UNIVERSITY OF Southampton

University of Southampton Research Repository

Copyright © and Moral Rights for this thesis and, where applicable, any accompanying data are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis and the accompanying data cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content of the thesis and accompanying research data (where applicable) must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holder/s.

When referring to this thesis and any accompanying data, full bibliographic details must be given, e.g.

Thesis: Author (Year of Submission) "Full thesis title", University of Southampton, name of the University Faculty or School or Department, PhD Thesis, pagination.

Data: Author (Year) Title. URI [dataset]

UNIVERSITY OF SOUTHAMPTON

FACULTY OF HUMANITIES

Modern Languages and Linguistics

Volume 1 of 1

L2 French lexical development of undergraduate students in a UK university

by

Virginie Pignot-Shahov

Thesis for the degree of Doctor of Philosophy

June 2018

FACULTY OF HUMANITIES

Modern Languages and Linguistics

Thesis for the degree of Doctor of Philosophy

L2 FRENCH LEXICAL DEVELOPMENT OF UNDERGRADUATE STUDENTS IN A UK UNIVERSITY

Virginie Pignot-Shahov

ABSTRACT

Growing a vocabulary of thousands of words is an essential part of foreign language learning (Milton, 2008: 236) and it is widely acknowledged that vocabulary size "plays a crucial role for L2 learners' communicative competence" (Stæhr, 2008: 139). And improving communicative competence itself is often what motivates L2 learners to persist with language learning to an advanced level. Therefore it is important to understand not only the trajectory of lexical development among advanced learners, but also its evolving relationship with overall proficiency and L2 motivation.

Whereas across the whole of Europe the teaching of foreign languages is increasing, particularly the teaching and learning of English, in the UK, foreign language learning is decreasing (Busse and Walter, 2013: 435). It is in this particular linguistic context, where learning a language other than English (LOTE) is not always considered useful, that the present study aims at investigating the lexical development of advanced L2 French learners at university, a language and an educational setting that are still seldom the focus of SLA studies. To determine the lexical development of first, second and final year undergraduate students, receptive and productive vocabulary tests were given to forty participants at the beginning and at the end of the academic year. The relationship between their L2 French lexical development on one hand and motivation, L2 proficiency and L2 exposure on the other was further investigated through the use of a motivation questionnaire, an Elicited Imitation test and a language learning background questionnaire. Based on these empirical data, this cross-sectional study presents some of the key findings on vocabulary development over the 3 years of the undergraduate degree, examining this development with reference to vocabulary known at different levels of frequency in French and the extent to which motivation, L2 proficiency and L2 exposure might explain L2 vocabulary variation. Findings from the study reveal that the development of receptive vocabulary is closely related to the development of overall L2 proficiency. L2 vocabulary development does display a frequency pattern, with increases in low frequency bands as proficiency develops. However, the lack of statistically significant difference at the end of the academic year between the year groups' receptive vocabulary size suggests that it is the quality rather than the size of vocabulary that improves over time. A period of one year living in a French-speaking country boosts overall proficiency and reduces intra-group variation. Factors positively influencing L2 vocabulary development were identified and include instrumental-prevention motivation and input frequency and quality. The study also reveals an imbalance in incidental activities undertaken by participants in the home setting, with little meaning focused output (writing and speaking) and fluency activities overall. The results of the study contribute to an understanding of L2 vocabulary development amongst L2 French advanced learners in relation to proficiency, motivation and input. Practical suggestions for pedagogy and future research were also identified.

List of contents

List of tables and figures

Research Thesis: Declaration of Authorship

Print name:	Virginie Pignot-Shahov

Title of thesis:	L2 French lexical development of undergraduate students in a UK university

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission

Signature:		Date:	01/06/2018

Acknowledgements

Firstly, I would like to thank wholeheartedly my supervisor Professor Rosamond Mitchell for her advice, guidance and encouragement throughout the study. I also wish to thank Dr Julia Hüttner for her help in times of need as well as my colleagues at the University of Southampton, in particular Arnaud Fresnel for his help with collecting data.

I am also very grateful to the forty students who gave up their time over the course of the academic year to take part in the present study. Without their help and enthusiasm, this study would not have been possible and it is a privilege to work alongside such fine students.

Finally, on a personal level, I wish to thank all my friends and family. My biggest thanks go to my husband Pande, my daughter Anastasia and my mother Jeannine for their ongoing support and patience during this research journey.

Abbreviations

ACTFL	American Council on the Teaching of Foreign Languages
A-level	Advanced level
BFP	Brainstorm Frequency Profile
BIA	Bilingual Interactive Activation
BICS	Basic Interpersonal Communicative Skills
BLC	Basic Language Cognition
BNC	British National Corpus
CAF	Complexity, Accuracy, Fluency
CAL	Centre for Applied Linguistics
CALP	Cognitive Academic Language Proficiency
CEFR	Common European Framework of Reference for Languages
CSE	Certificate of Secondary Education
DfE	Department for Education
DIWK	Depth of Individual Word Knowledge
EAT	Edinburgh Associative Thesaurus
EBacc	English Baccalaureate
EI	Elicited Imitation
EIT	Elicited Imitation Test
GCE	General Certificate of Education
GCSE	General Certificate of Secondary Education
HE	Higher Education
HEFCE	Higher Education Funding Council for England
HLC	Higher Language Cognition
IB	International Baccalaureate
LDT	Lexical Decision Task
LFP	Lexical Frequency Profile
LLC	Language Learning Common

LOTE	Language Other Than English
LP	Language Proficiency
L1P	L1 proficiency
L2P	L2 proficiency
ML	Modern Languages
PCA	Principal Component Analysis
PSTM	Phonological Short-Term Memory
SEM	Structural Equation Modelling
SLA	Second Language Acquisition
SOPI	Simulated Oral Proficiency Interview
STM	Short-Term Memory
SW	Stimulus Word
TOEFL	Test of English as a Foreign Language
UG	Universal Grammar
VLT	Vocabulary Levels Test
WA	Word Association
WM	Working Memory

Chapter 1 L2 French lexical development of undergraduate students in a UK university

1.1 Introduction

In the last few decades, vocabulary has regained an important part in second language acquisition studies and many theories have been developed putting vocabulary at the heart of language learning and advocating that the development of grammar knowledge is actually lexically driven. These include, for instance, Pienemann's (1998) processability theory, Kaplan and Bresnan's (1982) Lexical functional grammar, Kempen and Hoenkamp's (1987) procedural account of language generation and Ellis' (1997) lexical learning hypothesis.

Consequently, and even though there is no unified theoretical view of the place of vocabulary in L2 development, many empirical studies about vocabulary have been carried out, concentrating on topics such as:

- the development of the breadth and depth of L2 vocabulary (David 2008, Tidball and Treffers-Daller 2008, Richards, Malvern and Graham 2008);
- lexical development and its effect on skills (Hilton, 2008; Stæhr 2008);
- vocabulary tasks and vocabulary uptake (Pauwels 2012, Milton 2008);

Some of the recent empirical research has led to some strong assumptions about L2 lexical development. For instance, there is now very strong evidence for a frequency effect on L2 vocabulary learning (Milton, 2008). New avenues have also been opened by generative theory, such as determining what lexical threshold might be necessary to trigger grammatical knowledge (David, Myles, Rogers and Rule 2009).

However, despite active research in vocabulary learning, it is still not clear how vocabulary is learned. As Schmitt points out (2010: 36) "it will probably take a large number of studies using a combination of methodologies before the key developmental patterns become obvious". The relationship between passive and active vocabulary knowledge in particular needs to be clarified as well as the acquisition of discourse-related aspects of vocabulary knowledge such as register, the development of implicit/procedural vocabulary knowledge (DeKeyser 2007, Hulstijn 2007) and development of automaticity.

In this present research study, the principal gaps we aim at investigating are linked to L2 French and the specific learning environment of a university in the UK. Indeed, in much SLA research, the foreign language studied is English, but in order to understand the processes involve in second language learning, information from one language only is not enough. Although SLA research focusing on French is more and more frequent in schools and colleges, its study at university level in the UK is relatively rare (Milton, 2008: 228). While a few studies have been carried out for French vocabulary acquisition at school level, suggesting that overall vocabulary knowledge is small and the rate of development is limited, very little is known about vocabulary development post-A-Level, and how far the vocabulary deficit among school age students is compensated (see Chapter 3 for literature review). Therefore, there is a huge scope to investigate the vocabulary uptake and development of undergraduates of French in the UK across the four years of their studies. In particular, the study will examine development with reference to vocabulary at known different levels of frequency in French (drawing on recent corpus based studies of French vocabulary such as David, 2008). To the author's knowledge, no such frequency-based study of learner vocabulary has been carried out with English learners of French, though a parallel study for learners of Spanish has been done (Tracy-Ventura et al., 2014).

The present study also sets out to explore a range of possible factors influencing the rate and nature of French vocabulary development among undergraduates. The factors to be explored include:

- Learner motivation
- General L2 proficiency
- Extent and nature of exposure to French

Across the whole of Europe "foreign language teaching in secondary education is increasing" and "86% of pupils were learning English in 2006" (Busse, 2013: 115) whereas language learning is decreasing in the UK; this makes it particularly interesting to research the motivation of advanced L2 learners, and its relationship with proficiency development including L2 vocabulary. An improved understanding of the motivation of successful L2 learners in the UK can have a direct impact on the language learning environment as "improving retention rates is high on universities' agendas worldwide and is particularly important in English speaking countries" (Busse, 2013: 951).

Growing a vocabulary of thousands of words is an essential part of foreign language learning (Milton, 2008: 236) and it is widely acknowledge that vocabulary size "plays a crucial role for L2 learners' communicative competence" (Stæhr, 2008: 139). Indeed, all L2 related tasks, such as reading, require the L2 learners to have a large vocabulary in order to make progress. For Laufer (2010: 16) for instance, students need to know 95% of running words to understand a text, while for Hu and Nation (2000), this goes up to 98%. Milton reckons that "knowledge of less than 80% of the words in a text generally makes it incomprehensible even for gist" (2006: 199). Stæhr (2008) found a strong correlation between vocabulary size and listening and writing but the strongest link was with reading (0.83). That is to say L2 vocabulary development impacts on communication, and in turn on proficiency and further acquisition.

It may be assumed that L2 vocabulary knowledge is a central component of overall L2 proficiency. However, the relationship between proficiency development and the rate and nature of L2 vocabulary learning is less clear. Therefore the study will also explore the relationship between proficiency and vocabulary development over time, using Elicited imitation (EI) to investigate the proficiency of L2 French advanced learners. Indeed, only two studies have previously used EI to test proficiency in L2 French (Burger and Chrétien, 2001; Tracy-Ventura, McManus, Norris and Ortega 2014). This study will thus also participate in the accumulations of interpretable EI findings in French and make it possible to use these findings to compare them with those in other languages too.

Finally, the study will seek to explore the extent and nature of students' exposure to French (through formal instruction, private/ informal study, and language use e.g. in study abroad settings), and how this relates to L2 development.

To carry out this project, a variety of measurements have been used such as a receptive vocabulary test, a productive vocabulary test, an EI proficiency test as well as a motivation and language learning questionnaire. This deliberate use of multiple tools is supported by Schmitt's view according to which "good vocabulary research is advantaged by multiple measures of vocabulary to better capture a wider range of lexical knowledge" (Schmitt, 2010: 152). Another reason why the lexical development instruments are varied in this project is to investigate the variation between learners as "how much individual variation lexical variation there is in vocabulary skills really needs to be made a top priority in L2 vocabulary acquisition research" (Meara, 1996a: 36).

The main questions that this research aims at answering are:

- What are the rate and the variation of lexical development in L2 French over a 3 year period of instructed learning?
- Is the frequency factor observable in L2 learners' lexical development?
- To what extent can the individual difference factor of motivation explain learners 'variation?
- To what extent does overall L2 proficiency explain the pattern of L2 vocabulary development over time?
- What is the relationship between the extent of exposure to French and vocabulary learning?

1.2 Thesis outline

To investigate these questions, in chapter 2, we will first review the major MFL policies that have shaped language teaching and learning in the UK in the past decades in order to understand better the current climate. We will also review studies that have focused on L2 French vocabulary acquisition in UK schools. Chapter 3 is the literature review and as such defines the theoretical framework of this study by reviewing the main areas of the project, namely, the mental lexicon, proficiency and motivation. In chapter 4, we will detail the methodological approach and the research design. In chapter 5, the collected data will be analysed before being discussed in relation to previous research studies in chapter 6. Finally, in chapter 7, the importance of the study will be evaluated and practical applications and recommendations for further research will be presented.

Chapter 2 MFL policy and practice in the UK

2.1 MFL policy and practice in the UK

To investigate the L2 French vocabulary development of advanced learners, we have adopted two theoretical frameworks, Levelt's language processing model, related to usage-based theories of L2 development, and Nation's word learning in instructional contexts. In both frameworks that will be discussed in Chapter 3, the role of input is central to L2 learning and to L2 vocabulary learning in particular. This is one of the reasons why one of the goals of this research is to investigate the relationship between the exposure to French and vocabulary learning. However, the L2 French advanced learners in this study are in an instructional setting, a UK university, that offers few teaching contact hours, hence leaving students free to decide how much to access additional input, in the shape of the many L2 French resources available outside the classroom. This is turn allows them to shape their environment and learning opportunities through their own agency, and to choose their learning goals (Duff, 2012). These opportunities and the exercise of agency are in turn mediated by learner motivation, so that there is an overall relationship between L2 motivation, learner engagement with L2 input, L2 vocabulary development and L2 proficiency development, as this study will demonstrate. However, prior to studying at university, most participants will have had most of their L2 French input shaped by language practices of their L2 French classrooms, themselves driven by MFL policies of the country. Therefore, in order to understand the profile of the target participants in this study (UK undergraduates), as learners of L2 vocabulary, it is necessary to understand the educational context, i.e. past and current policies and practices in second language teaching in the UK. Therefore, in this chapter, we will review the main policies that have shaped foreign language teaching in the UK in the last few decades before concentrating on the data available about vocabulary development observed in UK schools (Graham et al, 2008 ; David, 2008; Richards et al, 2008). Finally, as we view L2 vocabulary learning as incremental, investigating the role of instruction in shaping participants' prior L2 vocabulary learning seems essential.

2.1.1 Recent changes in foreign language teaching policies

To understand the background to contemporary teaching and learning and the resulting proficiency profiles of advanced L2 learners in the UK, we will focus on the policy developments that happened from the second half of the 20th century onwards. Indeed, the last century saw a

radical change in the way languages in the UK are taught. In the 1960s and 1970s, the pedagogical changes were first driven by the democratisation of education and the development of comprehensive secondary education which changed the makeup of the students' cohort experiencing foreign language education. As a result, the grammar-translation approach which clearly influenced vocabulary learning in the past with "direct vocabulary instruction included only when a word illustrated a grammatical rule" (Coady and Huckin, 1997: 6) and which had been the main teaching method up to then in secondary education, was widely abandoned for situational teaching methods as these methods were perceived as more accessible (Mitchell, 2010). The new focus on functional language and everyday topics also affected the nature and scope of vocabulary learning. These changes were not uncontroversial and led to continuing debates about standards. From then on and up to the present, many have continued to perceive languages as a difficult subject. In addition to changes in languages education due to the general democratisation of education, the obvious fact that English was becoming a lingua franca, pushed the British government to review the rationales for teaching FL in the UK.

2.1.2 Foreign language teaching and the EU

As early as in the 1970s, the European Union has clearly voiced its ambition to promote language learning and linguistic diversity in all its member states. And since joining the EU in 1973, the UK has tried to adapt to some of its language promotion policies. For instance, in the 1970s, the promotion of French in primary schools was encouraged in the UK but issues "such as problems with teacher supply, and lack of continuity during transition to secondary" (Mitchell, 2011: 54) meant that "the initiative lost impetus and was not supported as mainstream policy" (ibid).

More recently, the Barcelona objective agreed in 2002 by the EU's heads of state and governments states that member states need "to meet an ambitious goal: enabling citizens to communicate in two languages other than their mother tongue". (European Commission, 2013). Even though this objective was not explicitly referred to in the successive versions of the National Curriculum for England from 1990 onward, it is understood as one can see from the purposes of study described by the government which clearly adopted a reconstructionist view of language education which was "committed to the production of useful citizens with an equal chance of succeeding in the world of work and public life" (Mitchell, 2010: 151). The latest version of the National Curriculum (DfE, 2014) reflects a view of the language learner as global citizen:

> Learning a foreign language is liberation from insularity and provides an opening to other cultures. A high-quality languages education should foster pupils' curiosity and deepen their understanding of the world. The teaching should enable pupils to express their ideas and thoughts in another language and to understand and respond to its speakers, both in speech and in writing. It should also provide opportunities for them to communicate for practical purposes, learn new ways of thinking and read great literature in the original language. Language teaching should provide the foundation for learning further languages, equipping pupils to study and work in other countries.

2.1.3 Examinations and assessments

In 1988, "a new-style unified national examination (the General Certificate of Secondary Education) was introduced"(Mitchell, 2010: 159), replacing two previous examinations: the more academic "O Level" General Certificate of Education (GCE) examination for foreign languages which included tasks such as translation and dictation, and the Certificate of Secondary Education (CSE), which provided a more functional/ situational curriculum (Mitchell 2011: 50). This new examination conceptualised languages learning in terms of the four skills of Listening, Speaking, Reading and Writing; compared with the former O-Level, considerably greater emphasis was placed on oral and situational proficiency. Importantly for this study, a word list was specified for this examination, covering everyday topics. However, languages were still an optional subject then and it is only from 1990, and with the introduction of the National Curriculum, that foreign languages became a compulsory subject for all students aged 11-16. Indeed, foreign languages became one of the foundation subjects of the new curriculum for England and Wales and "the first cohort experiencing this *languages for all* requirement started secondary school in 1992" (Mitchell, 2011: 50).

The emphasis on the four skills was continued in the National Curriculum, which was conceptualised for assessment purposes in this way. Following National Curriculum principles, all subjects were to be assessed throughout students' compulsory education with reference to 10 levels of attainment "to cover the full period of compulsory education (from age 5 to age 16: later reduced to an eight-level model, covering ages 5-14)" (Mitchell, 2010: 159). The 10 levels of attainment framework was also the basis for the national tests in literacy, numeracy and science taken at ages 7, 11 and 14. For subjects such as languages, which started only in the secondary school, this multi-level model posed a challenge, as the 10 levels prescribed for each of the four

skills had to be covered in 5 years rather than 12. Teachers became preoccupied with demonstrating children's progress through these levels; and in addition, with unified national tests, comparisons between schools could be made easily, leading to some institutions being more prestigious than others. Thus one outcome of the National Curriculum was that compulsory education in general became much more assessment-led. The breadth of the curriculum, plus the increased national focus on numeracy and literacy, also led to a reduction in teaching hours for particular subjects, including languages.

This pressure on school results and the examination format is not to be neglected as it can have a serious impact on language learning and in particular on vocabulary development rate. For instance, Milton (2004) used past examination scripts to compare 21 learners "who had studied O-Level French between 20 and 50 years ago with 24 best students from an equivalent cohort that took GCSE". The latter group produced only half to a third of the vocabulary of the O-Level students. This decline, probably due to the reduction of teaching hours, the different examination format and the pressure to increase the number of students learning a language, has, of course, a direct impact on students and their ability to communicate effectively.

2.1.4 Focusing on teaching MFL in primary

Following the introduction of the National Curriculum, the number of students taking a language up to age 16 peaked at 78% in 1999-2001. However after 2000 and with some relaxation of the National Curriculum, languages became compulsory up to the age of 14 only and as a result, the number of students taking an examination in a foreign language at 16, dropped significantly. In order to compensate for the drop, from 2002, the "National Strategy" for languages was published (DfE, 2002) whose main aim was to promote foreign languages in primary schools. Important means were dedicated to this new commitment supported by an outline curriculum and considerable investment in resources and in teacher development, a scheme which was widely adopted but did not become compulsory until 2014. However, the worrying and ongoing decline of entries in A-level in a foreign language was unaffected by these lower level changes.

Around the same time as when English as a foreign language started to be widely taught in EU primary schools, many parents and language educators in the UK also wished for languages to be taught from a younger age, both because it was commonly believed that "the younger the better" as far as developing language skills were concerned but also because it might ease the transition with secondary. As a result, so-called Pathfinder projects were funded to "experiment with

curriculum models and teaching methods for languages" which led to the publication in 2005 of the Key Stage 2 Framework for Languages.

As noted by Mitchell (2010), the English National Curriculum does not provide detailed guidance for teachers about the best pedagogical practices to adopt or detailed assessment methods. This was also true in the case of languages, apart from specifying that all four skills, reading, listening, writing and speaking need to be developed. For this reason, the National Strategies were created in the early 2000s as a framework for much of the lower secondary school curriculum (years 7, 8 and 9), including languages. In the 2003 Framework for Teaching Modern Languages (Ashton, 2003), behavioural objectives were set for 5 strands: "words, sentences, texts, reading and writing, listening and speaking, cultural knowledge and contact", i.e. a considerable departure from the four skills usually referred to in the National Curriculum, and with explicit attention paid to vocabulary learning for the first time. Language professionals were generally in favour of the 2003 Framework because of its ambition "to favour explicit teaching of grammar in place of the broadly situational and phrasebook-style learning which had come to characterise lower secondary school languages pedagogy" (Mitchell, 2010: 162). This document was replaced in 2009 with a modified Framework which was more in line with the 2005 (voluntary) Key Stage 2 Framework to prime a smoother transition between upper primary and early secondary. However, due to a policy time lag, the introduction of the Framework took place at the point when languages became compulsory for 3 years only. Therefore the Framework was no longer a launch pad for the formal GCSE examination in ML at age 16.

In 2014, languages in Key Stage 2 were finally given compulsory status. According to Board and Tinsley (2015), this has already had a positive impact as 99% of the schools taking part in the annual Language Trends questionnaire, claimed to teach a language. However, there are still some serious concerns about foreign languages teaching in UK primary schools and some were already issues 40 years ago. For instance, some secondary schools are concerned that some primary schools do not have the resources and the staff to teach to a standard that will enable students to build on their prior knowledge when they arrive at Key Stage 3.

Since the development of comprehensive education, which initially led to the expansion of languages in secondary schools, we have seen that languages are often regarded as a difficult subject leading to a decreasing number of students taking a national examination in a language. Thus if in 2015, students understood the positive value of languages, they were still reluctant to take them at GCSE and A-level for fear or not obtaining a high grade which could have an impact on their college and university choice. More worryingly still, in both the state and private sectors,

pupils are sometimes excused from studying a language at all, for a range of reasons. “In the most economically deprived category of schools”, the numbers reach 17% at Key Stage 3 and 44% at Key Stage 4. (Board and Tinsley, 2015).

