Press.
- Chandler, D. (1997). An introduction to genre theory. *The Media and Communications Studies Site*.
- Horst, M., Cobb, T., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a foreign language*, 11(2), (p. 207-223).  
[https://www.lexutor.ca/cv/beyond\\_a\\_clockwork\\_orange.html](https://www.lexutor.ca/cv/beyond_a_clockwork_orange.html)
- Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a foreign language*, 13(1), (p. 403-430).
- Laufer, B. (1989). What percentage of lexis is necessary for comprehension?. *From humans to thinking machines*, (p. 316-323).
- Lightbown, P. M., & Spada, N. (2013). *How languages are learned 4th edition-Oxford Handbooks for Language Teachers*. Oxford: Oxford University Press.
- Meschyan, G., & Hernandez, A. E. (2006). Impact of language proficiency and orthographic transparency on bilingual word reading: An fMRI investigation. *NeuroImage*, 29(4), (p. 1135-1140).  
<https://doi.org/10.1016/j.neuroimage.2005.08.055>
- Ming-Tzu, K. W., & Nation, P. (2004). Word meaning in academic English: Homography in the academic word list. *Applied linguistics*, 25(3), (p. 291-314). Oxford: Oxford University Press.
- Miralpeix, I., & Muñoz, C. (2018). Receptive vocabulary size and its relationship to EFL language skills. *International Review of Applied Linguistics in Language Teaching*, 56(1), (p. 1-24). <https://doi.org/10.1515/iral-2017-0016>
- Muñoz, C. (Ed.). (2006). *Age and the rate of foreign language learning* (Vol. 19). Multilingual Matters. Clevedon: Cromwell Press Ltd.
- Nation, I. S. (2001). Vocabulary and reading and writing. In *Learning Vocabulary in Another Language* (p. 204-262). Cambridge: Cambridge University Press.
- Nation, I. (2006). How large a vocabulary is needed for reading and listening?. *Canadian modern language review*, 63(1), (p. 59-82).  
<https://doi.org/10.3138/cmlr.63.1.59>

- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a foreign language*, 18(1), (p. 1-28).  
<http://nflrc.hawaii.edu/rfl>
- Pellicer-Sánchez et. al. (2018) L2 reading and Reading-while-listening in multimodal learning conditions: an eye-tacking study. *ELT Research Papers 18.01* (p. 1-29). London: British Council 2018.
- Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning*, 57(2), (p. 165-199).  
<https://doi.org/10.1111/j.1467-9922.2007.00406.x>
- Taft, M. (1991). Evaluation of models. In *Reading and the mental lexicon* (p. 11-57). UK: Lawrence Erlbaum Associates Ltd.
- 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), (p. 130-163). <http://nflrc.hawaii.edu/rfl>
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied linguistics*, 28(1), (p. 46-65). <https://doi.org/10.1093/applin/aml048>
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of cognitive neuroscience*, 8(3), (p. 231-256).  
<https://doi.org/10.1162/jocn.1996.8.3.231>