In January 2011, the Conservative/Liberal Democrat Coalition government conceived the EBacc (English Baccalaureate). Primarily a school performance measure, it was also presented as an award for pupils who gained C grade or above in core curriculum subjects which are English, mathematics, history or geography, the sciences and a language. This initiative was thought to help slow down the decline in take-up for languages which fell from 78% in 2001 to 40% in 2011 (Board and Tinsley, 2015: 11). It has had a timid but still positive effect on the percentage of end of Key Stage 4 students sitting a language at GCSE which went from 48% in 2013 to 51% in 2014 in all state-funded mainstream schools (ibid).

2.1.5 Languages in Higher Education

Having considered the ML policies concerning primary and secondary schools in England and Wales, let’s now turn to Higher Education. For almost two decades, languages in the HE sector have benefited from special financial support from the Government and other agencies, including £41.5m of new investment from HEFCE and other agencies for languages between 2005 and 2012 (Worton, 2009: 2). On the other hand, HE as opposed to other sectors, has been less referred to in the National Languages Strategy and according to the Review of Modern Foreign Languages provision in higher education in England: “there has been insufficient ‘joined up’ thinking about the role of foreign languages in the UK over the past decade; there remains no sense nationally or internationally that the UK is committed to multilingualism and thereby to informed cultural interactions” (Worton, 2009:2). For Worton, the absence of foreign languages from the National Curriculum after 14 has sent a negative message to students choosing HE disciplines. Consequently, there is “a small but steady decrease in the numbers of undergraduate students studying languages over the past several years” (Worton, 2009: 6). The picture differs from one HE institution to another but nowadays, ML departments are found essentially in pre-1992 universities and, indeed, mainly in Russell Group universities. (ibid)

But apart from the decrease in numbers of students, the main concern of ML staff in HE is about “the quality and preparedness of students” (Worton, 2009: 3). Respondents stressed the importance of supporting first year university students as “the majority felt that current A Level provision does not provide students with either the grammatical knowledge or the language learning skills necessary for university-level study” (Worton, 2009: 26). Interestingly, several

respondents noted that "*ab initio* candidates often outperform those entering with some prior knowledge of a language" (ibid). In other words, students studying a language from scratch at university can overtake someone who has been studying a language since year 7 in less than 2 years.

In response to this situation, there is an expectation that universities in the UK will themselves take the responsibility to develop "a clear and compelling identity for Modern Foreign Languages in an increasingly competitive higher education context, one which presents a convincing case for the contribution that Languages Departments make to the strategic objectives of their institutions and more widely" (Worton, 2009: 4). This has resulted in some moves such as the one away from specialised 'Single Honours Programmes' to Joint or Combined Programmes seen by students and employers are more useful degrees. So key elements to a vibrant ML department at HE are diversity and a strong understanding of the reasons why students and employers favour languages.

Despite considerable efforts, the number of students taking a language post-14 is declining. A number of factors influencing this decline have been considered, including the rise of English as lingua franca, the fact that languages are now an optional subject for GCSE, and their perceived difficulty at both GCSE and A-level.

Having described the overall background of UK for languages in education, and declining patterns of participations, in the next section, we will review studies about L2 French vocabulary development in schools, colleges and universities in the UK.

2.2 L2 French vocabulary learning in UK schools

In the last thirty years, vocabulary has become a central research focus of language studies but whereas studies of vocabulary development in instructed learners of L2 English seem to have flourished, considerably fewer have been carried out about instructed learners of L2 French, particularly in the UK. Undoubtedly though, the analysis of L2 French in these studies (David 2008; Graham, Richards and Malvern, 2008; Milton, 2006; Richards Malvern and Graham, 2008) can contribute to the development of the SLA research agenda as well as provide valuable information for the teaching of modern foreign languages in the UK.

2.2.1 Empirical studies of L2 French vocabulary development in the UK

Milton echoes Stæhr's (2008) opinion when stating that "the principal lesson that emerges from the study of vocabulary learning in foreign languages is that learners need to learn lots of vocabulary if they are to achieve any level of independent communicability" (2009: 249). It is also known that the learning of vocabulary in the classroom is a slow business. Different studies reviewed by Milton and Meara (1998) that focus on the rate of vocabulary learning and time spent in the classroom showed that even in the most fruitful classroom environment, students learn on average 6 words an hour.

Häcker observes in the assessment-dominated UK education system the classroom may not be an optimal vocabulary learning environment; she claims there is a negative wash back effect where teachers restrict input, including vocabulary, to those limited topics that occur in examinations (2008: 219). Häcker's comments are confirmed by David's (2008) findings about the relatively small vocabulary size of UK students taking GCSE. And compared to other pupils in the EU, Milton (2006a) suggests that UK secondary school pupils' L2 vocabulary growth rate and size is lower. It has been calculated that English students from year 7 to year 11 learn receptively as little as 170 words per year (Milton, 2006a). This, in turn, has implications in Higher Education as undergraduate students identify vocabulary as the main source of difficulty when reading in a foreign language (Gallagher-Brett, 2006).

The relatively low vocabulary growth from year 7 to year 11 is documented by Richards et al (2008) and also by David (2008: 172) who reports that the mean scores for an X-Lex test for year 8 is 439 and for year 11 students is 564. These results contrast greatly with the uptake happening over the two years after GCSE when pupils learn on average 535 words a year (David, 2008). The most important L2 vocabulary spurts which have been documented in the UK education system happen mostly between GCSE and A-level. "Learners take their B1 GCSE examinations with, on average, 850 words of French and the B2 Advanced level with just under 2000 words" (Milton, 2009: 84). The spurts of L2 vocabulary between GCSE and A-level shown in these two studies seems to suggest that when learners choose to take their learning of French beyond compulsory education, this decision has a direct impact on their vocabulary development. This is the case despite the finding of Graham (1997) that students in post-compulsory classes found the amount of new vocabulary to be learned in year 12 and 13 a challenge because of the lexical gap between GCSE and A-level, and that this had a direct impact on their perceived progress and their motivation.

Some variation in the rate of vocabulary learning has been noted in different parts of the UK school system. Richards et al (2008) compare the findings of their study of year 12 pupils in 20 schools in the South of England with 3 other studies: A small pilot study in a selective grammar school where students' receptive vocabulary was tested using the computerised version of the X-Lex test, a one year study of non-selective comprehensive schools and a study in 15 comprehensive schools. The X-Lex results revealed that year 12 grammar school pupils' receptive vocabulary ranged from 1900 to 4150 with a mean of 3129 whereas year 12 comprehensive students' scores ranged from 1250 to 3500 with a mean of 2437. (Richards et al, 2008: 204).

As we have seen earlier, when measuring the vocabulary size of second language learners the frequency factor is very often referred to. The assumption that learners add more frequent words to their L2 lexicon before learning words that are not so frequent has been well documented across many languages. Milton (2009: 242) notes that "the importance of frequency in vocabulary learning is as near to a fact as it is possible to get in L2 acquisition". There are however some critics such as Tschichold (2007) who are concerned that factors such as the age of the reference corpus or the extent of availability of cognates could bias analyses of frequency, and are rarely taken sufficiently into account in frequency tests. In the present study, the age of the corpus has been, to some extent, an issue when analysing some of the participants' answers to the productive vocabulary tests. For instance, in the Flex15 test, *internet* (internet) was often given as an answer to the stimulus word *ordinateur (*computer). However, the item internet was not included in Baudot's French frequency list. This however, did not have a negative effect on the frequency profile of participants as the words was counted as not *found*. Rather than the age of the corpus, we found that word categories of words in receptive vocabulary tests can bias analysis of frequency. For instance, in the X-Lex Nov test, all three year groups were 2K deficient. However, upon analysing closely the word categories, we realised that the 2K list included four closed category words, a pronoun, a conjunction, an interjection and a preposition. The other bands (1K, 3K, 4K and 5K) also included closed category words but only one or two per frequency band. So it seems that word category can bias frequency analysis and this was reinforced a few months later when in the X-Lex May test, none of the groups were 2K deficient and none of the selected words in the 2K band were from closed classes. In Chapter 6, this finding will be discussed further in the light of the frameworks we have adopted.

As far as L2 French in UK schools is concerned, Milton (2006a: 188) notes that "frequency effects are observable in learners after as little as 100 hours of tuitions at which point, breadth of exposure to language overcomes the idiosyncratic selections of individual textbooks". Milton also notices

that after a few years of French language instruction with a variety of textbooks, the frequency profile described by Meara emerges (Milton, 2006a: 196). That is to say despite the fact that textbooks which are organised thematically may contain low frequency words, learners tend to learn high frequency words before low frequency words. David also notices the frequency effect in her study of 483 learners of French in the UK: "average scores for each frequency band increased consistently from year-12 learners to native speakers" (2008: 173), apart from years 8 and 11 when little uptake of new vocabulary takes place.

Concerning lexical development beyond the secondary school stage, the study of David (2008) remains relevant, as she investigated the lexical development of students of French from year 8 to UG4 in UK schools and universities. This study provides some interesting findings about the vocabulary size of first and second year university students. For instance, David reports a fairly small mean difference of 330 words between UG1 and UG2, with mean scores of 2524 and 2854 words respectively. However the range of scores was very wide for both year groups, with minimum scores of UG1 and UG2 respectively of 750 and 1250. This means that the lowest scores for UG1 and UG2 reported by David are even smaller than the expected vocabulary size at A-level which Milton estimated at just less than 2000 words (2009). This again underlines the fact that even if it is possible to predict how groups will progress in their lexical development, there are still large individual variations.

It is possible to speculate that this variation is due to different overall proficiency levels related to memory, learning strategies and environmental exposure or motivation, and some of these factors are pursued further in this thesis. As our participants are motivated language students who may be dedicating a lot of their time outside the classroom to the development of their language skills, it will be interesting to observe how their vocabulary develops across the years and in particular to investigate the impact of resources used. For instance, do students who spend more time learning incidentally know considerably more vocabulary and rarer words than those who do not? Are there some particular incidental activities that are better for acquiring more vocabulary?

2.3 Conclusion

In this chapter, we have reviewed the landscape of L2 teaching and learning over the last couple of decades. It was necessary to contextualise L2 learning and teaching in the UK as ML policies have an impact on curriculum and therefore on the content taught and also on students' perception of the usefulness of languages and therefore on motivation. Because we view learners

as individuals who can be influenced by the environment in which they study, we felt it was relevant to analyse some of the most recent ML policies and their impact, including in particular the downward pressures in recent decades on instructional time and the consequences for vocabulary learning. We then reviewed available evidence on trends in the uptake of French vocabulary in UK schools and university courses, including the vocabulary spurts that happen a few months after the beginning of non-compulsory language courses (e.g. at the beginning of GCSE and A-level) , and the wide variations in vocabulary knowledge among university students. We will now turn to the literature review to determine the theoretical framework of this study considering in particular the mental lexicon and its development and the constructs of L2 proficiency and L2 motivation, whose relationships with vocabulary development will be explored later in the thesis.

Chapter 3 Theory: the mental lexicon, proficiency and motivation

3.1 Introduction

In the previous chapter, lexical development in L2 French has been reported in terms of vocabulary size, that is to say the number of words students learn over a period of time such as over an academic year for instance. However, L2 vocabulary acquisition, in addition to size, raises many other issues such as: What are the types of cognitive mechanisms that support L2 comprehension and production? Are these mechanisms similar to those responsible for L1 comprehension and production? Does L2 vocabulary development differ for individuals with different proficiency levels, with different types of word, or in different learning tasks? (Tokowicz, 2015: 1). Therefore, it appears clear that size is only one aspect of vocabulary development and in order to investigate vocabulary growth, additional questions need to be considered as well and a theoretical framework needs to be determined. Even though there is no single widely accepted model of L2 lexical organisation and processing (de Bot, Paribakht and Wesche, 1997: 309), Levelt's processing model with its lexicon components of concepts, lemmas and lexemes and his take on the organisation of the mental lexicon (1989) is the model that we have selected to underpin this study. Also, due to the potential pedagogical implications of studying the lexical development of L2 French learners at university, Nation's definition of a word (Nation, 1990) and the circumstances that seem to promote or impede the acquisition process are also considered. In 3.2, entitled "the mental lexicon", we will detail these two theoretical frameworks; and in the second and third sections, we will explore proficiency and motivation, the two significant individual differences which this study of L2 French lexical development focuses on.

3.2 The mental lexicon

3.2.1 What is the lexicon?

The concept of a lexicon implies that the words we know are stored and organised in ways that facilitate their fast retrieval so that we can communicate efficiently, and as Juffs states "no matter one's theoretical perspective, the lexicon is a key component of language" (2009: 181). The psycholinguistic study of the mental lexicon's nature has shed light on some part of its organisation but the way the mental lexicon is organised overall is still not clear. What we already know though is that it has to be organised in some way, otherwise, it would be very difficult for

speakers to hold in their memory randomly organised words. But despite its being the original meaning of the term "lexicon", the analogy of a dictionary is not a particularly helpful one in understanding how words are stored in the brain. Slips of the tongue such as "he told a funny antidote" with "antidote" instead of "anecdote" (Aitchison, 2003: 11) show that the initial sound does seem to play a role in the organisation of words in the lexicon, like initial letters in a dictionary, but this is by no means the most important factor driving the organisation as we will see in this chapter.

Another difference between the mental lexicon and a dictionary is their content. Whereas the number of words in a dictionary is fixed and new words will be added only in a new edition, new words are created and added all the time in the mental lexicon. Also, the information given about a word in a dictionary is limited whereas "there is no known limit to the amount of detailed information … which may be associated with a lexical item. Existing dictionaries, even large ones, specify lexical items only incompletely" (Hudson, 1984: 74). Dictionary entries are also presented in isolation whereas words in the lexicon are related to each other. For instance *warm* is related to *cold, tepid, hot* (Aitchison, 2003: 14).

Another difference is the fact that speakers are able to differentiate common words and rare words, i.e. are sensitive to word frequency, information that has only recently started being added in dictionaries. One last difference between the mental lexicon and dictionaries is that dictionaries present only one phonetic form for each word whereas speakers are able to recognise a word even if it is pronounced slightly differently.

So an alternative metaphor for the lexicon and one which is more suited is that of a web. Aitchison herself has described the lexicon as "a gigantic multi-dimensional cobweb" (2003). The analogy with a web has felt appropriate in trying to capture the multiple levels of interconnections between items in the mental lexicon in which different types of semantic link are supplemented by phonological, orthographic, syntactic and encyclopaedic connections between words (Wilks and Meara, 2002: 303). This complex nature of words is therefore one of the reasons why it is so difficult to establish what the connections are in the mental lexicon. Another consequence is that it becomes increasingly difficult to separate the lexicon from grammar.

It is not hard to see then why understanding the lexicon as a whole is a very challenging task and therefore different strands of research on the lexicon have focused on different goals. For instance, "psycholinguists such as Kroll and her colleagues (De Groot and Kroll, 1997; Kroll and Sunderman, 2003) focus on the relationship between L1 and L2 forms, meanings and storage"

(Juffs, 2009: 181). It is actually psycholinguists who refer to the lexicon as the *mental* lexicon or the *internal* lexicon (Schönefeld, 2001: 9). For Ellis and usage-based theorists who "are concerned that although language behaviour can be described as being rule-like, this does not imply that language behaviour is rule-governed" (Ellis, 1998: 638), the focus of mental lexicon studies is on the role of frequency in word acquisition, collocations and morphosyntactic patterns. Nation (2001) and others interested in word learning in instructional settings also concentrate on frequency of words and in particular how this is demonstrated in corpora of existing texts. Finally, sociocultural researchers focus on word development when learners are participating in culture (Juffs, 2009: 182). So different fields have approached the study of the mental lexicon from different angles and with different research agendas. They are all important and necessary, but in this research, we prioritise the viewpoint of researchers who are interested in word learning in instructional contexts such as Nation, and also those interested in the mental lexicon including its content and organisation, based on Levelt's L1 speech production model. But before exploring Levelt's model, we will first of all review in more detail how words have been defined, and what is understood by knowing a word.

3.2.2 The nature of word meaning

There are two main views on the meaning of words. On one hand, some believe that the meaning of words is fixed in terms of a set of features, and in this view, "the semantic entries in one's mental lexicon will be fairly cut and dried, and failure to achieve this ideal state will be due to lack of education or to mental laziness" (Aitchison, 2003: 41). The alternative point of view considers that words do not have fixed meaning but rather that "Natural language concepts have vague boundaries and fuzzy edges" (Lakoff, 1972: 183) and that meanings can shift.

According to both viewpoints, people use "concepts" to interpret the world and these concepts are linked to words such as *house* and *table*. So for instance, *cat* "refers not to the image of a cat but a somewhat more abstract amalgam" (Aitchison, 2003: 45). From the first perspective, "necessary and sufficient conditions" are essential to capture the meaning of a word, and word meanings comprise a set of "critical attributes" which speakers check one by one. So for instance to know the meaning of *cat* is to know that it is a four-legged animal that purrs and so on. The difficulty of course is to know what precise details should make up such a checklist. According to this theory, some attributes are more important than others and make up the core meaning which should be located in a sort of linguistic dictionary, whereas other attributes which are not essential are probably "stated in an encyclopaedia of general knowledge" (Aitchison, 2003: 46).

There are two main problems with this theory, the first one being that it is not possible to define the essential and non-essential characteristics of a word, the second one being that some words do not seem to have any necessary conditions.

On the other hand, some theorists believe that word meanings are fuzzy, so that it is not clear where the meaning of an individual word ends. In this view, meanings overlap and are interconnected. They believe that some words have a fixed core meaning but most words belong to a family of items "which all overlap in meaning but do not share any one common characteristic" (Aitchison, 2003: 51). Prototype theory deals with this fuzziness of meaning stating that speakers analyse a word against a prototype and even if the characteristics of a word do not match all the characteristics of the prototype, the two are linked anyway.

Prototype theory (Lakoff, 1999) has given rise to a variety of proposals for mental lexicon frameworks, some which are fixed "with empty slots into which items are inserted" and others which are built as the speaker searches for words. This theory and the frameworks which rely on it, that are discussed later in section 3.2.7 give us some clues about the elements the mental lexicon might be made up of but still not about its organisation. So words are linked together but how this happens is still unclear.

Here again, different theorists have different opinions on the way words are linked together. One of the theories is associated with early versions of Noam Chomsky's view of language, and follows from the view that word meanings are composed of semantic components. These semantic primitives are believed "to exist only at an underlying level" (Aitchison, 2003: 79). In this view, then words "might be stored in the human mind in a disassembled state" (Aitchison, 2003: 80). Logically, according to this view, words that are made up of more components would take more time to process. However, empirical research findings have been inconclusive and it is not possible at this point to differentiate between the time which is needed to retrieve more complex words and the time needed to retrieve simpler words, that is to say words that are in theory made up of fewer components. The existence of a universal store of semantic primitives has not been demonstrated and linguists do not actually yet agree on what makes up these primitives.

3.2.3 What is a word?

First of all, the study of vocabulary development and the rate of this development imply some measuring of words. In this context, it is helpful to distinguish between word types and word tokens, the former referring to the number of different words in a text or a corpus and the latter to the total number of words including repeats (Milton, 2009: 8). For instance, in the sentence *the cat ate the mouse;* there are 4 types but 5 tokens.

But should words be considered as wholes or should they be divided into smaller components, that is to say into morphemes, "the smallest grammatical unit"? (Aitchison, 2003: 126). Are they stored as wholes or as morphemes that are assembled together to form wanted words? One could argue that this latter option seem more plausible as storing whole words might put a strain on memory. On the other hand, assembling morphemes might put a strain on processing. There is strong evidence which shows that it depends on the morpheme. Thus, derivations seem to be more attached than inflections and regular inflections are added as we speak. However, frequent regular inflections like "s" in *peas* can be added to the word already. So, words can be stored both as a whole in the main part of the lexicon, or as smaller components (ibid).

Also, all languages divide words into grammatical categories such as nouns, adjectives and so on and all these categories play an important part in the sentence. However, content words, i.e. nouns, adjectives and verbs are considered by many as being the "lexicon proper" (Aitchison, 2003: 103). Some linguists believe that "speakers select a syntactic frame for a sentence, such as noun-verb-noun and then put appropriate words into the slots" (Aitchison, 2003: 103). But sentence structure is also very dependent on the choice or words, therefore, one can say that semantics and syntax overlap. According to both Aitchison (2003) and Schmitt (2010), words can be regarded as coins with meaning and word class on one side and sounds on the other. This combination of meaning, word class and form is described more technically as a lemma (Levelt, 1989), an important concept in the study of the lexicon. The storage of words, lemmas, lexemes and morphemes is complex and controversial though. We will come back to this point when reviewing Levelt's and De Bot's mental lexicon models.

A lemma "consists of a headword and some of its inflected and reduced forms" (Nation: 2010: 7). The lemma concept is based on the psycholinguistic finding that the mind stores the base form of the word, together with its syntactic role and connections. For instance, the lemma of *slow* includes *slower* and *slowest* which are also adjectives but not the adverb *slowly* because it is a different part of speech from that of the headword. One difficulty with lemmas arises with

irregular forms of a headword such as *bring/brought*. Whether they should be counted as separate lemmas or not is still under discussion. As Nation (2010) points out, when lemmas are used as a measuring unit, the number of units in a corpus can reduce greatly: for instance, the 61,805 types in the Brown Corpus were reduced by almost 40% to 37,617 lemmas. But as Vermeer (2004: 175) states, the lemma is "the most reliable unit of counting words". It is used to count and estimate vocabulary knowledge in vocabulary tests, such as Nation's Levels Test (Nation, 1990) and the X-Lex test (Meara and Milton, 2003), the former being the test that has been selected to measure receptive vocabulary in the present study and that we will detail in the Chapter 4.

Word families are another unit of vocabulary measurement similar to lemmas but which include all words related to the headword regardless of their word class. Therefore, *teach, taught, teaches,* but also *teacher* or *teachable*, belong to the same word family. Using word families instead of lemmas as a type of vocabulary count will produce even smaller vocabulary size figures as words that would be counted separately in a lemmatised count would fall under a single headword (Milton, 2009: 12). Word families have been used by Goulden, Nation and Read (1990) to estimate the vocabulary knowledge of native speakers. This is a logical choice as native speakers are most likely to know almost all the derivations and inflections of a word.

3.2.4 What does it mean to know a word?

As well as explaining what is meant by a word, it is also important to be clear about what it means to *know* a word. One commonly agreed distinction is that between receptive or passive lexical knowledge and productive or active knowledge. Productive knowledge usually refers to being able to use words in speaking and writing while receptive knowledge is associated with recognising words in listening and reading (Laufer and Goldstein, 2004). But the distinction between the two is not always so clear when we consider that "good passive skills often require the reader or the listener to actively anticipate the words that will occur" (Milton, 2009: 13). In other words, while listening and reading, the learner also draws upon productive knowledge.

Also, according to Laufer and Goldstein (2004: 405), "there is no consensus as to whether this distinction [between passive and active knowledge] is dichotomous or whether it constitutes a continuum". For Melka Teichroew (1982), receptive and productive knowledge are placed on a continuum, on which receptive knowledge gradually moves towards productive mastery as a result of the learner learning more about individual lexical items. This gradual cline from passive

to active has been widely accepted but the threshold at which receptive knowledge becomes productive is not clear (Laufer and Goldstein, 2004; Schmitt, 2010). The receptive and productive tests used in the present study, do not enable us to investigate the potential gradual learning that would turn receptive vocabulary into productive vocabulary. However, as we will see in Chapter 5, they yield some interesting findings in relation to overall proficiency.

For Meara (1997), the two types of knowledge actually represent different types of associational knowledge and therefore cannot be a continuum. He proposes a lexical organisation in which productively-known words are connected to a productive item in the L2 lexicon, whereas receptively-known words are not connected to any words in the lexicon. There is no natural progression from a receptive to a productive state in this view. This idea is linked to connectionism and measurement of the strength of productive mastery "would presumably require determining the relative number of links to other members in the lexicon" (Schmitt, 2010: 82).

Another word knowledge convention suggested by Anderson and Freebody (1981) is the distinction between breadth and depth. Breadth knowledge refers to the number of words known and depth knowledge to how well a word is known and therefore how well it can be used by the learner. But as Milton notes "simple binary divisions like receptive and productive, or breadth and depth do not really do justice to the complexity of word knowledge" (Milton, 2009: 14).

Nation attempts to describe the complexity of what it involves to know a word by dividing word knowledge into knowledge of form, knowledge of meaning and knowledge of use. He then subdivides each of these three categories into three subcategories. For instance, knowledge of form is further divided into spoken knowledge, written knowledge, and knowledge of word parts. Nation maps receptive and productive distinctions to the nine subdivisions, so that the form/spoken/receptive division aims at answering the question "What does the word sound like?" and the form/spoken/productive division investigates "How is a word pronounced?" (Nation, 2001: 17). Nation's dimension framework "is certainly not the only way of modelling lexical knowledge" (Milton and Fitzpatrick, 2014: 174) but it is a useful one to investigate part of lexical knowledge, in the case of this study, receptive and productive knowledge and to consider it in relation to overall proficiency.

Knowledge of word meaning also has three subcategories, of which the first is "form and meaning"; this corresponds to making the link between a word and its meaning or in a L2 learning context to link a word with an L1 translation for instance. The other two divisions of meaning are

"concepts and referents" and "associations" and indicate that from one language to another, both the meanings and the associations of apparently similar words can differ. Knowledge of word use has knowledge of "grammatical functions" as its first subcategory; this involves knowing which part of speech a word is, and how it combines with other speech parts. For instance, L2 French learners learn that most adjectives like *bleu* (blue) come after the noun as in *le ciel bleu* (the blue sky), and not in front of it as in English. The second division "collocations" implies knowing what other part of speech a word can be often found connected with. In English for instance, prepositions and certain verbs are often found together to make up phrasal verbs such as *keep up, keep off, keep out*. The last division of word use is "constraints on use" which corresponds to knowing where, when and how often one can use this word. Receptive and productive knowledge are mapped onto the subdivisions.

An additional important factor which is not actually part of word knowledge but rather control over it is fluency (Laufer and Goldstein, 2004: 401) which Schmitt (2010) refers to as adequate recognition/comprehension speed when reading or listening, and adequate retrieval/production speed when speaking or writing. In this study, we will concentrate on some aspects of vocabulary knowledge such as receptive and productive knowledge as well as fluency. Receptive knowledge will be tested using the X-Lex test, productive vocabulary knowledge will be tested using a test inspired by the Lex30 test and for fluency, the Elicited Imitation Test will be used. These instruments used to measure vocabulary knowledge and development will be detailed in the next chapter.

3.2.5 L1 and L2 mental lexicons: same or different?

Weinreich (1953) is one of the earliest to have proposed different possibilities regarding the storage of words in the multilingual lexicon (De Bot, Lowie and Verspoor, 2005: 43). According to him, words might be stored in three different ways: compound, referring to a multilingual lexicon with common concepts and different words in each language for these; co-ordinate, a lexicon that is completely separate with its own concepts and words for each language; and subordinate, where concepts are shared as in compound organization, but, the L2 item can only be reached through the L1, in other words there is no direct link between the concept and the L2 (De Bot et al, 2005: 43). Over past decades, these three possible multilingual organisations have been tested using different experimental tasks such as a translation task in which learners of different proficiency levels are asked to translate words from L1 to L2 and from L2 to L1. Kroll and Stewart (1994) found that the response time was slower for translations from L1 to L2 than from L2 to L1

and that the response time was even greater for less proficient learners (De Bot, Lowie and Verspoor, 2005: 44). These results led them to propose a model according to which the mental lexicon moves from a subordinate organisation to a compound organisation at more advanced language levels (De Bot, Lowie and Verspoor, 2005: 44).

If the idea of a lexical organisation that develops with increasing proficiency, as proposed by Weinreich, is in line with current thinking, the idea that there are separate lexicons is on the other hand more difficult to maintain. De Bot notes for instance that in addition to proficiency, the response time on translation tasks can be influenced by words' semantic characteristics too. For instance, a concrete word might be translated faster than an abstract one, as would cognates (De Bot et al, 2005: 44). Under these circumstances, De Bot proposes that we should view the mental lexicon as organised in a dynamic way, in which direct links between concepts on one hand and certain L2 words on the other, such as concrete words and cognates, would be established.

But the notion of activation of links between words and concepts also raises the question of whether a bilingual can "turn off" a language and function as a monolingual when listening, speaking, reading or writing (Tokowicz, 2015: 2). Since the 1990s, the research and the debates on selective versus non-selective access that examines whether or not bilinguals consider words from both languages when they are listening and speaking has been flourishing, leading to models of activation such as the Language Mode Hypothesis (Grosjean, 1985) or the Inhibitory Control Model (Tokowicz, 2015: 7). Green (1998) proposed that a language could be in one of three possible states: selected, when the language is being used; active when the language plays a role in the background; and dormant, when the language does not play any role (De Bot et al, 2005: 45). There is growing evidence and almost a consensus that multilinguals do not "switch off" the language that might be "dormant" but rather access all their languages when listening and speaking. In other words, methods such as lexical decision tasks (LDT) have shown that processing involves non-selective lexical access (De Bot et al, 2005: 45). This activation of lexical items across languages is the foundation of the Bilingual Interactive Activation (BIA) model that can account for code switching for instance. It has to be noted too that proficiency also plays a role in activation as L2 words at lower proficiency levels have usually "a lower level of activation due to lower frequencies and less interaction as the network of this language subset will be relatively small" (De Bot et al, 2005: 46).

3.2.6 Word association

The second major notion of a dynamic lexical processing model which we must consider is "association", which is concerned with the types of links in the mental lexicon that could explain its organisation.

Traditionally, research into L1 development has employed word association (WA) tests, because of the belief that "associative responses to stimulus words (SWs) reflect the functioning of thought processes of an individual" (Zareva, 2007: 124). Cramer (1968) believed that more complex processes of thinking could be better understood by exploring simpler units of thought and Deese (1965) took the idea further and stated that "researchers could get insights not only into the processes of thinking but also in to the way individuals construct their set of meanings" (Zareva, 2007: 125). Soon afterwards, word association tests were also adopted by L2 researchers.

In L1 and L2 word association tests, participants are given a stimulus word and then they are asked to produce the first word or words that come to their mind. The hypothesis is that "automatic responses which have not been thought out will consist of words which have the strongest connections with the stimulus word in a person's lexicon" (Schmitt, 2010: 59). Thanks to word association tests, linguists have been able to identify four particularly strong types of link in the mental web. They are, starting with the most common one: coordination (words which cluster together, like *salt* and *pepper*, in this case opposites), collocation (words that are likely to be found next to the stimulus word in connected speech, such as *salt water*), superordination (more general and more specific terms, for instance *colour* for *red*) and synonymy. Even in spite of recognising these different types of links, it is not possible yet to fully map the mental lexicon "since the structure of a group is likely to depend on the type of word involved - objects, colours and actions might be treated rather differently" (Aitchison, 2003: 90).

Word association responses can also be described qualitatively, that is to say according to what categories they belong to. Traditionally, there are three categories of associations: firstly, clang associations which are associations evoked by phonological similarity with the stimulus word rather than by meaning; Schmitt gives the example *reflect-effect* (Schmitt, 2010: 60). Secondly, syntagmatic associations are recognised, when the response is a word that collates well in a grammatical string and has semantic affinities. The response is usually from a different word class than the stimulus word. For instance, syntagmatic associations can be adjective-noun pairs such as *black-magic* or verb-adverb pairs like *walk-slowly* (Schmitt, 2010: 60). Finally, paradigmatic associations are semantic in nature; they often include synonyms, antonyms, and hyponymy etc.

and the responses are from the same word class, such as verb-verb pairs like *eat-drink* or noun-noun pairs like *house-home*. In WA tests, it has been shown that nouns elicited nouns in 80% of the time and verbs and adjectives elicited items from the same word class around 50% of the time (Aitchison, 2003: 105). Function words on the other hand seem to be divided from these content words but it is not clear how. The scope of the present study did not allow us to consider word associations for all the responses to the two productive tests but we considered word associations of participants' first answers to stimulus words. The results will be discussed in Chapter 5.

Researchers from the semantic field tradition, according to which "meanings of words must be understood in relation to other words that shape a given semantic domain" (Lehrer and Kittay, 1992) agree that the two main types of meaning relation are paradigmatic (synonyms, antonyms, hyponymy etc) and syntagmatic (words that collate well in a grammatical string and have semantic affinities) (Zareva, 2007: 126). However others (Ervin, 1961, Deese, 1965, and Tiffany 1972) adopt "a distinction between paradigmatic and syntagmatic relations based on lexical class rather than on semantics alone" (Zareva, 2007: 126). Also, the paradigmatic/syntagmatic distinction tends to show that it is difficult to distinguish between meaning and syntax. Next, we review some of the findings about these three types of associations.

The first conclusion that was quickly drawn from word association tests in L1 was that responses to a stimulus are not random and that participants tend to give systematically the same answers. For instance, *black* often elicits *white* and "the top three responses often account for half or more of the total number" of responses (Schmitt, 2010: 59). This is what Cramer refers to as response commonality, "used to refer to the frequency of occurrence of any three most commonly given associations to a stimulus word (SW) determined in terms of their absolute frequency of occurrence in a WA data set" (Zareva, 2007: 125). This means that the lexicon for all native speakers of a language is organised in a similar way and finding out about this organisation, some argue, will be beneficial for non-natives too and should be considered in L2 teaching and learning (Schmitt, 2010: 61).

Another characteristic of the L1 mental lexicon is that there is a shift in the word association categories depending on the age of the L1 speaker. Indeed, it has been widely proven in many different languages that the shift in responses goes from clang to syntagmatic to paradigmatic word associations. As opposed to adults, children produce many more clang associations, in other words, word similarity plays an important role in the first stages of L1 lexical organisation. But with age and more complex lexicon links, clang associations are less common and make way to syntagmatic responses which in turn make way to paradigmatic answers. With a preference

for paradigmatic responses in adulthood, the lexicon of mature L1 speakers seems to be organised in a more meaning-based way, as they sense the relation between words in terms of word classes and relations with other words (Schmitt, 2010: 61).

But according to Entwisle, Forsyth and Muuss (1964), the shift from syntagmatic to paradigmatic occurs at different points in time for different word classes (Schmitt, 2010: 61). The first word class to shift seems to be nouns, followed by adjectives and then verbs, and Schmitt also notes that a clang-syntagmatic-paradigmatic shift might not actually happen for every word but rather "indicates the general evolution of lexical organisation patterns as a learner's language matures" (Schmitt, 2010: 61). Also, despite the large number of research studies backing this shift, the validity of the syntagmatic-paradigmatic shift has been questioned in the light of research (Nissen and Henriksen, 2006) which has shown that native speakers actually seem to have a preference for syntagmatic responses.

As we have just seen, the study of the L1 mental lexicon has a long and strong history relying on fields such as psychology and using psycholinguistic tests such as word association. But research on the L2 mental lexicon has managed to adapt this methodology to investigate the organisation and the development of L2 learners' vocabulary in order to assess whether the L1 and the L2 lexicons are similar or not.

It was in 1982 that Meara first proposed to use word association methods to investigate the L2 lexicon. This was an important and innovative step in investigating what the L2 lexicon might look like and how it might differ from the L1 lexicon (Fitzpatrick and Barfield, 2009: 39). Meara's findings pointed towards the idea that L2 learners' associative patterning mirrors that of L1. Indeed, he and other researchers have found that L2 learners produce more clang and syntagmatic associations, like L1 speaking children, before moving on to more paradigmatic associations as their proficiency increases (Ervin, 1961).

Traditional word association research did not investigate the underlying psycholinguistic processes which led to the shifts in associative patterns which it described. To research form-meaning development in L2, from a psycholinguistic perspective, Meara and other psycholinguists have adapted Levelt's (1989) model of the lexical entry, represented in Figure 3.1 below

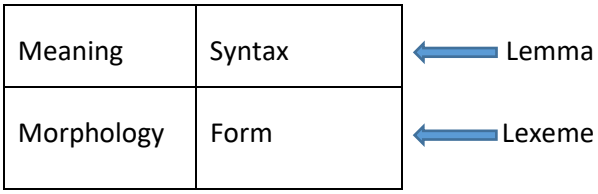


Figure 3.1: A lexical entry (Levelt, 1989: 188)

Using a variety of tasks such as "speak a word associated with a picture out loud; translate a word into, or from, another language; decision tasks of various kinds in which a participant in the experiment must rapidly decide whether a word is a possible word in a language or not" (Juffs, 2009:186), some conclusions about form-meaning acquisition were drawn. First of all, Meara (1997) and Schmitt (1997) demonstrated that L2 beginners tend to associate words with similar forms "even within the L2" (Juffs, 2009: 190), hence producing more clang associations as mentioned above. So it seems that form acquisition precedes the acquisition of lemma (semantic) acquisition in L2 too. With growing proficiency, the activation of form gives way to meaning activation, hence suggesting that L2 lexical processing does resemble L1 lexical processing as it develops.

The view that there is a shift from syntagmatic associations to paradigmatic ones in the L2 lexicon has dominated L2 word association work (Söderman, 1993). However, findings from more recent studies are also starting to question the existence of this shift in L2. For instance, Zareva's (2007) study of 87 adults both non-native and native speakers of English, all of them students at a US university, supports Söderman's view according to which "adult learners pattern their responses very much like adult native speakers do, even at the lower level of proficiency" that is to say with predominantly more paradigmatic associations (Zareva, 2007: 145). In Söderman's (1993) work, the shift of word associations seem to depend on the type of words too. Therefore, whilst Meara considers the L2 mental lexicon evolving globally towards the organization of a L1 mental lexicon with growing proficiency, for Söderman, every lexical item has "its own processing history, passing from a more phonological to a more semantic profile as it becomes more integrated into the internalized system" (Singleton, 1999: 136).

Wolter (2001) proposes a similarly dynamic account of the structure of the mental lexicon, i.e. the depth of individual word knowledge model, also known as the DIWK model. According to this model, the connections in both the L1 and L2 lexicon are driven by how well individual words are known to a speaker (Wolter, 2001: 46) rather than solely by general proficiency or word frequency. In this model, the traditional view of the lexicon as pre-existing structure in which

words would be slotted is refuted and instead, the DIWK deals with the connections between words a speaker knows. One assumption is therefore that the L1 and the L2 mental lexicons are different because the L2 lexicon contains fewer words than the L1 lexicon. Also for Wolter, the mental lexicon is unstable not only because new words are learned while others are forgotten but also because there are degrees of knowing a word. In other words as stated above, the strength and nature of its connections depends on the depth of knowledge of a word. At the centre of the mental lexicon, Wolter positions the core, made up of well-known words, as described by Nation, and successive layers are made of less well-known words. Associations within the core would therefore primarily, but not only, be paradigmatic and associations within the layers would be more syntagmatic and phonological.

An important finding related to proficiency is that the number of responses given to a stimulus in WA tests tends to increase with proficiency level, hence "confirming a relationship between the number of meaning connections and proficiency level" (Zareva, 2007). In the present study, we have not used a word association test but in the productive vocabulary test, as we will see in Chapter 5, we have also noticed that less proficient learners are not able to produce as many responses to stimulus words as more proficient participants. However, the types of responses given by L2 speakers are much more varied than the ones given to L1 speakers. Indeed, we have seen earlier that for L1, response commonality is high, that is to say that the three most common L1 responses to a WA tests are often the same and make up for a high proportion of the WA test's results. Considering our earlier example of *cat* as a WA test stimulus: the first three L1 responses to it are *dog*, *mouse* and *black,* and these account for 64% of the total responses. This lack of consistency or commonality in L2 answers compared to L1 answers made Meara (1982) state that the connections in the L2 learners' mental lexicon are less stable and that the words are more loosely connected with each other. For Meara (1996), "when L2 learners develop vocabularies that are sufficiently large in size, it is the organisation of that knowledge that becomes important" (Zareva, 2007: 126) and so measuring vocabulary size becomes less significant. An L2 intermediate vocabulary size of about 6000 words has "fewer links among words, a lower degree of commonality and lesser heterogeneity of meaning connections" (Zareva, 2007: 144). On the other hand, larger vocabularies of over 9000 words display more connectivity and more associations between words learners know, so larger vocabularies are richer in connections, both in size and commonality (ibid).

For Laufer (1997) who reviewed studies about the "intralexical difficulties" of words such as pronounceability, word length and part of speech, morphological and lexical complexities (that is

to say information relating to register and synonymy for instance, for the latter) are the most difficult aspects of a word to master. She concludes that "most of the problems in vocabulary learning pertain to meaning" (Singleton, 1999: 140). Sonaiya (1991) had already formulated a similar idea explaining that "the primary task in vocabulary acquisition is seen as one involving continuous refining of meaning and readjustment of boundaries between lexical items that have already been acquired and subsequent items that are encountered" (Singleton, 1999: 145). L1 and L2 lexical systems are communicating with each other to process the meaning of new lexical items but whether this is via direct connections between individual L1 and L2 lexical nodes or via a common conceptual store or both, remains unclear (Singleton, 1999: 189).

To sum up, L2 learners' lexicon tend to be smaller in size than the lexicon of native speakers. The organisation of the L2 lexicon, for some, mirrors the organisation of the L1 lexicon and the L2 lexicon becomes more nativelike with developing proficiency, as highlighted in the shift from form to meaning when learners go from clang, to syntagmatic to paradigmatic associations. This present research aims at determining the vocabulary size of first, second and final year undergraduate students of L2 French. It is a main aim of the study is to determine whether and how the lexicon size and organisation differ from one group to another, and also whether and how the lexicon size and organisation differ within each group.

3.2.7 Levelt's model of processing

There are, as we have seen, many different definitions of a word in second language acquisition research, and different views of how the L2 lexicon is structured, and the way words are processed and learned is even less consensual. However, Levelt's processing model (1989) that has already been mentioned has become the most accepted model for monolingual speakers (De Bot, Lowie and Verspoor, 2005: 39) and, as we will see in this section and the next, it is also relevant for L2 lexical processing. Subsequently De Bot adapted it to propose a bilingual or multilingual version that we will detail in the next section.

Levelt's speech production model first proposed in *Speaking. From Intention to Articulation* (1989) and successively detailed over the years proposes a two-step approach to lexical access, in the initial stages of speech production, i.e. once a message has been conceptualised. The first step consists in a lexical selection that aims at "retrieving the one appropriate word from among thousands of alternatives" (Levelt: 1993: 3). The second step is phonological encoding, that is to say "computing the phonetic shape from the selected item's phonological code or form specification as it is stored in the mental lexicon". (Levelt, 1993:3). In this model, syntactic

patterns are then determined by the mediating lexical items and their features (Levelt, 1989: 181); this is the "lexical hypothesis". "The lexical hypothesis entails, in particular, that nothing in the speaker's message will by itself trigger a syntactic form, such as a passive or a dative construction. There must be mediating lexical items, triggered by the message, which by their grammatical properties and their order of activation cause the Grammatical Encoder to generate a particular syntactic structure" (Levelt, 1989: 181). In other words, "the mental lexicon, mediates between conceptualisation and grammatical, morphological and phonological formulation" (Li Wei, 2006: 91) and as such plays a crucial part in the model. In Levelt's model and its adaption for bilingual speakers by De Bot, "the lexicon is assumed to contain declarative information, while the processing modules contain procedural information" (Schmid and Lowie, 2011: 270-271).

The Levelt model proposes a modular approach to speech production with three speech processors: a conceptualiser, a formulator, and an articulator. One of its advantages is that the model "permits an explanation for the speed of speech processing and it allows for the subdivision of speech production into clearly distinct and intuitively appealing stages and components" (Schmid and Lowie, 2011: 268). In productive speech, the speaker formulates communicative intentions or meaning intentions in the conceptualizer, that is the preverbal message (De Bot, Lowie and Verspoor, 2005: 41), "a message containing all the necessary information for converting meaning into language but that is not itself linguistic" (De Bot, Paribakht and Wesche, 1997: 312).

The formulator then interprets the preverbal message and converts it into lexical entries, themselves made up of two levels: the lemma that includes semantic and syntactic information, and the corresponding lexeme with morphological and phonological information. In the lemma, the lexical entry's meaning and syntax are represented, while morphological and phonological properties are represented in the lexeme (De Bot, 1992: 4).

The final output is a phonetic plan that can be executed by the articulatory motor system (Levelt, 1993:5). "For unilingual speakers there is substantial evidence to show that sounds are not the units of speech planning. It is more likely that speech is encoded and produced in larger units. Levelt assumes that syllables are the basic units of articulatory execution" (De Bot, 1992: 15). Speech error research has also shown that words need to be constructed over and over again and that their "phonetic form is not a ready-made template that can be retrieved as a whole" (Levelt, 1993: 9).

In speech, the correct lemma needs to be accessed, that is to say lexical access "involves essentially recognizing the most entailing predicates in the concept and finding the unique lemmas that have these as their core conditions" (Levelt, 1989: 214). But little is known about how the lemmas "become activated by fragments of the message" (Levelt, 1989: 198). However, errors in accessing the correct lemmas show that "the process of lexical retrieval in speech is parallel" (Levelt, 1989: 222). In word exchanges for instance, the two words are always of the same syntactic category suggesting that the developing surface structure creates particular syntactic requirements. (Levelt, 1989: 222).

In speech production, factors such as "visual effects, conceptual effects and word-frequency effects" (Levelt, 1989: 223) can have an effect on word retrieval. For instance, research have demonstrated the effect of priming on naming response speed. Levelt gives the example of the target word "church" that is elicited using a related prime word (house), an unrelated prime word (car), an identical prime (church) and a neutral prime (XXXX), that is the baseline. Retrieving the word is facilitated by the identical prime whereas the unrelated ones and the related ones slow down the naming. For Levelt, interferences of this sort happen at conceptual level (Levelt, 1989: 229). The lexical stage is also as mentioned sensitive to word-frequency (Levelt, 1989: 230). So a word that occurs often in language use is named faster than a word that is less frequently used.

So in Levelt's model, the mental lexicon contains all the information about words that the speaker needs for speech production. "This information involves at least, the meaning of each item and its syntactic, morphological and phonological properties" (Levelt, 1989: 233). And even though it is not a bilingual or multilingual model, De Bot (1992) and Li Wei (2006) believe that its "theoretical assumptions about the distinctive and connected components or levels of speech production are especially relevant to the study of the nature of the multilingual mental lexicon" (Li Wei, 2006: 90-91). As far as this research project is concerned, the central aspect of this processing model that is pertinent for the theoretical framework of this study is the lemma component that mediates between preverbal concepts that originate in the conceptualiser and word forms, or lexemes and its relation with proficiency.

3.2.8 De Bot's bilingual adaptation of Levelt's speaking model

Singleton noted that Levelt's model "seeks to address all aspects of language processing" (Singleton, 1999: 106) and in adapting this model for bilingual or multilingual speaker, De Bot places the speaker at the centre of the model. This approach fits well with the present research project too as it views the speaker, that is to say the L2 French learner, as a whole.

One of the first questions De Bot addresses is to determine what part of the system is responsible for the choice of the language in the speech process, to answer this point, he adopts Paradis's (1987) conclusion according to which there is no difference between registers in unilingual speakers and languages that multilingual speakers speak (De Bot, 1992: 7), hence placing the language decision in the conceptualiser. In other words, the preverbal message contains information about the register or the language and "subsequently plays a role in the selection of register-specific lexical items as well as the way in which these items are encoded" (De Bot, 1992: 7). To overcome the fact that in Levelt's model, the conceptualiser is actually language-specific, De Bot proposes that there are two processes in the conceptualiser, the macroplanning that is not language-specific and the microplanning that is language-specific (De Bot, 1992: 8).

The second point De Bot deals with is the question of how speakers access words from different languages and based on neurolinguistics research with bilinguals, he proposes that the Subset Hypothesis of storage of two languages in the brain is the most in line with the description of items in the mental lexicon in Levelt's model. This hypothesis "assumes the use of a single storage system where links between elements are strengthened through continued use" (De Bot, 1992: 11). Elements of the same language will be strongly linked to each other whereas elements from another language may form subsets. The Subset hypothesis "is closely related to current models of the lexicon which are based on *activation spreading* (De Bot, 1992:11) and overtime in bilingual speakers, "links between elements in different languages will be as strong as links between elements in one language". (De Bot, 1992: 11). The notion of activation implies that languages have different levels of activation and these levels "depend on the amount of contact, use, level of proficiency reached, maybe method of instruction, age of acquisition and many more variables" (De Bot, 2004: 26).

De Bot refers to Green's (1986) three possible levels of activation: selected, active and dormant. In the selected level of activation, "the selected language controls the speech output" (De bot, 1992: 13). An active language works in parallel with the selected language but "has no access to the outgoing speech channel" (De Bot, 1992: 13), that is to say "the phonetic plan of the active language is not fed into the articulator" (De Bot, 1992: 13). As De Bot notes, this parallel production can explain "phenomena associated with fluent and frequent code-switching" (De Bot, 1992: 13). Even though a dormant language does not play a role in speech production, it is still stored in the long-term memory (De Bot, 1992: 13). Therefore, De Bot proposes a bilingual model in which the mental lexicon will be the same for all languages and that the system of processing is language-specific with a formulator for each language.

For phonological encoding, De Bot proposes a unique articulator for all languages. In other words, rather than being stored separately, De Bot notes “that the existing collection of normalized sounds in L1 can be extended with additional sounds when a new language is acquired and it is not inconceivable that L1 norm apply to L2 as long as possible” (De Bot, 1992: 13). This point is supported by the fact that many proficient language users have issues mastering L2 pronunciation and intonation fully hence showing that it is unlikely that there are two language-dependent articulators. For De Bot “the quality of the L2 norm will depend on the frequency of use of the language, the amount and quality of language contact, and the extent to which subtle differences between L1 and L2 sounds can be perceived” (De Bot, 1992: 16).

To summarise, in the bilingual model, the conceptualizer is partly language-specific (micro-planning) and partly not (macro-planning). The mental lexicon stores all the words for all different languages. The formulators are language-specific and convert the preverbal message into a speech plan. Finally, the different formulators submit their speech plan to an articulator which is “not language specific and which stores the possible sounds and prosodic patterns of the languages”. (De Bot, 1992: 21) “The perception side of the model contains the mirror image of the Formulator, which is called the Speech Comprehension System (or Parser) and a mirror image of the Articulator called the Acoustic Phonetic Processor” as seen below in figure 3.2 (Schmid and Lowie, 2011: 270-271).

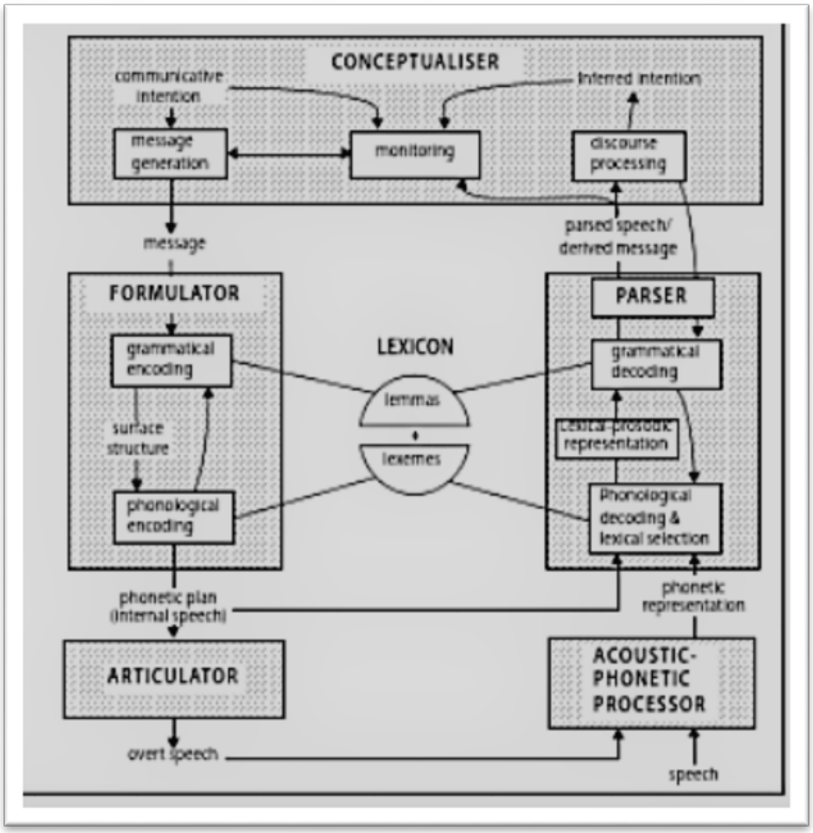


Figure 3.2 Adapted Levelt's (1989) blueprint of a speaker

These two models are interesting for the present research because according to them processing language also means learning and also because words are central to learning, as we have seen grammar is lexically driven, this is the lexical hypothesis. As processing is learning, the notion of frequency is also important is these models. One tends to learn high frequency words first because they occur more often in language use. So this study will enable us to find out if indeed more frequent words are known and the impact of the input on the lexical development of the participants and how this development influences general proficiency.

3.2.9 Lexicon development and language input

All language learning theories, in particular from usage-based and emergentist approaches, agree that input is the most important factor in learning a particular language. For De Bot, Lowie and Verspoor, "learning is essentially a matter of assimilation and accommodation" (2005: 52). The first mapping that occurs for a new word is between form and meaning at word and sentence level. Then, gradually, other cues about lexical items, such as agent identification, are learned based on their frequency of occurrence (De Bot et al, 2005: 56). In other words, the input will lead to learning, a process that can itself be divided into two main mechanisms. On one side, learning

is about incorporating new information into the system (De Bot et al, 2005:58), what Piaget called assimilation and defined as "the integration of external elements into evolving or completed structures of an organism" (Piaget, 1970: 706-707) and what Anderson (1983) referred to as consolidation in his theory of semantic networks. On the other hand, learning is also about matching a word's formal characteristics with semantic content, which Piaget refers to as accommodation, "any modification of an assimilation scheme or structure by the elements it assimilates" (De Bot et al, 2005: 57) and Anderson as semantization. Piaget's accommodation/assimilation and Anderson's semantization/consolidation are interrelated and "if a word is not adequately semanticized, consolidation cannot take place" (De Bot et al, 2005: 58).

In the context of Levelt's model, this means that when reading for instance, a learner needs to match a string of letters to a lexeme and if enough information is available, this activates a lemma which in turn activates a concept. Whether syntactic and semantic information about the lexical item are also internalised during this encounter is unclear. However de Bot, Paribakht and Wesche (1997: 316) suggest that in reading or listening, the information is parsed by the comprehension system sometimes bottom-up (letters or sounds) and sometimes top-down (knowledge of the world, the discourse setting etc). This implies that proficiency plays a key role as more proficient learners will have more syntactic and semantic knowledge that will in turn make the task of learning new words less difficult. It also implies that at lower proficiency levels, more reference to the L1 will be necessary until the L2 system is more developed. That again, implies that vocabulary size is an important factor for lexical processing and strengthens the point made earlier that proficiency and vocabulary size are very important factors for the organisation of the mental lexicon. In the next chapter, we will detail the instruments that were used to investigate the relationship between vocabulary size and proficiency, the receptive vocabulary test X-Lex and the Elicited Imitation test.

While there are few empirical findings on the uptake of syntactic and semantic information of new words when encountered for the first time, "the literature on effective and efficient vocabulary acquisition generally points to the necessity of association and activation with the terms elaboration and rehearsal" (De Bot, Lowie and Verspoor, 2005: 50). This means that on one hand, words should be encountered several times in different contexts so that a maximum number of associations are made for these new lexical items; this is elaboration. Trying to identify the type of input that can lead to the best L2 lexical processing, Hulstijn (2013) reviewed many studies and notes that processing of new lexical information that is done more elaborately leads "to higher retention than processing new lexical information less elaborately" (Hulstijn, 2013:

551). This suggests that learners should explicitly pay attention to word characteristics, "from orthography and prosody to the word's syntactic and semantic features" (De Bot, Lowie and Verspoor, 2005: 50). In addition, access to lexical items need to become more automatic and therefore needs to be rehearsed, i.e. encountered and used frequently. As far as L2 learning is concerned, there is general agreement that a great deal of input and opportunities to rehearse vocabulary are necessary. However, there is little empirical information about these opportunities to rehearse vocabulary amongst advanced L2 learners. This study therefore addresses this point by investigating the type and the frequency of input participants experience and how this might drive their L2 French vocabulary development.

However, the input is not a "one-way stream of information from the outside to the inside of a system" (Verspoor, Lowie, De Bot, 2008: 74) and the learner "is not a passive container that receives some input, but a very active problem solver who anticipates, predicts, and thereby interacts with the input continuously in a constructivist manner" (Verspoor, Lowie, De Bot, 2008: 78). In that sense, Verspoor, Lowie and De Bot propose that the term input should be replaced by processing and that processing as defined in Levelt's and De Bot's models can be equated to acquiring or learning (Verspoor, Lowie, De Bot, 2008: 73). Therefore, while the learner is processing input for meaning he or she may also be processing and acquiring related forms in the same input. It is very likely that this processing is supported by all the processing that the learner has done up to that point. In other words, the learning process is dynamic in that everything the learner has processed before will be available to support subsequent processing of new material (Verspoor, Lowie, De Bot, 2008: 78). The quality of processing depends on many individual factors such as age, proficiency, motivation, memory or learning strategies. In the present research project, the type of input accessed by students outside the classroom and its frequency are investigated through a questionnaire and the relationship between input, proficiency and vocabulary development in particular is considered.

Now that we have established the importance of input for acquisition through lexical processing, based on Levelt's model and also on the dynamic model of the lexicon, we examine more closely in the next section the nature of input in instructed settings, and the development of L2 vocabulary in such settings.

3.2.10 Which words to learn? The frequency factor

Given than languages are likely to contain more than a million words, it is fair to say that both native and non-native speakers will never learn all of them. According to Goulden, Nation and Read (1990: 341), "well-educated adult native speakers of English have a vocabulary of around 17,000 base words" and it is widely accepted that a frequency factor plays a key role in building both receptive and productive vocabulary knowledge. Indeed, Palmer noted long ago that "the more frequently used words will be the more easily learnt" (Palmer, 1917: 123). Words that are frequent in a language are highly likely to be learnt at an early stage, simply because they are encountered more often and each encounter leaves a trace which makes them easier to retrieve.

High frequency content words (such as nouns and verbs) are not semantically constrained like lower frequency words. That is to say that they can be used in a variety of contexts because there are no very specific connotations or collocations attached to them. Function/ grammar words such as prepositions, pronouns and determiners are also typically short, and high in frequency. Low-frequency words which do not recur often in the language may for example be proper nouns, or technical terms belonging to particular subject areas; they make up around 5% of an academic text for instance (Nation, 2010).

Lists of lexical frequencies have been created in various languages, but most commonly in English, through the analysis of language corpora. Such lists typically identify and rank the first 1,000 most frequent words found in the corpus, the first 2,000 words, and so on. Nation (2010: 15) concludes that learning "the 2,000 most frequent words of English …. is still the best decision for learners going on to academic study", as those 2,000 words generally cover 80% of a text. In instructed settings, such arguments have led to attempts to include target sets of high frequency words and other sets of words with desired characteristics in pedagogic materials, so as to maximise exposure to these words in input, hopefully leading to more efficient L2 processing and learning. Gathering data in many languages on word frequency, range and dispersion has been simplified by the building of numerous corpora and the development of computer software to analyse them (Laufer and Nation, 2012: 168). In this study, the word list used for investigating the development of L2 French advanced vocabulary is Baudot's (1992) list which we describe in more detail in the next chapter.

3.2.11 Word knowledge and input in instructed settings

Researchers interested in vocabulary learning in instructed environments, such as Laufer, Milton, Meara, Schmitt, Nation and Webb, are concerned with the contribution of incidental versus intentional learning to the acquisition of vocabulary from instructional input.

Before reviewing their findings, it is important to clarify what is understood by incidental and intentional learning. First, incidental vocabulary learning takes place when meaning-focused learners are engaged in a communicative task and acquire vocabulary incidentally during activities such as reading for pleasure. But incidental does not necessary mean without awareness. For instance, learners engaged in a reading for pleasure activity can decide to look up new words in the dictionary; in the classroom, a teacher can support incidental learning by explaining new vocabulary too. In other words, as Gass (1999) pointed out that incidental learning might more intentional than one would think. Ellis's (1999) definition of incidental vocabulary learning as learning words as a by-product of a task is now widely accepted. Intentional vocabulary learning on the other hand corresponds to focus on form or forms and it "may occur through pedagogical techniques such as input enhancement or corrective feedback" (Loewen, 2015: 100). In instructed L2 learning, because the input is often limited, most of the L2 vocabulary is deliberately learned (Cobb, 2007; Laufer, 2003, 2005; Webb, 2008).

As incidental learning is widely accepted to take place in L1, Nagy, Herman and Anderson (1985) proposed the Incidental Learning Hypothesis according to which, free reading is a major source of vocabulary acquisition during the school years. To be most effective, it should involve a lot of input, but also a good quality input in which new words are encountered more than once. Despite limited evidence for the incidental learning hypothesis (Webb and Nation, 2017: 54) and despite the fact that it is sometimes referred to as the Default Learning Hypothesis, because L1 learners do not learn vocabulary from lists, it has been widely accepted and in SLA, some like Krashen (2003) advocate that incidental learning through reading is the most efficient way to learn L2 vocabulary too. Given the restricted amount of classroom time, it seems wise to suggest that not all vocabulary acquired by L2 students can be taught in class and that therefore, they need to undertake some of this learning outside the classroom. This study is therefore interested in finding out how the development of L2 French vocabulary develops over time and relates to activities undertaken outside the classroom.

L2 vocabulary learning is indeed possible from incidental tasks as studies have shown that L2 students learn new words that are encountered repeatedly in different contexts (Horst, Cobb and

Meara, 1998). However, studies also show that the vocabulary gain from incidental learning is usually small, and always smaller than the gain from intentional learning (Laufer, 2003). Also because, as stated earlier, the input must be large and varied, "few L2 learning programmes are capable of providing the time and resources for learners to read this amount of text" (Webb and Nation, 2017: 51). Despite these drawbacks and the fact that incidental learning is time-consuming, Webb and Nation argue that it should be part of any L2 learning programme and should be combined with intentional learning because it is of real value for L2 vocabulary development. Certainly, one must be aware of the limitations of incidental learning and the conditions required for optimising its usefulness. However, we believe it is essential for it to remain a part of L2 vocabulary learning, even more so for advanced L2 learners who need to develop a large vocabulary size. This study also aims at shedding some light on the contribution of incidental learning over a period of three years of instruction. The longitudinal nature of this study adds to its potential contribution, because measuring vocabulary growth after only a short period "cannot be considered an accurate evaluation of incidental learning because it only measures one moment in a very long process" (Webb and Nation, 2017: 57).

While reading is often mentioned in studies of incidental learning, there is much less discussion of spoken input and its effect on L1 and L2 vocabulary, one possible reason being that it is usually believed that spoken input does not contain many low frequency words that could be unknown and therefore be learned incidentally (Webb and Nation, 2017: 52). However, recent studies have shown that watching TV programs and films could also be a useful resource for incidental learning as they can contain words beyond the 10K limit and also provide repetitions (Webb and Rogers, 2009). The research on vocabulary acquisition from watching videos is also encouraging and consistent in showing incidental learning (Neuman and Koskinen, 1992; Rice and Woddsmall, 1988). And even if speech offers fewer low frequency words than reading, Webb and Nation speculate that the gain might actually be similar in the long run as people spend on a daily basis a lot of time watching and listening (2017: 53), compared to reading. In any case, they believe that an extensive watching programme and an extensive reading programme are both necessary for L2 learners.

Let's now consider a little more the point about the frequency of words in reading. While incidental learning is likely to involve less frequent words, if there are too many such words, incidental learning will not happen for the reasons we have explained when discussing Levelt's processing model; there won't be enough information at the lexeme level to trigger a lemma. In fact, Liu and Nation (1985) state that "95% of the running words need to be already familiar to the

learners", that is, there should be no more than approximately one unknown word in every 20 running words, for learning to occur from reading input. Therefore learners' vocabulary size is crucial in incidental learning, and the learning burden of new words will decrease with larger vocabulary size as it will become easier to guess the unknown words (Webb and Paribakht, 2015; Liu and Nation, 1985). According to Webb and Chang (2015), a difference in vocabulary size of 500 words can lead to 30% more unknown words learned from context. This type of correlation is often referred to as Matthew's law: the more vocabulary one knows the more he or she will learn new words. This can also explain how more proficient learners are better at learning new words. It also shows that incidental reading does not mean any reading. Both teachers and learners need to be aware of their vocabulary size in order to select resources that will be appropriate and enhance learning in the long run.

Strategies to deal with unknown words in context and strategies to learn how to become an autonomous learner are important skills that need to be taught to all students (Webb and Nation, 2017: 145-146), but even more so to intermediate students and advanced students who need to find opportunities to access reading and listening or watching to develop beyond the 3K word limit for the former and the 8-9K word limit for the latter. And because correct guessing from context also correlates with vocabulary knowledge as we have just seen, but also with reading comprehension and verbal IQ (Hafner, 1967), it is important for learners to improve their reading skills (Nation, 2010: 245). In accordance with the dynamic model of processing, learners should also find opportunities to use vocabulary productively in order to make lexical processing more automatic. Another way of lowering the learning burden in incidental learning and building up syntactic and semantic knowledge is for example to read different articles on the same topics to increase the chances of encountering less frequent words more than once. Even though there are few empirical studies, Nation concludes that it is possible to add more and more information to words using the different contexts they are encountered in, in line with Levelt's model and the need to build the lemma and its syntactic and semantic information. For Nation, "learning from context is a cumulative process where meaning and knowledge of form are gradually enriched and strengthened" (2010: 236). Varied contexts are also important to create a range of instantiations that is to say a word's different meanings (Anderson and Shifrin, 1980). For instance, a learner might know the word *ball* but have different instantiations depending on whether one talks about a *golf ball* or a *child's ball* (Nation, 2010: 240).

There are fewer studies about L2 incidental vocabulary learning than about L1 but some useful ones shed some light on the rate of learning. For instance, Horst, Cobb and Meara (1998) gave

two kinds of vocabulary tests to their participants who had read a graded reader and found that one in five unknown words was learned to some degree (in Nation, 2010: 237). In terms of unknown words actually learned, this averaged at about five (ibid). But Laufer (2005) notes that number of words learned does not correlate with the length of the text. So for instance, just because 5 words are learned from a 200 word long text, this does not necessarily mean that 50 words will be learned from a 2000 word text.

As we can see, determining specific learning outcomes from incidental learning is not easy as "such learning seems to be determined by complex interactions of contextual, word textual factors" (De Bot et al, 1997: 314-315). In addition to these factors, we need to add the learner factor: incidental learning depends not only on the context and on the learner's vocabulary size as we have seen, but also on the learner's need. For instance, Laufer's Involvement Load Hypothesis states that in order for incidental vocabulary learning to happen, there are three factors that need to be present: need, search and evaluation. Need as in, is the word necessary to complete a task? As for search, learners need to figure out the meaning of the word, and evaluation entails comparing the word with other known words and deciding whether it fits the context. For Schmitt too, engagement is the optimum condition for vocabulary learning and under this umbrella, implicit, focus on form and explicit learning can all fit. Schmitt first proposed this term in 2008 in order to cover all the factors that facilitate vocabulary learning that have been developed in the literature (Schmitt, 2010: 28):

- Increased frequency of exposure
- Increased attention focused on lexical item
- Increased noticing of lexical item
- Increased intention to learn lexical item
- A requirement to learn lexical item (by teacher, test, syllabus)
- A need to learn/use lexical item (for a task or personal goal)
- Increased manipulation of lexical item and its properties
- Increased amount of time spent engaging with lexical item
- Amount of interaction spent on lexical item.

Taken together, Laufer's "Involvement Load Hypothesis" and Schmitt's "Engagement" both imply that learners' activity and motivation for L2 vocabulary learning are both essential for learning to be effective.

3.2.12 Conclusion

The mental lexicon is constantly reorganising itself as words are added to it and others are forgotten. Form-meaning is the first type of knowledge about words that learners process when encountering a new lexical item. In reading for instance, this means decoding the lexeme to create a lemma that will be linked to a concept. In the lemma, syntactic and semantic information about the lexical item will be built over time with repeated exposure to the words in different contexts. Some recent research studies validate the idea that the development of syntactic and semantic knowledge is gradual (Pigada and Schmitt, 2006; Webb, 2007), therefore supporting the idea that knowing a word is not actually the opposite of not knowing a word. Based on Levelt's processing model, activation and association are key concepts that explain the need to use lexical knowledge to make it more automatic and to encounter it regularly in different settings to strengthen the links with other parts of the lexicon.

The way the lexicon is organised is not fully clear because of its complexity and dynamic, changing character, leading De Bot, Lowie and Verspoor (2005: 50) to conclude that the mental lexicon is a chaotic affair. At the lexeme level, links can be phonological, for instance, words starting with the same letter can be associated together; at the lemma level, words can be linked either syntactically, for instance, word categories such as verbs can be associated, or semantically, through associations between words that belong to the same lexical field. As we have seen reliance on the different types of links is broadly related to different proficiency levels and to vocabulary size.

Based on this theoretical framework of the mental lexicon, Nation and other researchers interested in vocabulary learning in instructed learners propose certain pedagogical approaches that could lead to effective L2 vocabulary learning. For instance, they advocate the use of both explicit and implicit vocabulary learning and stress that incidental learning needs to take place through different tasks such as listening, reading and watching. Learners of different proficiency levels will have different needs, for instance, more links between the L1 and the L2 are necessary at lower proficiency levels. However, that does not exclude more proficient learners too as it has been found that both groups learn more vocabulary through translation tasks (Vidal, 2003). A lot of varied input is required to gradually learn more and more word characteristics. Finally learners should have the opportunity to reuse words regularly, so as to be able to access them more automatically. As we have seen, learners' proficiency, motivation and engagement are also important factors in lexical processing, and we examine these in following sections.

3.3 Proficiency

3.3.1 Introduction

Proficiency is a key concept for everyone involved in language learning, for learners first, whose aim is generally when learning a language to become proficient in the language, and for teachers whose knowledge of what might impede or facilitate proficiency will in turn help their students. Proficiency is a key factor in SLA research and "should be taken seriously because it is fundamental in the understanding of language acquisition and bilingualism" as stated by Hulstijn (2012: 429). If there is consensus about the importance of proficiency, defining it is not always easy, as we will see in this section. As one of the aims of the present study is to investigate the extent to which overall L2 proficiency can explain the pattern of L2 vocabulary development overtime, we will review both the construct of proficiency and proficiency assessment.

3.3.2 Proficiency in education and professional contexts

One approach to defining and assessing proficiency is to be found in educational and professional settings, where proficiency is often described as what an individual can do with the language in the real world. In this view, the components of language "are derived not so much from empirical research or theoretical models of language ability, but from the opinions of experts with first-hand experience teaching the local curricula and assessing student performance" (Purpura, 2008: 53). These experts are charged with the task to list the skills and the knowledge that are required to communicate effectively in the target language. Based on this information, curricula and assessments for teaching and evaluating those skills and this knowledge at different levels of proficiency are written. Some well-known teaching and assessing standards developed following this approach include the ESL Standards for Pre-K-12 Students (TESOL, 1997), a framework outlining "the goals, performance levels and descriptors of the behaviors underlying each standard and observable indicators of how progress on these standards might be assessed" (Purpura, 2008: 54). The Oral Proficiency Interview (2012) of the American Council on the Teaching of Foreign Languages (ACTFL) in the United States and the Common European Framework for Languages in Europe (Council of Europe, 2001, 2018) are other popular frameworks which aim at specifying what learners should do with the language at given levels and what they should do to progress from one level to another.

The Common European Framework for Languages (CEFR) is the proficiency framework which we will refer to the most, given that the setting of this research project is a UK higher institution. This framework was first piloted in 1996 before being published in 2001. The framework "organises language proficiency in six levels, from A1 to C2, which can be regrouped into three broad levels: Basic User, Independent User and Proficient User" (Council of Europe, 2018). The CEFR levels are used as basis for many language assessments in Europe, such as for instance, the DIALANG self-assessment tests (http://dialangweb.lancaster.ac.uk/), an online diagnosis tool that enables L2 learners to test their writing, reading, vocabulary and grammar in over 14 different languages and reports their level against the CEFR levels.

Testing learners' proficiency in the four skills (speaking, writing, listening and reading) is also important in educational settings. The emphasis of the CEFR and ACTFL frameworks is very much on language use, that is to say on what learners can do with the language in different communication situations, therefore testing within these frameworks is done as much as possible in situations that imitate the real world. For instance, ACTFL tests that are specially designed to measure oral proficiency, assess a "person's unrehearsed ability to speak a language in the real world regardless of where, when, or how the language was acquired" (Wu and Ortega, 2013: 681). The revised version of CEFR (2018) also emphasizes an "action-oriented approach" to language learning and teaching and presents the "language user/learner as a social agent, acting in the social world and exerting agency in the learning process" (CEFR, 2018: 25).

The theoretical model on which the CEFR scales are based, views the development of proficiency along two dimensions: the quantity dimension and the quality dimension. The first one is operationalised as "Communicative Activities" itself made of four types of activities (Reception, Production, Interaction, and Mediation). Each of these activities is then branched out within the CEFR. For instance, within "Interaction" a spoken and a written branch are distinguished. The spoken branch of "interaction" includes "understanding a native speaker, informal discussion, obtaining goods and services, conversation, formal discussion, interviewing and being interviewed". The second dimension, quality, is operationalised into "Communicative language competences" which include three types of competences: linguistic, sociolinguistic and pragmatic which are all seen as essential to effectiveness in language use. Linguistic competence is further developed into "range" (which includes vocabulary range) and "control" (Council of Europe, 2018). The elaborated CEFR theoretical model described above is based on multi-trait and communicative language models that have been popular in the last few decades and which we will explore in the next section.

Frameworks such as the CEFR and their descriptors of language proficiency levels are undeniably useful not only to measure students' learning outcomes but also to evaluate a program quality, or to accredit instructors for instance (Byrnes, 2006; Tarone, 2013). In educational and professional contexts proficiency has been defined as "what individuals can do with language in terms of speaking, writing, listening and reading in real-world situations in a spontaneous and non-rehearsed context" (ACTFL, 2012). However, this definition leaves many questions unanswered when investigating the development of L2 proficiency. For instance, is the development of proficiency a staged process? Do learners have to go through all the different stages in all the activities to improve their proficiency? Do all learners at the same proficiency level develop their L2 in a similar way? Also, what are the exact components of proficiency and what is the relationship between them? If those questions are relevant for both education and L2 research, in order to shed some light on them, one needs to turn to definitions of proficiency from an SLA perspective.

3.3.3 Proficiency in SLA

Some earlier definitions of proficiency in SLA include that of Thomas (1994) who defined it as "a person's overall competence and ability to perform in L2" (Thomas, 1994 in Wood Bowden, 2016: 649) and that of Higgs (1984) for whom proficiency was "the ability to function effectively in the language in real-life contexts". These two definitions seem relatively similar to the definitions of proficiency proposed by frameworks such as CEFR, themselves also grounded in sociolinguistic models of communicative competence (Canale and Swain, 1980 in Leclercq, Edmonds and Hilton, 2014: 6). However, in such definitions the concept of proficiency itself is not clearly defined (Hulstijn, 2010; Leclercq et al, 2014).

If many definitions of proficiency have been suggested in SLA over the years, none have been unanimously adopted, one of the reasons being that one's preferred definition depends on one's "particular theoretical stance on language acquisition" and on "what we believe knowing an L2 actually entails" (Leclercq et al, 2014: 5). If it is largely agreed now that "proficiency refers to a multicomponent phenomenon underlying one's knowledge of and ability to use a language" (Leclerq et al, 2014: 11), what these components are exactly and how they interact and develop still needs to be investigated.

One of the earliest models of language proficiency is Lado's (1961) which originates from behaviorism and American structuralism in linguistics (Hulstijn, 2010: 186; Purpura, 2008: 56).

This model consists of 16 cells, crossing four "language aspects" (phonology/orthography, morphology, syntax and lexicon) and four skills (listening, speaking, writing and reading). This "skills-and-elements" model of L2 proficiency paved the way to the teaching and testing of L2 for decades to come (Leclercq et al, 2014: 6). In Lado's model, language development can occur at a different pace in the four skills and as such, he "considers the degree of achievement in each of these separate skills as variable to be tested" (Leclercq et al, 2014: 6). In other words, "testers could design tasks to assess each discrete component (or element of L2 proficiency) through one or more skills" (Purpura, 2008: 56).

Following Lado, Carroll (1968) also proposed a "skills-and components" model, including phonology and orthography, grammar (morphology and syntax) and lexis as components and speaking, listening, reading and writing as the four skills (Purpura, 2008: 56). However for Carroll, in the assessment of proficiency, "discrete-point" tasks, measuring a single type of knowledge or abilities (Leclercq et al, 2014: 13) "needed to be complemented by integrative tasks, those which assess the capacity to use several components of language knowledge at the same time" (Purpura, 2008: 56). This suggestion to use both discrete-point and integrative tasks was very popular and became the basis for the first TOEFL design. In 1961, Carroll also included in his model of L2 competence the concept of "facility" in the use of language (Leclercq et al, 20014: 6), but neither Carroll nor Lado show skills happening in communicative situations in their models (Hulstijn, 2015: 37).

Carroll's call for integrative tasks led Oller (1979) to reject Lado's "skills-and-elements" model and to propose a general or global proficiency model instead. Oller's model was based on empirical evidence for strong correlations between "performance on cloze tests, reading, standardized discrete-point reading tests and measures of oral verbal ability" (Cummins, 2000: 119). This led him to propose language proficiency "could be conceptualized as a unitary global dimension that was largely indivisible from intelligence as measured by both verbal and non-verbal IQ tests" (Cummins, 2000: 119). Oller's proposal was also backed up by findings in psychometric research that showed evidence of a strong general proficiency and intelligence factor (Cummins, 2000: 120). Based on his findings, he developed the theoretical notion of a "pragmatic expectancy grammar" and used "pragmatic tests" (such as cloze tests and dictations) to elicit learners' performance. For Oller, an integrative task such as these was sufficient to elicit the language processing required in real-world situations and therefore, there was no need to recreate those real-world situations. However, in 1983, Oller acknowledged that the results of his analyses had been over-interpreted and that "the strongest form of the unitary hypothesis was wrong". He therefore proposed a revised and less strong global proficiency model, but yet again this was

rejected and a multicomponent approach to language ability was pursued over the next few decades.

Hulstijn (2015) notes that even if Oller's analysis of his results are psychometrically wrong, "the findings did not falsify his idea of a global factor of language proficiency, reflecting what he called expectancy grammar" (2015: 92). However for Gu (2014), one of the reasons why the five tests (composition, vocabulary, grammar, phonology and dictation) given to 164 non-native speakers of English correlated highly with a general factor was probably due to the fact that no speaking or listening tests were included. This highlights the issue of selection of test types when measuring L2 proficiency, a point that will be discussed further in this section and in the methodology chapter.

In 1972, Hymes's article "On communicative competence" widened the construct of L2 competence models which had up to then been influenced by Chomsky's (1965) "linguistic competence" and, which, had led many linguists to focus "quasi-exclusively on the knowledge of grammatical structures as an indication of the mastery of a language" (Leclercq et al, 2014: 6). Hymes called Chomsky's "narrow interpretation of the notion of "competence" into question, insisting on the pragmatic aspects of language use." (ibid), proposing "the notion of communicative competence, including not only knowledge of the language system but also knowledge of the appropriateness of language use depending on the communicative situation" (Hulstijn, 2015: 38).

Canale and Swain (1980) took up Hymes's idea and proposed a three component framework for communicative competence consisting of grammatical, sociolinguistic and strategic competence. Grammatical competence was defined as "knowledge of lexical items and of rules of morphology, syntax, sentence-grammar semantics and phonology" (1980: 29). Sociolinguistic competence was "made up of two sets of rules: sociocultural rules and rules of discourse" (1980: 30). Sociocultural rules include the appropriate use of propositions and communicative functions within a given sociocultural context whereas rules of discourse refer to *the cohesion* (grammatical links) and *coherence* (appropriate combination of communicative functions) of groups of utterances" (1980: 30).

Bachman and Palmer (1996) further developed Hymes's communicative model proposing "a three-level hierarchical model of language ability, distinguishing organizational language knowledge (grammatical and textual knowledge) and a component of strategic competence (metacognitive components and strategies)" (in Hulstijn, 2010: 186). Even though this model was only partly supported by data (Harley, Allen, Cummins and Swain 1990), Bachman and Palmer's

model is nowadays the main theoretical framework underlying communicative language teaching and testing (Purpura, 2008: 57), in the CEFR mentioned above.

Skehan (1989) proposed the components of complexity, accuracy and fluency (CAF) as developmental dimensions of L2 proficiency, and this framework has gained popularity in conceptualising advanced L2 proficiency in particular. But Housen and Kuiken (2009: 463) note that these notions are not straightforward L2 proficiency measurements. For instance, the notion of complexity can be applied to the task itself or to the language being used; in the latter case, "it remains unclear to what extent complex linguistic structures entail greater processing complexity" (Leclercq et al, 2014: 8). Accuracy, which is of course a concept used in language education, is more problematic as a proficiency measure because it entails comparing L2 performance to that of L1 users and as we will see when discussing Hulstijn's proficiency model, it is difficult to determine what is the norm for L1 speakers as they tend to vary also in their performance. As for fluency, it is not clear which dimensions of this concept are best included in CAF (Leclercq et al, 2014: 9)

Chapelle (1998) and Purpura (2004) have proposed other multicomponential L2 models, including grammatical knowledge, lexical knowledge, discourse knowledge, sociolinguistic knowledge and pragmatic knowledge. They aim at representing an overall framework while at the same time testing the different components to understand their interrelationship (Purpura, 2008: 53). However, if most researchers (e.g. Chalhoub-Deville, 1997, 2003; Kunnan, 1998) agree on the multicomponential nature of L2 ability, they have not yet reached an agreement on the specific components. "Some define them in terms of the formal elements of the language, others in terms of standards, and still others in terms of the language skills" (Purpura,2008: 55). There is no consensus on the way these components interact and how they could explain individual differences.

3.3.4 BLC-HLC Theory

Hulstijn (2011, 2012, 2015) has recently proposed a wider construct of language proficiency which aims at explaining the individual commonalities and differences among L1 and L2 learners. Indeed, how is it that L1 speakers who vary in their use of L1 still communicate effectively? Also, how come L2 learners can differ in their success in acquiring L2 even when they share the same motivation, and are exposed to the same L2 input over the same period of time? These are the two main questions his theory aims to answer.

In his language proficiency construct, Hulstijn makes the distinction between basic language cognition (BLC) and higher language cognition (HLC). BLC is what "all native speakers have in common: implicit knowledge of phonetics, prosody, phonology, morphology and syntax; explicit knowledge in the lexical domain (form-meaning mappings), and the automaticity with which these types of knowledge are processed" (Hulstijn, 2011: 230). It is "restricted to spoken and aural language (speaking and listening) and to highly frequent items and structures common to all adult speakers of their first language" (Wood Bowden, 2016: 649). The frequent lexical items and grammatical structures of BLC may occur "in any communicative situation, common to all adult L1ers, regardless of age, literacy, or educational level" (Hulstijn, 2011: 230). Hulstijn's distinction between BLC (restricted to speech reception and speech production) and HLC is in line with American and European structuralism regarding "the reception and production of speech as a more fundamental human attribute than literacy" (Hulstijn, 2011a: 231).

Higher language cognition is a complement of BLC, and it is identical to it except that HLC includes "lower-frequency lexical items or uncommon morphosyntactic structures that can be found in written or spoken utterances" (Hulstijn, 2011: 231). According to this construct, HLC is therefore the domain where individual differences can be observed, as BLC is what all native speakers have in common (Hulstijn, 2011: 230) and HLC depends on the learner's intellectual skills and level of education.

The notion of low and high levels of language ability is not new and the dichotomy of BLC and HLC is similar to the notions of basic interpersonal communicative skills (BICS) and cognitive academic language proficiency (CALP), proposed by Cummins (1980 in Hultstijn: 2011a: 232). Despite positioning himself in the emergentist school rather than the generative one, Hulstjin in describing BLC also agrees with Chomsky who also points out that "the phonological, syntactic, and lexical skills necessary to function in everyday interpersonal contexts are universal across native speakers" (Hulstijn, 2011: 233). However, the main difference between Cummins's BICS and Hulstijn's BLC is that the latter does explicitly explain what the components of BLC are.

With this model, Hulstijn (2015) hypotheses that under normal acoustic conditions, all adult L1ers can understand and produce, correctly and quickly isolated utterances made up of high-frequency lexical phrases and high-frequency morphosyntactic structures. The second hypothesis states that in BLC there will be hardly any individual differences and that most of the L1ers will perform very well in it whereas individual differences will be larger at HLC (Hulstijn, 2011: 232). But as far as L2 speakers are concerned, according to Hulstijn, while BLC can be attained in the domains of vocabulary and most grammatical structures, late L2 learners will not gain a native-like

pronunciation or “the production of some grammatical features in spontaneous, unmonitored speech”(Hulstijn, 2015: 50). They can however be as proficient in HLC as native speakers of the same “intellectual, educational, professional and cultural profile despite some deficiencies in their L2 BLC” (2011: 242).

Hulstijn agrees with Carroll (1981), Canale and Swain (1980) and Bachman and Palmer (1996) who proposed that language proficiency was more than linguistic competence and that it should include for instance pragmatic knowledge, sociolinguistic knowledge and knowledge of discourse organization. But Hulstijn proposes that L2 proficiency components are either core components or peripheral components. For him, empirical research consistently shows that measures of lexical and grammatical knowledge are associated with the four skills (Hulstijn, 2015: 41). Therefore for Hulstijn, linguistic knowledge and “the speed with which this knowledge is processed” are core components of his L2 proficiency model. By linguistic knowledge, he refers to vocabulary and grammar (as well as pronunciation), taking the usage-based approach which does not distinguish between lexis and grammar (Hulstijn, 2015: 43). Thus, the model is not particularly interested in finding out whether one or the other core components has more weight on language skills (speaking, listening, reading, writing) as it views vocabulary and grammar as one. In this view, BLC components can only be core, whereas HLC can be divided into core and peripheral components. Again, this echoes Chomsky’s ideal speaker-hearer’s intrinsic competence according to which all adult native speakers share the same core grammatical competence (Hulstijn, 2015: 55).

Thus in proposing this L2 proficiency model, Hulstijn’s aim is to tackle one of the most central and “puzzling phenomena” in language acquisition: individual differences. Why do L1ers with different experience of the language achieve similar knowledge? On the other hand, why do L2ers with similar input experience, level of motivation and length of study differ in their L2 acquisition? While the question is not a new one in language cognition studies, in suggesting the dichotomy of BLC and HLC, Hulstijn opens up a different theoretical framework to investigate it. In the present study, we adopt Hulstijn’s model and use an Elicited Imitation test to document overall proficiency and explore its potential relationship with lexical development.

3.3.5 Measuring proficiency

To understand how language proficiency develops in different contexts and to investigate individual differences in language acquisition, it is important to measure proficiency as precisely as possible. However, its multidimensional nature makes it methodologically difficult to test. As a

result, language proficiency assessment and the types of test employed have varied over the years as we will review below.

Bachman and Palmer, some of the pioneers of the componential model of language proficiency, designed a language proficiency test that reflected their three-trait language proficiency theory. Each of these three traits were tested with four test tasks: an oral interview, a writing task, a multiple-choice test and a self-rating task. The multidimensional LP definition was clearly the driving force behind their methodological choice. Using this "three-traits-by-four-methods" approach, (Hulstijn, 2015: 94) Bachman and Palmer attempted to measure the LP of 116 ESL students in the US, but the results did not validate their theory. Instead the results tended to show that the best fit for the model "consisted of a general factor and two specific factors: grammatical/pragmatic competence and sociolinguistic competence" (ibid).

Since Bachman and Palmer (1982) and Canale and Swain (1980) no serious attempt has been made to test their models or some modified version of their models. This is most probably because "it is almost impossible to find evidence for a trait model in a single correlational study" (Hulstijn,2015: 111). Instead, researchers have in majority focused on exploring the contribution of LP component skills to integrated language skills such as reading comprehension, composition writing, speaking proficiency or listening comprehension (ibid). However, this is a challenging undertaking; for example, a meta-analysis conducted by Jeon and Yamashita (2014) of 59 studies of components of L2 reading ability show that most of the studies failed to provide clear information about participants' levels of L2 proficiency.

Other reviews show that language proficiency assessment is not always included in SLA studies and that when it is, techniques of assessments can be very varied. For instance, Thomas (1994) reviewing 157 experimental or observational studies between 1988 and 1992 from 4 major journals, reports that four main proficiency assessment techniques were used: "impressionistic judgement (seemingly spontaneous, unsupported, characterization of learner's competence in L2); institutional status, learner's membership in a specific group (typically an academic course); in-house assessment, proficiency defined by locally-developed instruments and standardized tests (e.g. the Peabody Picture Vocabulary)" (Thomas, 2006: 283). With the substantial increase in the number of L2 proficiency studies, Thomas conducted another meta-analysis of 211 articles from 5 major journals from 2000-2004 and noted that the distribution of test types used was very similar to her first analysis. On the other hand, she notes that when proficiency is better reported, it is much more used. She also notes a trend for studies to downplay the importance of reporting empirically L2 proficiency. (Thomas, 2006: 292).

Tremblay's (2011) meta-analysis of 144 studies published between 2000 and 2008 from 3 journals conclude that only just over one third of those studies reported L2 learners' proficiency independently (and only two of these were concerned with L2 French).The most common method was one or more "original or simplified sections of existing standardized proficiency or placement tests (e.g., Greek Language Proficiency Test, Michigan Test, Oxford Proficiency Test etc), a cloze test or a C-test, and oral interviews or accent ratings" (Tremblay, 2011: 342). Studies that did not use these types of test, estimated proficiency based on L2 participants' year of instruction, or used proficiency scores from assessments such as TOEFL (ibid). Tremblay found that in the language background questionnaire of her 169 L2 learners of French, "of all the single underlying factor that was found to account for 52.9% of all the variance in the language background information, years of instruction in French yielded the strongest positive loading coefficient, followed by self-rating of proficiency, percent weekly use of French, and months spent in a French-speaking environment" (Tremblay, 2011: 347). However, even if years of instruction in the target language as well as other learning variables can correlate with proficiency, it seems reasonable that methods that are more robust should be used (Tremblay 2011).

Hulstijn (2012) considered 224 empirical papers published in "Bilingualism: Language and Cognition" from 1998 to 2011. Of those, 142 were concerned with "various types of group comparisons such as (i) comparison of bilinguals dominant in either two languages (ii) comparisons of bilinguals (languages A and B) with native speakers of either language (iii) comparison of L2 learners at different levels of L2 proficiency with native speakers of their L2" (Hulstijn, 2012: 423). 37 of these papers were published between 1998 and 2004 and in only 7 was language proficiency assessed with a test. Of the 103 published between 2005 and 2011, 56 used an objective language proficiency measurement. Echoing Thomas's finding that with time the importance of language proficiency is taken more systematically into account, Hulstijn still notes that in 28 (20%) out of 140 studies, language proficiency was not assessed at all (Hulstijn, 2012: 424). He concludes that the notion of language proficiency be it in a first language (L1) or second language (L2) is often taken for granted in the literature on bilingualism (2012: 423). Carlsen (2012) shares this point of view and notes that proficiency level remains a "fuzzy variable" in many studies.

For many years, scholars have been calling for using, whenever and wherever it is possible, measurement that can make comparisons between studies easier (at least between those that investigate the same target-language) (Tremblay, 2011: 343). In SLA, examples of careful screening of proficiency can be observed as in Tremblay's (2011) study of 169 L2 learners of French who were selected and assigned to three levels of L2 proficiency by means of (i) language-

history information, (ii) self-assessed L2 proficiency and (iii) an L2 cloze test (validated in another study) (Hulstijn, 2012: 425). A range of proficiency assessment tools, and proficiency-related tools, like questionnaires (e.g. the Language Experience and Proficiency questionnaire, LEA-Q), corpus analysis, and standardized tests, are now available to use and to share. Being able to compare the results of the present study with other studies was also one of the reasons why the French version of Ortega's original Elicited Imitation test, which is discussed below, was selected.

3.3.6 Elicited Imitation Test

Elicited imitation, as a research tool, has been used for the last 40 years in domains such as child language, neuropsychological research and second language research. In SLA, this technique that requires participants to listen to a stimulus and to then repeat it as closely as possible has been used since the 1970s to investigate different aspects of L2 language learning, including the assessment of implicit knowledge, morphology and listening comprehension (Kim, Tracy-Ventura and Jung, 2016: 657). In the field of L2 research, EI was initially used for assessing grammar proficiency (Vinther, 2002: 56). Naiman (1974) and Hameyer (1980) were among the first to suggest this technique as a measure of global proficiency that others then followed (Savignon, 1982; Scheibner-Herzig, Sauerbrey and Kokoschka, 1991). The theoretical rationale behind this proposal was that sentences could only be accurately repeated if learners have comprehended and parsed them through their developing grammars (Tracy-Ventura, McManus, Norris and Ortega, 2014: 146). In other words, accurate repetition is only possible if the listeners have enough lexical and grammatical knowledge about the target language (Gaillard and Tremblay, 2016: 427).

However Yan, Maeda, Jing and Ginther's (2016) meta-analysis showed that in over 44 years, less than two studies a year used EITs to measure proficiency, that is, 76 studies in all from 1970 to 2014. This number decreases to 21 if one takes into account studies that "met all the inclusion criteria for the meta-analysis which examined the construct-related validity of EITs" (Kim et al, 2016: 657). The two main criticisms that are often made against EIT is that, first of all, this method cannot be used to assess spontaneous language (Prutting, Gallagher and Mulac, 1975 in Vinther, 2002: 55) and secondly, that it does not tap into comprehension. For some (Munnich, Flynn and Martohardjono 1994 in Vinther, 2002: 58), Elicited Imitation is a mere imitation exercise in which participants only repeat strings of sounds without knowing the meaning of words they repeat (Vinther, 2002: 54).

Even if EITs do not create a normal communicative environment, however, Gallimore and Tharp (1981) found stable EI test-retest correlations over a period of years that is related to language behaviour in natural settings (in Vinther 2002: 55). Before then, Epstein (1978) found that the results of an L2 English EI test "for Danish schoolchildren correlated positively with the language teachers' assessment of these children's language proficiency" (in Vinther, 2002: 55). The second criticism about EITs, parroting, raises the question of the role played by working memory (WM) in EITs. If the stimulus is too short, it is feasible that it will be stored in the short-memory so "if they are short enough to be retained in immediate memory as an acoustic image" (Vinther, 2002: 58), it is possible for learners to repeat accurately sentences that they have not understood.

Therefore, in the debate of the role of working memory in EITs, length of sentences (or stimuli) is important and as Slobin and Welsh (1968, 1973) proposed, when the length of a sentence exceeds short-term memory capacity, participants will not be able to parrot it. Instead, the sentence will need to be decoded, recalled and reconstructed based on participants' grammar (Wu and Ortega, 2013: 683). Therefore, for EI to become a reliable tool, it is necessary to know how long a sentence should be in order to exceed working memory capacity (Vinther, 2002: 59). Over the years, proposals for a memory span limit have varied from Miller (1956)'s "magical number seven" (Vinther, 2002: 59) of chunks beyond which language users, even more advanced ones, are unable to process information to Simon (1974)'s five chunks (Vinther, 2002: 59). Estimates of the number of syllables necessary to exceed short-term memory capacity vary as well, and can go up to 15 for Naiman (1974). For Eisenstein and Madden (1982), Savignon (1982) and Kelch (1985 in Vinther, 2002: 61), it is 7, which seems to be the number of syllables that could go beyond short-term memory capacity and therefore, many recent EITs contain sentences with a minimum of 7 syllables.

There are other issues concerning EITs' design which have not yet been solved and that could potentially have an impact on findings, such as the question of whether or not the imitation should be immediate or on the other hand delayed. Based on Spitze and Fischer (1981)'s finding that "STM fades rapidly if not renewed" (Vinther, 2002: 59), some propose immediate repetition, but others like McDade (1982) note that if the stimuli have been understood, they will not fade away. Some calls have emerged in recent years for future research to investigate EI designs that include delayed repetition and its impact on phonological short-term memory (PSTM) (Kim et al, 2016: 667; Yan et al, 2016).

Despite the above questions and since Ortega (1999) started the revival of interest in EITs, many SLA researchers have used them and common findings about their usefulness and characteristics

as proficiency tests are emerging (Wood Bowden, 2016: 651-652). For instance, it seems that sentence length is indeed the main source of difficulty. It accounted for 73% of the variance in item difficulty in Graham, McGhee and Millard (2010), supporting similar findings from Ortega, Iwashita, Norris and Rabie (2002), Hendrickson, Aitken, McGhee and Johnson (2010) (in Tracy-Ventura, 2014: 146), Christensen, Hendrickson and Londsdale (2010) and Okura and Lonsdale (2012) (in Wu and Ortega, 2013: 683). In addition to sentence length, other characteristics contribute to item difficulty such as lexical frequency which according to Graham et al (2010) accounts for 8% of the variance in item difficulty. Investigating the question of item difficulty too, Erlam (2006) designed an EIT containing both grammatical and ungrammatical sentences as well as a comprehension test after each stimulus to check whether they tapped into form or meaning. Results show that there was a correlation between repeating accurately grammatical sentences and responses on the comprehension tasks. All of these results support Slobin's assumption that EITs are reconstructive in nature, that is to say that they require learners to process the stimuli and therefore these tests are not mere rote imitation.

In a 2016 study, Gaillard and Tremblay investigated phonological short –term memory (PSTM), the component of working memory which is the most widely researched, and came to the same conclusion about the importance of sentence length as a source of difficulty in EITs. They based their French EIT on Ortega (1999)'s 30-item Spanish test, with an additional 20 items, and claim that their EIT overall showed sufficient variability in length, in phonological complexity, in lexical complexity, morphological complexity and syntactic complexity so that the sentences were "representative of the French language as a whole" (2016: 426). The performance of their 94 L2 French participants again decreased as the sentence length increased. They also observed that EIT items discriminated well between "lower and higher level L2 learners" (2016: 441). Gaillard and Tremblay note that in other studies using different scoring systems to the one they used (Ortega, 1999), the level of difficulty of test items is very stable, "making the task a promising tool for the assessment of processing efficiency and thus linguistic proficiency in L2 learners" (Gaillard and Tremblay, 2016: 426).

3.3.7 Ortega's Elicited Imitation tests

As already mentioned, in recent years, more SLA studies have been using the same EIT design to measure L2 proficiency. This test was first designed by Ortega (1999) and then versions of it were created in other languages so that they "could facilitate global (oral) proficiency measurement and comparability across languages" (Wood Bowden, 2016: 652). The French version of this test is

the one that we have selected to test our participants' overall proficiency. The original test was designed to test L2 Spanish and consisted of 30 sentences ranging from 5 to 17 syllables. In 2002, Ortega et al created four EITs in English, German, Japanese and Spanish. The 30 grammatical sentences in each of these four languages included, as in the original test, a variety of grammatical structures at different levels of syntactic difficulty, and ranged from 7 to 19 syllables. A delay of 3 seconds is included between playing stimuli to participants and their repetitions. Participants' answers are scored on a 5 point-scale, which will be discussed in more detail in the methodology chapter.

In 2009, Zhou and Wu designed a Mandarin version of the test that was further developed by Wu and Ortega (2013). In 2014, Tracy-Ventura, McManus, Norris and Ortega created a French version, followed in 2016 by the Korean version developed by Kim, Tracy-Ventura and Jung. This EIT now available in 7 languages has demonstrated "a great amount of evidence for validity and reliability" (Kim et al, 2016: 656). The internal consistency scores of all these EITs for instance have been reported to range between 0.92 and 0.97, with 0.92 in French in Tracy-Ventura et al (2014: 155)'s study and the highest in Spanish and Chinese with 0.97 (Tracy-Ventura et al, 2014: 155). In reporting stable and consistent results such as internal consistency, the first aim of the shared EIT format, which was to "allow for crosslinguistic SLA comparisons and accumulation of interpretable findings across L2s" (Tracy-Ventura et al, 2014: 148) seems plausible. In the next section, more empirical findings yielded by these EITs are discussed.

3.3.8 EIT empirical findings

Despite the increasing use of EITs in SLA studies over the years, they have been little used in French apart from in Burger and Chretien (2001), Erlam and Loewen (2010) and Markman, Spilka and Tucker (1975) who all designed their own EITs (Tracy-Ventura et al, 2014: 148). Since the development of the French version of Ortega et al's (2002) EIT by Tracy-Ventura et al (2014), Gaillard and Tremblay (2016) have developed another French version in a study aiming to "provide evidence for the validity reliability, discriminability and practicality of an EIT developed for assessing the linguistic proficiency of L2 French learners" (Gaillard and Tremblay, 2016: 426). Tracy-Ventura et al (2014) and Gaillard and Tremblay's (2016) empirical findings will be discussed below. Due to the limited EI data available in French, data from recent studies using versions of Ortega's EIT in other languages will be referred to; they include Wu and Ortega (2013: Mandarin), Bowden (2016: Spanish) and Kim et al (2016: Korean).

In developing the French version of the EIT, Tracy-Ventura et al (2014) aimed at measuring the proficiency of L2 French learners while exploring the reliability and validity of this new test. Participants were 29 French-degree students (25 L1 English students and 4 L2 English speakers) who had just finished their second year at a university in the UK and 10 native speakers. Instruments included the 30 sentence EIT ranging from 7 to 19 syllables; a background questionnaire (including questions about participants' age, length of time studying L2 French, age of first exposure and other languages studied); a general oral interview; an oral retelling of a picture-based narrative; a written argumentative essay and a vocabulary recognition test. In addition, the end-of-year grades of each participant were collected. Using Ortega's scoring system, the EIT items were scored on a 0-4 point scale.

As expected, the French native speakers performed at ceiling level on this test. For the L2 group, positive correlations between EI scores and other measures of proficiency were found, the strongest being between EI scores and end-of-year exam grades (r= 0.78). As we will see in Chapter 5, a strong positive correlation between EI scores and end-of-year exam grades was also found in the present study (r= 0.64). However, the wide variety of EIT scores from this cohort does show too that, institutional levels are not the most reliable measurement of proficiency. Other positive correlations were found between EI scores and oral interview accuracy scores (r = 0.62) and oral picture-based narrative accuracy scores (r = 0.67). However, the vocabulary recognition X-Lex test did not correlate well with EIT scores. As already mentioned, and in line with other Ortega based EITs, the EI scores were highly reliable (α = 0.92). As for item difficulty, Tracy-Ventura et al noted that sentence length was an important variable but that other predictors of difficulty included morphosyntactic complexity, interaction between phonology and syntax, prosody and register. However, one of the limits of the study is that predictive validity of the EIT is limited because all the participants were from the same institution and the same institutional level. Tracy-Ventura et al call for repetition of this test but with distinct populations so that learners' comparisons can be made. By testing three different institutional levels, the present study aims to contribute to further validation of the EIT instrument in the measurement of proficiency development in L2 French. As we will see in Chapter 5, the EIT is a useful tool when used across different populations of learners as it is differentiates between low proficiency learners and high proficiency ones.

Unaware of Tracy-Ventura's French EI version, Gaillard and Tremblay (2016) used the one developed by Gaillard for her doctoral dissertation (2014). The aim of Gaillard and Tremblay (2016) was to assess the L2 French proficiency of 94 L2 French learners, of whom 74 had English as their L1 and the rest a variety of L1s. Participants were learners of French from beginners' level

to graduate level in the US. Six native French participants were also included to verify whether, as expected, the L1 baseline would be at ceiling level or not. As in Tracy-Ventura et al, Gaillard and Tremblay included a background questionnaire (to collect self-reported information about age, length of time studying French, age of first exposure, percentage of French weekly use); participants' self-evaluation of their proficiency in the four skills and Tremblay's (2011) cloze test. The EIT included 30 grammatical sentences well as 20 extra sentences including vocabulary and grammar structures drawn from a standard textbook. Sentence length ranged from 8 to 32 syllables. The cloze test was first given to participants, followed by the EI test which was divided into two sections and in between these two blocks, students completed the background questionnaire. The EIT was then scored on a 7-point scale assessing meaning, syntax, morphology, vocabulary and pronunciation. The scoring system in Gaillard and Tremblay's study is different from Tracy-Ventura et al's not only in terms of number of scale points, but also because Gaillard and Tremblay aim at scoring different components of proficiency whereas Tracy-Ventura et al aim at assessing global oral proficiency (following Ortega, 1999).

By performing a principal components analysis (PCA) relating the non-linguistic factors to the L2 French learners' averaged EIT scores, Gaillard and Tremblay found that class level, self-rated listening proficiency and self-rated speaking proficiency were the strongest predictors of proficiency (0.93), followed by cloze test scores (0.87), years of French instruction (0.76) and months of residence in a French-speaking environment (0.75).

Two findings also corroborate Tracy-Ventura's results, the first one being that EI scores are very varied within a group-level, therefore, institutional levels cannot be used straightforwardly as a measure of proficiency. Secondly, sentence length is the main source of item difficulty, with participants' EI scores decreasing as sentence length increases. Gaillard and Tremblay also noted that the EI test was discriminating well between lower and higher L2 level participants. Participants' institutional tests were not available for this study and therefore it was not possible to verify whether as in Tracy-Ventura et al., they correlated well with EI scores too.

Kim et al (2016)'s Korean EIT is one of the latest additions to the growing number of EIT available on IRIS (https://www.iris-database.org). In addition to evaluating the validity of the Korean EI test, Kim et al also investigated the influence of phonological short-term memory on EI test performance. 66 L2 Korean learners living in Korea took part, divided among beginners, intermediate or advanced learners. The EIT was designed as closely as possible to the English version and was made up of 30 sentences, ranging from 7 to 19 syllables. Participants' responses were scored following Ortega et al (2002)'s procedure. A standardised listening test was also

included, the listening section of the TOPIK (internationalised standardized Korean proficiency test). Kim et al (2016) also compared the EIT scores and the TOPIK scores to CAF measures. The participants also did a digit span test as a measure of phonological working memory.

The first key finding from Kim et al is that once again the analysis of internal consistency found high reliability (=.96). Using Ebel (1979)'s guidelines, Kim et al also concluded that 29 out of the 30 items in the EIT were very good at discriminating between high and low-scoring participants. Correlation was strong between EI scores and speaking scores (r=.77) and between EI and TOPIK listening scores (r=.62). As for CAF dimensions, the strongest correlation was found between EIT and fluency (r= .62), the number of morphemes per clause (r=.55), the proportion of accurate clauses (.47) and the number of clauses produced (r=.40). This leads Kim and al to conclude that EIT does indeed measure the integrative use of oral language. However, no correlation was found between EI scores and digit span scores, as Okura and Lonsdale (2012) had found too, thus demonstrating that PSTM does not play an important part in EIT performance. Overall, Kim et al validate a number of findings from previous research: the high reliability of the EIT and the fact that working memory is not an important predictor for EIT scores, which supports the claim that the test taps into language processing rather than being a parroting test. It also supports the point that the most important item difficulty factor remains sentence length.

Wood Bowden (2016) explores to what extent the EIT can discriminate amongst L2 learners from a wide range of levels, hence answering Tracy-Ventura et al (2014)'s call for more comparisons between learners. The study included 37 English native speakers who are L2 Spanish learners from three different levels: low-experience L2 speakers, advanced-experience L2 speakers and very-advanced L2 speakers. The EIT is based on Ortega (1999)'s original test, itself in Spanish, to which Wood Bowden made a few changes, with Ortega's approval. The EI scoring again followed Ortega's procedures. The SOPI, an oral proficiency test in Spanish designed at the Centre for Applied Linguistics (CAL) in Washington DC, was also used. This test is meant to tap into both BLC and HLC and follows ACTFL speaking proficiency guidelines. A background questionnaire was included (to collect information about age of exposure, months of immersion, years of formal Spanish and self-ratings of proficiency).

Key findings from Wood Bowden that support other EIT studies are, once again a high level of internal consistency (.98) and the fact that it is a tool that discriminates well between low L2 learners and advanced learners. However, the study also seems to suggest that the test is not so useful at discriminating between advanced and very advanced learners. According to the analysis, three subgroups can be defined by EIT scores. The external validity of the test is also shown as EIT

scores correlate very highly with SOPI scores (r=91). Out of the four self-rated proficiency strands, speaking is the one that correlates the best with the SOPI (r=.75) and EIT (r=.73), a similar finding to Kim et al (2016) who also reported a strong correlation between speaking and EIT scores (r= 77).

Overall it seems that recent studies using similar EITs in different languages yield some very consistent findings. First, the EIT appears both internally and externally consistent. Second, the item difficulty factor that seems to have the strongest effect on EIT scores is sentence length. Third, EITs are about language processing and not parroting, as there are no consistent correlations between working memory scores and EIT results. Mounting evidence will keep strengthening the validity of the EIT as a measure of global oral proficiency and will "aid readers of research when deciding the extent to which findings can be generalized to other samples and populations" (Norris and Ortega, 2012, in Tracy-Ventura et al, 2014: 144). Given the wide variety of languages in which the test is available and the practicality of the test, which can be administrated in 10 minutes and scored in 15, this is a likely target.Using EIT (along with a language learning background questionnaire as well as receptive and productive vocabulary tests), the present study thus aims at exploring L2 French learners' overall proficiency and to determine whether they are any relationships between EIT scores and vocabulary development scores. We will now review motivation, the other main individual difference this study focuses on.

3.4 Motivation

3.4.1 Individual differences

Individual differences in language learning and in particular in vocabulary learning are a complex and yet unsolved issue as Milton states: "We understand something about how groups of learners grow vocabulary in a foreign language, we really do not understand and cannot yet predict how individuals will behave" (2009: 243). Individual differences have been listed and categorised in different ways in SLA. However, age, intelligence, language aptitude and motivation are regularly referred to as major individual differences (Dörnyei, 2005; Ellis, 2008) whereas others like willingness to communicate and learners' beliefs are less frequently listed (Loewen, 2015: 162). Due to the nature of our cohort, individual differences such as age and ability are not so relevant. Indeed, our target cohort is quite homogeneous concerning age (all are young adults), and also

concerning academic ability/ language aptitude (all of them have been very successful in formal education, achieving high grades and choosing to study languages at university).

On the other hand, motivation is a pertinent individual difference for vocabulary learning at all levels. Indeed, as Williams and Burden suggested "understanding students' reasons for studying a language is important in order to understand students' sustained involvement during the course" (Williams and Burden 1997 in Busse and Williams, 2010: 68) and sustained involvement is in turn essential for L2 vocabulary learning. In this section, we will briefly review motivational theory development from Gardner to nowadays before looking more closely at the model we have chosen for this study, namely the L2 Self Motivation model, its relevance in understanding individual differences in L2 learning and in particular in L2 learning of LOTEs (languages other than English) in higher education in the UK.

3.4.2 Motivation theories

3.4.2.1 From the social psychological period to the process-orientated period

The modern and rich L2 motivation field we know originates from a social psychological perspective and owes a lot to two social psychologists in particular, Wallace Lambert and Robert Gardner who worked "in the bilingual social context of Canada" (Dörnyei and Ushioda, 2011: 41). A principal key to their motivation theory is that an individual's L2 learning behaviour is influenced by his/her "attitudes towards the L2 and the L2 community, as well as their ethnocentric orientation" (Dörnyei and Ushioda, 2011: 41). According to Gardner (1985), a truly motivated individual displays the three components of L2 motivation which are motivational intensity or effort, desire to learn the language and positive attitudes towards learning the language. Two concepts in particular from Gardner's motivation theory have attracted a lot of attention over the years and have been extensively studied; they are integrative orientation and instrumental orientation. The former relates to the "willingness to be like valued members of the language community" (Gardner and Lambert, 1959: 271) and the latter to "the potential pragmatic gains of L2 proficiency, such as getting a better job or a higher salary" (Dörnyei and Ushioda, 2011: 41).

But by the late 1980s and early 1990s, issues with Gardner's motivation theory started to surface and there were calls to find alternatives to it. One of the criticisms voiced by Crookes and Schmidt (1991) concerning social psychological theory was that it had little practical relevance for

teachers, and they therefore called for “a more practitioner-validated concept of motivation shaped by insights from motivation research in education” (Dörnyei and Ushioda, 2011: 47). The second main criticism against Gardnerian theory was that it did not reflect the cognitive revolution that was taking place in mainstream motivational psychology at the time. A growing gap between L2 motivation theories and these emerging new concepts in motivational psychology was highlighted by Oxford and Shearin (1994, 1996). Under these pressures, the social psychological period moved onto a period that Dörnyei (2005) called cognitive-situated which was characterised by motivational theoretical frameworks becoming more complex and grounded in the classroom environment. The social psychological framework was not completely wiped out though but rather extended.

In the 1990s, while classrooms became more common settings to investigate motivation, two significant new theoretical frameworks of L2 motivation were developed. Firstly, Dörnyei’s (1994a, 1994b) three-level framework of L2 motivation conceptualised L2 motivation within language level, learner level and learning situation level. The *language level* relates to “aspects of the L2, such as the culture and the community, as well as the intellectual and pragmatic values and benefits associated with it” (Dörnyei and Ushioda, 2011: 47). The *learner level* refers to the learner’s individual characteristics during the learning process and the *learning situation level* refers to “situation-specific motives rooted in various aspects of language learning within a classroom setting” (ibid). This last level comprises *course-specific motivational components*, *teacher-specific motivational components* and *group-specific motivational components.*

The second framework was developed by Williams and Burden (1997), drawing on a wide range of motivational theories from mainstream psychology, rather than L2 motivation ones. The framework is organised to cover both internal factors and external factors. The former includes factors such as intrinsic interest of activity, or perceived value of activity, for example; and the latter, significant others (parents, teachers, peers), the nature of interaction with significant others and the learning environment (Dörnyei, 1998: 126). It is also during this period that autonomy in language education started to attract a lot of attention. This new focus and the increasing field of task-based research, in turn, lead to a focus on motivational self-regulation and a move from the cognitive-situated period to a more process-orientated period during which the temporal aspects of motivation were analysed, both at micro level, for instance during a task, and macro level, during the course of a programme of study (Dörnyei and Ushioda, 2011: 60).

Whereas in the social psychological and in the cognitive-situated periods, quantitative approaches were used to build an overall picture of motivation, in the process-orientated period, in which

"motivation is viewed as dynamic, fluid and changing" (Loewen, 2015: 164), the methodology had to be adapted to analyse the dynamics of L2 motivational change. Questionnaire-type instruments were used in many studies that aimed at investigating changes either over the complete course of study or over several years of learning, including in Chambers (1999), Gardner, Masgoret, Tennant and Mihic (2004) and Ushioda (1998, 2001). The latter for instance, conducted a longitudinal study of 20 Irish learners of French and found that positive learning history had the most impact on 16 out of her 20 participants whereas future goal-orientation was "more appropriately conceived as a potentially evolving dimension of language learning motivation rather than its necessary rationale" (Ushioda, 1998: 81-82). This finding supports Dunkel and Anthis (2001) and Waterman (1982) for whom future self-beliefs are particularly dynamic between the ages of 18 and 22 (Busse, 2013: 392).

Furthermore, a "gradual convergence of self-theories and motivation theories in mainstream psychology" has resulted "in the introduction of a number of self-specific mechanisms that link the self with action" (Dörnyei and Ushioda, 2011: 80). From this perspective of self-theory, some possible selves that represent "the individual's ideas of what they might become, what they would like to become, and what they are afraid of becoming" (Markus and Nurius, 1986 in Dörnyei and Ushioda, 2011: 80) are important in academic achievement.

3.4.2.2 The L2 Motivational Self System

In the early 2000s, as English was clearly becoming the dominant international language, dissatisfaction grew even more with the Gardnerian concept of integrativeness which was suited for very specific bilingual contexts; after all, out of 75 studies conducted using the Gardner's Attitude/Motivation Test Battery, only nine took place outside Canada. It is in this climate and with the increasing importance of self-theories in mainstream psychology, and drawing in particular on possible selves theory (Markus and Nurius, 1986), that Dörnyei (2005) proposed the L2 Motivational Self System that was made up of three components: *Ideal L2 Self, Ought-to Self and L2 Learning Experience*. The two first components draw on Self theories. *Ideal L2 Self* "is a powerful motivator to learn the L2 because of the desire to reduce the discrepancy between our actual and ideal selves" (Dörnyei and Ushioda, 2011: 86); it represents a self who would have mastery of the L2. The *Ought-to Self* refers to "the attributes that one believes one ought to possess to meet expectations and to avoid possible negative outcomes" (ibid). *L2 Learning Experience* refers to the learning environment and experience and the possible impact of the peer

group, the teacher or the curriculum (Oakes, 2013: 180). This last component that does not originate from self theories, provides evidence that increasing classroom research that started during the cognitive-situated period made a lasting impact on the field of L2 motivation.

Dörnyei (2009) renewed the model following the well-known large scale longitudinal research investigation, in fact the largest L2 motivation study ever (Taguchi, Magid and Papi, 2009: 66), which he and Csizér conducted in Hungary involving over 13,000 students over a period of 12 years and that focused on the participants' attitudes to English, German, French, Italian and Russian. Since then, many studies were conducted to test and validate the L2 Motivational Self System in different learning environments. All of them show that *Ideal Self* and Integrativeness are indeed closely related, hence validating the L2 Motivational Self System, in which the *Ideal Self* offers a helpful reformulation of integrativeness (Csizér and Kormos, 2009; Taguchi et al., 2009). Taguchi et al (2009: 88) conclude that "our findings support the underlying tenet of the L2 Motivational Self System that integrativeness can be relabelled as the Ideal L2 Self". This motivational model is the one used in this present research and details of the instruments adopted and adapted from Taguchi et al (2009) will be discussed in the methodology chapter.

Thus, the three sources of motivation to learn a second language are the "learner's internal desire to become an effective L2 user; social pressures to become an effective L2 user" and "the actual experience of being engaged in the L2 learning process". (Dörnyei and Al-Hoorie, 2017: 457). But Dörnyei (2009) proposes a key condition to the model, "namely that the future self-image should be in harmony, or at least not clash, with the expectations of the learner's family, peers and other elements of the social environment" (Dörnyei and Al-Hoorie, 2017: 460). In other words, clashes with the image of Self can be counterproductive. Another point about future Self visions was raised by Markus and Nurius (1986) who noted that for a possible Self to trigger a motivational potential, the vision itself must be quite detailed. This means that the individual's imagination plays an important part in possible self theories and as such, some are even suggesting ways to use mental imagery in the L2 classroom to support motivation of learners (Busse, 2013: 380).

As mentioned, many studies (Schmidt and Watanabe, 2001; Masgoret and Gardner 2003; Csizér and Dörnyei, 2005; Bernaus and Gardner, 2008) have shown the positive effect of motivation for L2 learning and the development of overall proficiency. On the other hand, the relation between motivation and vocabulary learning in particular has been less systematically studied. Tseng and Schmitt (2008) believe that the dynamic and multidimensional nature of motivation make it indeed relevant to vocabulary study. They developed a model of motivated vocabulary learning, through structural equation modelling (SEM) "a modern multivariate statistical technique that

allows a set of relationships to be examined simultaneously" (Tseng and Schmitt, 2008: 360). The elements they selected for the model were drawn from their understanding of the literature and some previous findings from their studies of motivation and vocabulary learning. These six elements are (Tseng and Schmitt, 2008: 360):

1. Initial appraisal of vocabulary learning experience

2. Self-regulating capacity in vocabulary learning

3. Strategic vocabulary learning involvement

4. Mastery of vocabulary learning tactics

5. Vocabulary knowledge

6. Postappraisal of vocabulary learning tactics

They examined the interrelationship of vocabulary knowledge and motivation amongst 49 university students from a Taiwanese university and 210 students from a Chinese university and concluded that learners' view of themselves as "proactive agents in vocabulary learning" (Tseng and Schmitt, 2008: 389) was enhanced by "conceptualising motivated vocabulary learning". Learners understand the need to proactively exercise control over their vocabulary and this has a positive effect on vocabulary learning. The present study takes account of these proposals in a different educational setting, and the relationship between vocabulary learning/motivation and proficiency/motivation among L2 French learners in a UK university will be investigating referring to the L2 Motivational Self System.

3.4.2.3 The motivational Self System and LOTEs

L2 motivation is a popular research subject that "far outweighs in volume the amount of research published in relation to other L2 learner characteristics and many areas of inquiry within second language acquisition" (Ushioda and Dörnyei, 2017: 451). And even if second language acquisition (SLA) is concerned with the learning of all second languages, it is increasingly evident that there is an imbalance in the target languages researched. As far as the field of motivation is concerned, "over 70% of the motivation surveyed during this decade were set in contexts where the focus was on learning English" (Ushioda and Dörnyei, 2017: 452). In this context, a special issue of *The Modern Languages Journal* (2017) was dedicated to motivation in which many voiced their concern about the impact which this bias could have on theorising motivation. In this publication,

Ushioda, Dörnyei and others call for more research on language learning motivation beyond "the lingua franca of the *global village*" (Ushioda and Dörnyei, 2017: 452) and invite researchers to consider critically the relevance of the current mainstream theoretical perspectives for motivation to learn languages other than English (LOTEs). Another key point is that it is important to investigate the impact global English has on motivation to learn another language. By investigating the motivation of L2 French learners in the UK, the present study contributes to gather data about motivation in learning LOTEs.

Dörnyei and Al-Hoorie (2017: 456) point out several reasons why a focus on LOTEs might be a "valid and fruitful research direction to enrich the notion of L2 motivation" (2017: 456), the first one being because the current theoretical framework might not be adapted to shed light on certain important features of motivation for learning LOTEs as opposed to learning English. For instance, as seen earlier, the notion of Ideal Self replaced Integrativeness because it was found that "identification with an external group such as the L2 community" was not particularly useful when researching L2 English learning outside of Canada. But LOTEs on the other hand can perhaps be more easily associated with a specific community. Therefore, it might be beneficial when investigating the motivation of LOTEs learning to reconsider "the attitudinal link between the learner and the respective community" (Dörnyei and Al-Hoorie, 2017: 459). In other words, while integrativeness was once problematic because too broad to differentiate between its strong form, that is to say the desire to integrate, and its weak form, to be interested in the L2 culture and language (Oakes, 2013: 179), both these forms could be potentially important in explaining the motivation to learn LOTEs.

The other Self that might need to be considered differently when focusing on studying motivation in learning LOTEs is the *Ought-to Self.* Indeed, whereas it is usually an important factor of motivation to learn English, the preeminent global and international language, learning a language other than English may not so much be regarded as useful. In this context, Dörnyei and Al-Hoorie go as far as wondering "whether it still makes sense to speak about the Ought-to L2 Self in the case of LOTE learning" (2017: 460). Others suggest that the predominance of the English language might actually push some individuals to learn other languages. Thompson and Vásquez (in Dörnyei and Al-Hoorie, 2017: 461) called this type of motivation anti-Ought-to Self. Other case studies seem to point towards an anti-Ought-to Self or a rebellious profile as described by Lanvers (2017) who found that some individuals are drawn to learning LOTEs because they refuse to comply with what is expected from them, that is to say they are rebelling against learning English because it goes against their own ideal L2 self. For Oakes (2013) who researched university

learners of LOTEs in the UK, one of the motivating factor is also of defiance, such as the defiance of the common belief that the British are mostly monoglote English speakers.

With a motivation questionnaire including closed and open-ended questions, the research project aims at investigating students' reasons to learn French at university at gathering useful examples of individual reasons why they have elected this subject as "LOTE learners will have particularly unique and often unusual stories to tell" (Dörnyei and Al-Hoorie, 2017: 462). This information is crucial as Dörnyei and Al-Hoorie note that "long-term LOTE learning may only be successful if students have had the opportunity to construct aspects of their narrative identities that support their unique language-learning enterprise" (2017: 462).

Another reason for reframing of L2 motivation when LOTEs are involved is that they are often learned alongside another foreign language which is usually English. Therefore the question arises of knowing what effect the prevalence of English has on the motivation for learning another language. But the context of our study does not permit us to shed any light on the question. On the other hand, the linguistic landscape (L1 English) and political situation of the UK (not fostering multilingualism, as seen in chapter 2), makes it a particularly interesting setting to investigate what drives L1 English speakers who live in a society where monolingualism is common to study other languages.

3.4.2.4 Motivation for learning modern foreign languages in UK universities

With the vote for Brexit in June 2016, the UK is living a significant and transitional period that will impact many sectors in the country in ways that are not yet clear to anyone. This divorce from the European Union, in addition to the challenges modern languages already face (see Chapter 2), will certainly affect the teaching and learning of languages as well. It is therefore more opportune than ever to find out what motivates UK learners to study languages. Findings from primary and secondary education are essential to get the full picture of MFL teaching in the UK but because our participants are university students, in this section, we will review some of the most recent findings about motivation to learn languages in higher education only.

Given the international status of English around the world, we could speculate that it has diminished the extent of instrumental motivation for learning other languages. But many findings seem to suggest that this might not be the case. For instance, in his study of 378 university students of French and Spanish in the UK, Oakes found that instrumental reasons for studying languages were frequently given by students. A similar conclusion was drawn by Canning (2011)

who revealed that there is a clear link for language students between learning languages and improving their job opportunities. It has to be noted that in Canning's study, participants were non-language specialists but for Busse's 59 university students of German, career goals are also an important motivational factor which "may over time become incorporated into students' visions of their future self" (Busse, 2013: 392). So it seems that for both non-specialists and specialist language students, instrumental value is still important, despite English being the global language nowadays. As we will see in Chapter 5, instrumental-promotion is also a common motivational factor for the participants in the present study. However, as this will be discussed in Chapter 6, it is not always a motivational trait that is associated with high L2 French vocabulary knowledge.

Nowadays, universities in the UK have generally a solid employability strategy that aims to inform students of the different paths their degree could lead them on and departments and faculties are also proactively informing students of the benefit of studying a particular subject, including modern languages. Yet prior to that, modern foreign languages are often "hindered by inadequate institutional support, detrimental policy restrictions and financial cutbacks" (Leerman and King, 2015 in Dörnyei and Al-Hoorie, 2017: 464). But when investigating instrumental motivation, there is generally no information regarding the time and the place where students started developing instrumental motives. There is also no information on the type and the source of evidence they base their instrumental orientation on. And once students are at university, how do instrumental motives develop? Will the support of the institution in the employability sector and the effort to showcase examples where languages can be useful make instrumental motivation more prominent? These are important factors to consider as hindrance or promotion of the usefulness of languages in the students' experience in the UK is "bound to evoke conflicting attitudes not only at the conscious but also the unconscious level" (Dörnyei and Al-Hoorie, 2017: 464).

But even if instrumental orientation is important for learners of LOTEs in the UK, the most important factor for them is the *Ideal L2 Self* and the vision of themselves as competent speakers of the language (Busse, 2010, Busse, 2013, Oakes, 2013). And it seems that having English as L1 is not necessarily a demotivating factor as L1 English only students are more willing to develop proficiency in their L2 than students who have another language instead of English (or in addition to English) as their L1 (Oakes, 2013: 184). Busse and Williams (2010) and Oakes (2013) also found that it is the desire to improve speaking skills in particular that is motivating students the most, followed by the desire to improve listening comprehension, followed by writing and reading. On

the other hand as previously hinted, the *Ought-to* Self is not important to L2 learners in the UK. They do not feel pressured by peers, parents or teachers to pick this subject at university (Busse and Williams, 2010; Oakes, 2013). But whether or not Oakes' point about L1 English speakers being even more motivated to be fluent in their L2 because their native language is English is a sign of anti-Ought-to Self is not clear. Once again, it is evident that more investigation is needed into the Ought-to Self.

Finally, as suggested in the previous section, languages other than English can be linked to specific L2 communities, therefore, integrative motivation might become relevant again. The limited available data on the other hand, do not confirm this; for example, Oakes notes that his students of French and Spanish did not feel any particular affinity with the L2 communities and did not particularly feel like being more like them. And that is despite the fact that they enjoy meeting French and Spanish speaking people and going to French and Spanish speaking countries (Oakes, 2013: 184). Busse and Williams (2010) found similar results and suggest that integrative motivation in its more traditional conceptualisation is still too broad even when considering the learning of LOTEs. The main question of the present study as far as motivation is concerned remains to investigate to what extent this individual difference factor can explain learners' L2 French variation. However, whilst exploring this question, we will also be able to contribute to the discussion about motivation and learning LOTEs in the UK.

3.4.3 Summary

Motivation theories have evolved since the social psychological period but the legacy of Gardner's earlier work has influenced current ones such as the L2 Self Motivation System that has been widely adopted by researchers. In the last few years though, some emerging issues seem to be calling for a fresh retuning of motivation theory. The main reason being that the L2 Self Motivation System as we know it, might not do full justice to LOTE learning and might be most appropriate to L2 English learning as the system has been mostly based on L2 English learning data. The Ought-to self, the Ideal L2 Self/Integrativeness and instrumentalisation seem to require fine-tuning when dealing with LOTEs.

The UK offers an interesting ground for L2 motivation research as this an L1 English speaking country which has recently decided to part with the EU and which has already a high level of monolinguals compared to other countries. In this context, what makes students want to learn

languages? This is one of the questions that this research aims at answering through the use of the motivation questionnaire and the language background questionnaire that we will discuss in the next chapter. But the other major aim for investigating motivation is to find out whether there is a link between the type of motivation and the overall development of L2 French learners at university, and their vocabulary development in particular. As we have seen in this section, many studies of L2 learners in higher education in the UK show that becoming proficient in the language is a major motivation factor. If the present research project concludes as well that this Ideal Self trait is important, naturally, one wonders how these students work towards this aim. Is vocabulary development a major contributor to fulfilling their goals? If so how do they go about it? Is vocabulary development usually linked to Ideal Self-motivation only or can it be linked to other motivational factors? In other words, how far can the difference in motivation factor explain learners' variation both L2 vocabulary and in L2 proficiency?

3.5 Conclusion

In this chapter, we have reviewed the literature concerned with three areas of critical importance in this present study. Firstly, we examined the literature that addresses lexical processing and organisation with a particular focus on Levelt's model of processing with its three lexicon components: concepts, lemmas and lexemes. We reviewed empirical findings about the organisation of the lexicon and the significance of frequency, proficiency, depth and breadth. The importance of input and incidental learning in the theoretical framework we adopt was explored, in particular in educational settings with Nation's definition of a word. Secondly, proficiency, the "multicomponent phenomenon underlying one's knowledge of and ability to use a language" (Leclercq, 2014: 11) was explored in more detail with a particular focus on its theoretical development from Lado (1961) to Ortega (1999) and on empirical findings from the global proficiency measurement tool, EIT. Thirdly, motivation theories from the social psychological period to the process-orientated period were reviewed before exploring the current L2 Motivational Self system (Dörnyei; 2005), as well as the issue of its relevance for exploring motivation for learning languages other than English.

It should be pointed out that, until now, very little empirical research concerning longitudinal lexical development in L2 French appears to have been done using a combination of qualitative and quantitative methods. This present study aims at addressing this gap and at understanding not only the trajectory of lexical development among advanced learners, but also its evolving

relationship with overall proficiency and L2 motivation. Five questions are thus raised which will form the focus for my empirical study:

- What are the rate and the variation of lexical development in L2 French over a 3 year period of instructed learning?
- Is the frequency factor observable in L2 learners' lexical development?
- To what extent can the individual difference factor of motivation explain learners 'variation?
- To what extent does overall L2 proficiency explain the pattern of L2 vocabulary development over time?
- What is the relationship between the extent of exposure to French and vocabulary learning?

The wider questions raised throughout this chapter e.g. regarding the nature of motivation for learning LOTEs will be revisited briefly in Chapter 6 as part of concluding discussion.

Chapter 4 Methodology

4.1 Introduction

This chapter presents a general overview of the study and the research questions. It then includes a detailed discussion about the participants, the procedures and the instruments developed for the study. A multi-method design is proposed in order to answers the research questions.

4.2 Overview of the study

This study aims at describing and analysing the L2 French vocabulary development of undergraduate students of Modern Languages in a British university. Three year groups of students (first years, second years and final years) are studied, each over the 2014/2015 academic year. The design of the study, detailed below, is driven by the theory described in the previous chapter and by the main research questions which are:

- What are the rate and the variation of lexical development in L2 French over a 3 year period of instructed learning?
- Is the frequency factor observable in L2 learners' lexical development?
- To what extent can the individual difference factor of motivation explain learners' variation?
- To what extent does overall L2 proficiency explain the pattern of L2 vocabulary development over time?
- What is the relationship between the extent of exposure to French and vocabulary learning?

The study was primarily cross sectional, i.e. data was collected on students in 3 different year groups, and these will be compared in order to model development over a 3 year period. However for each individual year group, longitudinal data was also collected during the 2014/2015 academic year. To investigate vocabulary development, receptive and productive vocabulary tests were administered at the beginning of the academic year in November 2014 and at the end of the academic year in May 2015. The analysis of this data will determine the lexical profile of the participants within each year groups, and whether and how their profile has changed over the course of the academic year. It will also allow for cross year comparisons.

An elicited imitation test was used in December 2014 to determine the proficiency of the students and finally a questionnaire concerning participants' language learning activities and a motivation questionnaire were distributed in March 2015. Participants' end-of-year grades from their normal instructional programmes were collected and will also be analysed to explore relationships with the especially collected data. Figure 4.1 is a visual representation of the tests and their timescale.

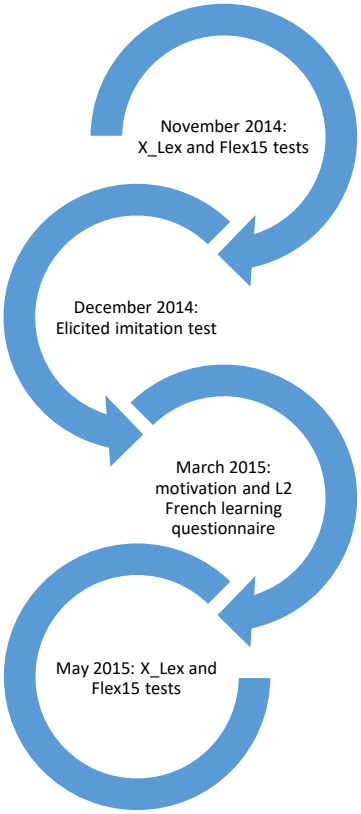


Figure 4.1 Timescale of tests

4.3 Pilot study

4.3.1 Aims of the pilot

Prior to starting the main data collection described below but after having obtained ethics approval, a small pilot study was run in order to try out some of the vocabulary tools and their efficiency, as well as questionnaire methods for investigating potential reasons for learners' variable development. In this pilot, the participants were first and second year students only (whereas the aim of the main study is to collect data from first, second and fourth year undergraduates of French in order to track the lexical development of learners over the full course of their language degree).

The main objective of this small pilot study was to trial a French version of the X-Lex test which evaluates the receptive vocabulary size of learners, to trial a questionnaire which aims at investigating possible reasons for learners' lexical variation and to explore the measurement of L2 French learners' productive vocabulary.

4.3.2 Pilot study participants

18 participants took part in the pilot: 7 first year undergraduate students (UG1), 10 second year students (UG2) and one native speaker (NS). However, one of the 7 UG1 had to be discounted as he ticked 5 non-words in the French version of the X-Lex test, making it impossible to know whether he guessed the meaning of the other words too or not. Therefore, I collected data from 17 participants. More information about the participants will be given when analysing their answers to the pilot questionnaire.

4.3.3 Piloting the X-Lex test

The test used in the pilot study to determine the participants' receptive vocabulary is Meara and Milton (2003)'s X-Lex test which is a yes/no test made up of 20 randomly selected words from the first five one thousand word frequency bands (based on Baudot's list) and 20 non-words. Participants tick the words they recognise, and they are awarded 50 points for each correct answer to obtain the raw score. Then, for each incorrect answer, 250 points are subtracted, and this is the adjusted score. The rationale for the selection of this well-known test for use in both the pilot and main studies can be found in Section 4.5.2 below.

All apart from one UG2 participant and the NS took the paper version of the test (see appendix A) which was emailed to them. They completed it without time restriction and then emailed it back once completed. All the feedback was positive, since students found the test and the instructions clear to follow. The UG2 and the NS who were still in Southampton came on campus to take the computerised version of the X-Lex test. They found the software engaging and easy to operate and they completed the test in less than 10 minutes. The advantage of the computerised version over the paper version was that both the raw score, the number of words participants claim to know, and the corrected or adjusted score were displayed on the screen as soon as the test was finished. The results page also determined the number of words known in each of the first five 1000-word frequency bands. In the main study though, the paper version was used as it was easier to administer to the group of participants which was larger than the group from the pilot study.

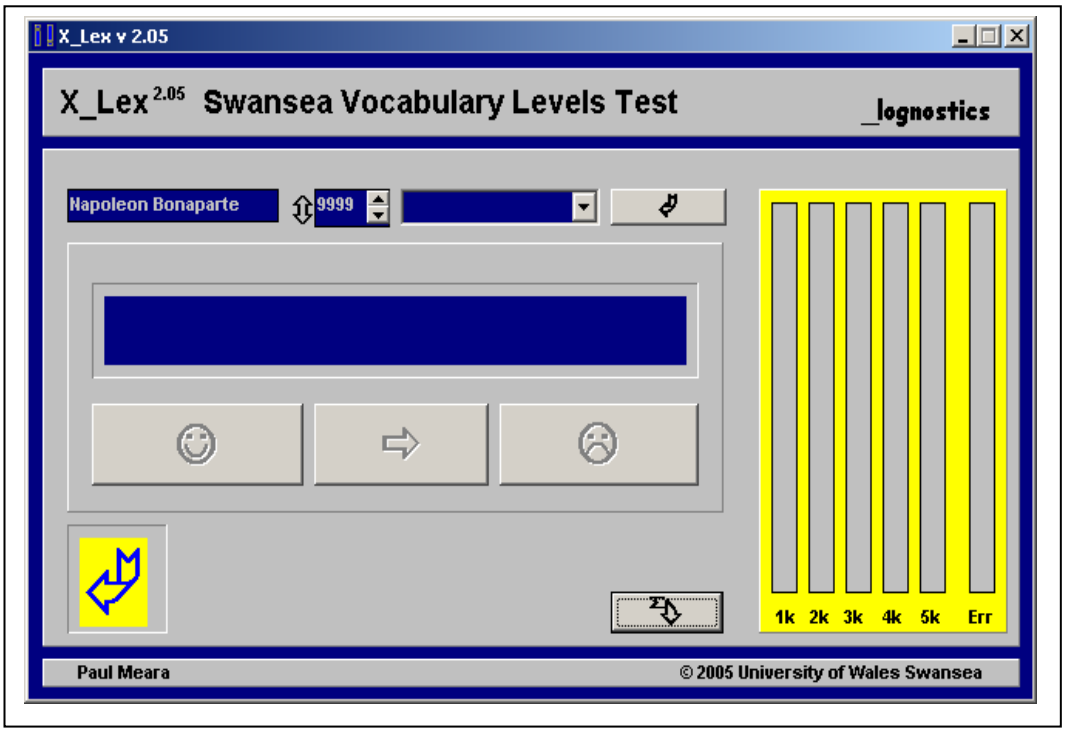


Figure 4.2 Opening page of the computerised version of X-Lex

4.3.4 Results for pilot X-Lex test: Vocabulary breadth size and vocabulary breadth size development between UG1 and UG2

Table 4.1 below shows findings for the UG1 and UG2 groups. UG1 scores varied from 2400 to 4750 with a mean of 3233. As for UG2, the lowest score was 2900 and the highest 4950 with a mean of 3720. As we can see there was a noticeable but by no means great difference of 487 on average between UG1 and UG2.

Participants	Maximum	Minimum	Mean	Standard deviation
UG1	4750	2400	3233	817
UG2	4950	2900	3720	574

Table 4.1 Descriptive statistics for X-Lex adjusted scores for first year and second year undergraduate students of French, pilot study

But in order to assess the credibility of these figures, it was necessary to ask how these figures compared to other studies' results. First of all, in her study investigating the lexical development of students of French from year 8 to UG4 in UK schools and universities, David (2008) also reports a fairly small difference of 330 between UG1 and UG2. In her study, which also used the X-Lex test to measure the receptive vocabulary of 483 participants, a more striking difference is the minimum score of UG1 and UG2 which are respectively 750 and 1250. In our pilot, the minimums were 2400 and 2900. The means in David's study are 2524 and 2854 whereas we found means of 3233 and 3720 for UG1 and UG2 respectively.

This pilot being so small, it was not possible to draw firm conclusions, but the differences between the receptive vocabulary findings for UG1 and UG2 in David's study and in this study are probably not due to the test itself, due to the inclusion of non-words in both tests, as a safeguard against "guessing and overestimation" (David, 2008: 170). So, one explanation could be that the L2 learners in the two studies had very different receptive vocabulary at the end of year 12. Another reason could be that their receptive vocabulary developed at a different pace over UG1 and UG2. Either way, the receptive vocabulary of L2 learners at the end of UG1 in this pilot study, 3233 words on average, is much greater than the receptive vocabulary of year 12 students reported both in Milton's study (1555) and David's (1577). This emphasised the importance of testing UG1 receptive vocabulary at the beginning of their first year of undergraduate studies, as was done in the main study.

As we have seen, important L2 vocabulary spurts have been documented before in the UK education system, mostly between GCSE and A-level. "Learners take their B1 GCSE examinations with, on average, 850 words of French and the B2 Advanced level with just under 2000 words" (Milton, 2009: 84). The spurts of L2 vocabulary between GCSE and A-level shown by both Milton and David, and the one apparently demonstrated in our pilot between A-level and the end of UG1, seem to suggest that every time learners choose to take their learning of French further, this decision may have a direct impact on their vocabulary development: they flourish. But is the spurt due to motivation? Is it due to the learning materials, or contact time, or individual engagement in out-of-class L2 practice?

Let's consider more closely the vocabulary breadth difference between year 12 and UG1 learners. For Milton (2006), the mean receptive vocabulary of year 12 L2 French learners is 1930 words, whereas, according to our data, participants at the end of UG1 had 3233 words, a difference of 1303 words. Given that UG1 students have 3 hours of scheduled L2 French language class's time per week during two semesters of 14 weeks (minus 4 weeks per semester for exams and revision),

students have a total of 60 hours French-medium contact time. Their content modules (history, literature) are delivered in English though students can read French materials in preparation for their weekly content classes. If all vocabulary learning were to take place in class hours, this would mean that during the 2-hour language seminars which focus on text comprehension, expression/vocabulary development and grammar as well as during the one-hour tutorials dedicated to speaking practice, UG1 learners of French were learning on average almost 22 words per class hour.

However, previous studies (e.g. Milton and Meara, 1998) that focus on the rate of vocabulary learning and time spent in the classroom showed that even in the most fruitful classroom environment, students learn on average 6 words an hour. So an average of 22 words an hour seems improbably, which implies that at UG1 level, a lot of the vocabulary learning takes place outside the classroom. This out-of-class learning clearly needed further study if the rate of vocabulary development was to be explained.

On the other hand, there seems to be a plateau between UG1 and UG2 with an average increase of 487 words between the two years in the pilot study. If we take the same 60 hours of classroom instruction as the main source of vocabulary learning, this would mean that, on average, students learn 6/7 words an hour which seems more in line with the data from other studies (Milton and Meara, 1998). However, it might also mean that UG2 learners' receptive vocabulary expands beyond the 5000 frequency band tested in X-Lex. This is a possibility as one of the pilot UG2 receptive vocabulary scores is a near perfect score at 4950. And it seems quite likely, since if L2 learners are expanding their vocabulary outside the classroom using authentic materials such as magazine/newspaper articles, films, news bulletins etc., they are of course exposed to low frequency words. To investigate this possibility, the uptake of vocabulary from both class materials and out-of-class activities needs to be addressed.

Finally, the pilot X-Lex results also highlighted the issue of individual learner variation. UG students, in theory, have had a very similar L2 learning experience up to A-level as shown in the questionnaire answers (see Table 4.2 below): they have taken the same examinations at the same age, often gaining very similar grades. They have had similar contact time with French assistants to practise their speaking in year 12 and 13, and most of them had participated in a trip to France or to a French speaking country. Their lessons were thematically very similar etc. And yet, even at the end of UG1, the vocabulary breadth variation found in the pilot study was very wide. Individual variation is an aspect that was therefore explored in more detail in the main study.

4.3.5 Vocabulary breadth size and lexical profile

The pilot study also showed that the X-Lex test can provide valuable evidence on the lexical profile of L2 French learners. It is now a well-established fact that the frequency factor can be observed when investigating L2 learners' knowledge. That is to say, learners tend to learn more frequent words before less frequent words, despite learning from theme-organised course books. According to Milton (2009: 34), 60% of learners have a "regular" lexical profile, i.e. they know most words from the most frequent bands, least words from the least frequent bands. However, as we can see from Figure 4.3 below, the pilot evidence suggested that UG1 and UG2 are "level 2 deficit learners" meaning that their knowledge of the second most frequent 1000 words is lower than anticipated.

This profile was already observed by Milton (2007). He also discovered that "Level 2 deficit learners" tend to score higher on the memory test, LAT_B, suggesting that "different learning strengths and styles really can influence the foreign language vocabulary that learners acquire in class" (Milton, 2009: 244). If we take into account that the 2000 most common words in a language need to be learned to move swiftly to successful communication, however, this type of profile might be problematic for the learner.

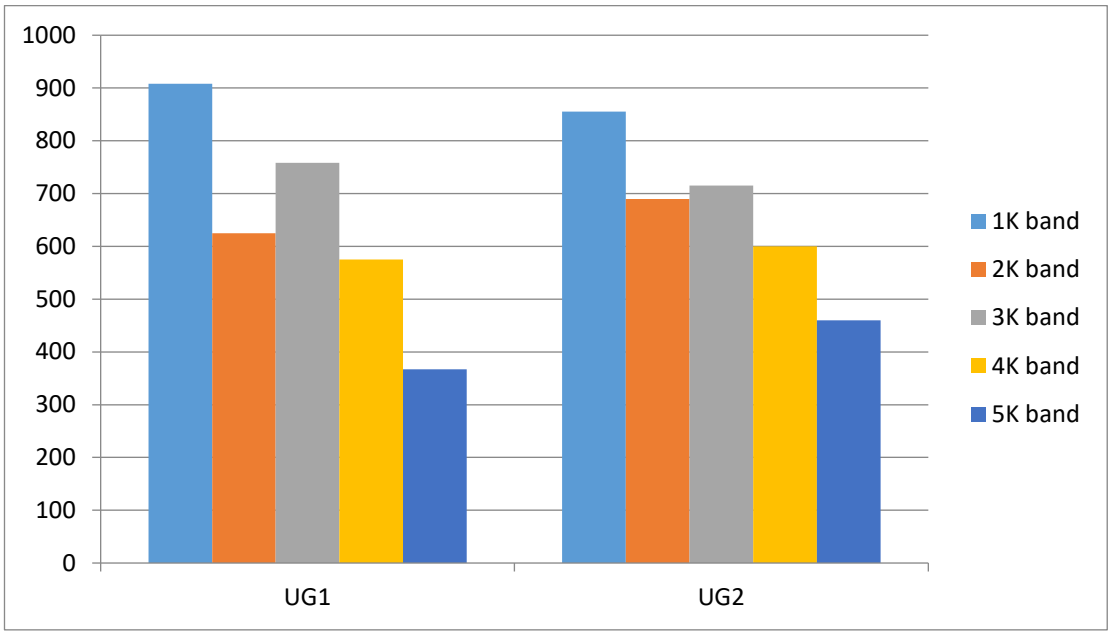


Figure 4.3 Mean number of words known per frequency bands, pilot study

4.3.6 Receptive vocabulary and exam grades, pilot study

As shown in Table 4.2, UG1 and UG2 learners in the pilot study had gained very similar grades for their A-Level in French and there was no obvious relationship between A-level grades and receptive vocabulary. For instance, the vocabulary breadth of A-grade-learners at the end of UG1 varies from 2400 to 3300 and the only B grade student scores 3250. However, there is a noticeable difference in the receptive vocabulary of the only A* student which is 4750 by the end of UG1. The lack of correlation between A-level grade and receptive vocabulary score was also noticed by David (2008) who notes "there was no significant difference in average vocabulary scores as a function of grades obtained in A-level French as measured by ANOVA (F (4, 106) = .742, p= .566)" (2008:174). In fact, in this pilot, the mean of receptive vocabulary of A-grade-learners for A-level is lower than of the B-grade learners.

Participants	**A-level grade**	**X_Lex results**
UG1.1	A	2400
UG1.2	A	2700
UG1.3	A	3000
UG1.4	B	3250
UG1.5	A	3300
UG1.6	A*	4750
UG2.1	B	2900
UG2.2	A*	3350
UG2.3	A	3400
UG2.4	A	3500
UG2.5	A*	3500
UG2.6	B	3550
UG2.7	Other exam (B2 level), 82%	3750
UG2.8	A	4000
UG2.9	B	4300
UG2.10	A	4950

Table 4.2 A-level and X-Lex results

4.3.7 Vocabulary size and end of year language grades, pilot study

"There is an assumption that vocabulary knowledge ties into some language skills rather better than others and that it can contribute more to some aspects of language performance in particular" (Milton, 2009: 171). For instance, Stæhr (2008) found a strong correlation between vocabulary size and listening and writing but the strongest link was with reading (0.83).

In the pilot study, in order to explore the relationship between vocabulary size and the four skills as far as was practicable, we used the UG1 scores for an end of year combined exam which consisted in a reading comprehension exercise, a translation (from French into English) and a composition; a listening exam and an oral exam (individual interview). We ran a statistical test (correlation) between these individual scores and the X-Lex results.

As we can see from table 4.3, there is a strong positive correlation between vocabulary size scores as measured by X-Lex, the three exam results and the overall grade with a particularly strong link with the listening and the oral skills.

	Combined exam (reading, translation, composition)	**Listening test**	**Oral test**	**Overall grade**
Vocabulary size	0.64	0.94	0.92	0.89

Table 4.3: Correlations between end of year scores and X-Lex scores, UG1 participants, pilot study

As we can see, the X-Lex-listening correlation is very strong at 0.94, compared to Staehr's (2008) 0.69 but the most surprising correlation is the one between oral test scores and vocabulary size as receptive vocabulary breadth does not usually predict speaking level (Milton, 2009; Adolphs and Schmitt, 2003).

However, as we turn to the UG2 learners' exam grades, there seems to be no meaningful relationship between vocabulary size as measured by X-Lex, the combined exam, which consists of a translation from French into English and of an article in English that L2 learners need to summarise in French, the listening exam, the oral exam and the overall grade. All correlations in this case were close to chance.

	End of year exam (translation, summary)	Listening test	Oral test	Overall grade
Vocabulary size	0.085	0.12	-0.11	0.09

Table 4.4: UG2 Correlations between end of year scores and X-Lex results, UG2 participants, pilot study

It goes against most findings (Milton, 2009: 176) to suggest that there is no link between vocabulary size and language skills. Thus the lack of link between the two at UG2 level suggests that the year 2 examinations may be testing more/ different factors other than language skills. For instance, at UG2, the reading comprehension element was replaced by a translation exercise (from French into English) which might involve other cognitive factors than the ones involved in a reading comprehension. Likewise, writing in French at UG2 involves summarising a text from English into French, not allowing students to use the words they know in a creative piece, but rather using the words they know about the theme in the text they have to summarize. The oral exam is also much more a general conversation about themselves rather than a presentation they can rehearse in advance, which is very much the case at UG1 level as well as at A-Level and GCSE. The place of vocabulary criteria in mark schemes might also explain the lack of correlation. For instance, if we take the written task assessment criteria for UG1 learners, vocabulary, together with style, accounts for 20% of the grade whereas in the UG2 written summary criteria, there is no explicit vocabulary criterion, though this is implied in the section *style, range of expression and syntax* which is worth 30% of the overall grade. So at UG2 level, L2 learners are required to display some knowledge of idiomatic expressions as well as varied vocabulary. Does this requirement have an impact on the vocabulary they learn? The lack of correlation between vocabulary size and exam grades at UG2 might be simply due to the fact that the pilot sample is too small or presents some oddities. However, an important lesson from this pilot was that end of year grades could not be straightforwardly relied on as a measure of L2 proficiency, and that another alternative measurement of proficiency would be required in the main study.

4.3.8 Data from the pilot questionnaire

The pilot questionnaire aimed at gathering information about the participants' language learning background. They were asked to provide personal details, French language background details (including their GCSE and A-Level grades) and information about other languages studies at university as well as about additional languages spoken at home. Their language learning practices were investigated through a series of open-ended questions (appendix B).

UG1 participants

Participants	Age	Years learning French	Other languages studied at university	Other languages spoken at home	GCSE grade	A-level grade	X-lex score
1	19	8	None	None	A*	A*	4750
2	19	8	Portuguese 1+2	None	A*	A	3300
3	20	8	Spanish 4 and Italian 1+2	Spanish	A	B	3250
4	19	12	Spanish 1+2	None	A*	A	3000
5	20	8	Spanish 1+2	None	A*	A	2700
6	20	8	Spanish 1+2	None	A*	A	2400

Table 4.5 Results of pilot questionnaire, background section, UG1

As summarised in Table 4.5, the UG1 learners in the pilot study were aged between 19 and 20 and had 8 years of French instruction, apart from one who had 12 years. They all sat GCSE and A-level achieving very similar grades. The academic profile of the learners was high as none of them received less that B for GCSE and A-level. Most pilot participants were British students who only spoke English at home apart from one who spoke Spanish.

We can also notice that most pilot participants, apart from one, studied another language in accelerated mode, meaning that at the beginning of the year they were complete beginners in that language and that by the end of the year, they reached post-A-level.

The first interesting observation we can make is that the breadth of vocabulary of UG1 learners as indicated in their X-Lex scores did not seem to correlate with studying another language. Indeed, the highest receptive vocabulary score of 4750 is achieved by the single student who studied French only and exposure to a new language at Stage 1 and 2 accelerated did not seem to have a visible effect on the receptive vocabulary of French. Indeed, we can notice that students who studied Italian, Spanish or Portuguese 1+2 varied widely in breadth from 2400 to 3300. Given the small number of students involved in this pilot, it is difficult to conclude whether learning another language has an effect on the development of breadth of French. More participants who study French only at university need to take the test in order to verify this. However, this issue was not pursued in the main study.

Next we considered the answers of pilot participants to the questionnaire section about their language learning. Interestingly enough, the two students with the highest scores (4750 and 3300) gave a clear definition of what a good language learner is. They both reported dedicating good amount of time to independent studies. One of them even gave 25 hours of listening practice a week.

Students who scored between 3250 and 2700 gave still a fairly good definition of what it is to be an independent learner but they seemed to dedicate less time to independent learning. They acknowledged that this was an important part of their development but for one reason or another, they did not manage to practise as much as they knew they should.

The learner who scored the lowest (2400) gave the vaguest definition of being an independent learner. His/her independent work activities were not very precise and no details were given about the amount of time dedicated to self-study even though this was specifically requested.

So these preliminary findings seem to indicate that the amount of time dedicated to self-study correlates with high or higher scores for vocabulary breadth, and therefore that an important part of vocabulary uptake at UG1 level takes place outside the classroom. The amount of self-study seemed linked to self-motivation rather than to an understanding of strategies. This led to the important decision to expand the questionnaire in the main study, to include an investigation of learner motivation alongside the investigation of their language learning practices. "Motivation and interest are important enabling conditions for noticing" (Nation, 2010: 63).

Interestingly enough, one pilot participant who scored 3250 on X-Lex had spent 9 months working in France as an au pair in a French speaking family, and she was also a fluent speaker of Spanish. In this case, why is this score so low? She seemed to have a good understanding of what is required of a language learner but the self-study activities that she listed are conversations with French ERASMUS students, with friends, and watching DVDs, rather than reading, grammar practice etc. This suggests that the quality and type of L2 practice activities needs further investigation, and not only their amount. This was taken into account in the main study and the participants' self-study activities were categorised following Nation's four strands.

UG2 participants

Participants	Age	Years learning French	Other languages studied at university	Other languages spoken at home	GCSE grade	A-level grade	X-lex score
1	21	7	Spanish 5, Italian 1+2, Russian 2	Greek	A*	A	4950
2	21	12	Spanish 3+4	None	A*	B	4300
3	20	16	Italian 3+4	None	A*	A	4000
4	21	6	Spanish 3+4	Hungarian	A*	A	3750
5 6	20	8	German 1+2	French	A*	B	3550

7	23	12	Spanish 5, Japanese 2	None	A*	A	3500
8	20	9	Spanish 5, Portuguese 3+4	None	A*	A	3500
9							
10							
	20	7	Spanish 5	None	A	A	3400
	21	9	German 5	None	A*	A*	3350
	21	8	None	None	A*	B	2900

Table 4.6: Background of UG2 participants, pilot study

The ten UG2 pilot participants were aged between 20 and 23 and studied French for an average of over 9 years. Most of them also studied another language apart from one who studied French alongside English literature.

As just seen, for the UG1 students, other languages did not seem to influence their French vocabulary breadth. However, as we turn to UG2, the lowest score (2900) was from the only student who did not study another language. She gave a good definition of what made a good language learner but described independent activities only in vague terms, which could be the reason for her lower score.

The student who scored almost 5000 (4950) spoke Greek at home, and studied Spanish Stage 5, Italian accelerated 1-2 and Russian stage 2. She was clearly language motivated and in her definition of what makes a good language learner, she emphasised the importance of understanding the history and the culture of the language studied as well as developing a great variety of vocabulary and expressions. The other two students who scored 4000 and 4300 gave a clear definition of independent learning and listed a fair amount of independent activities.

There was a strong correlation between motivation, independent learning and growth. However, this data also confirmed that language learners can vary tremendously in their lexical development, highlighting the need to investigate other potential reasons for this language

variation. Hence, the questionnaire in the main study also included a motivation section and the Elicited Imitation test to determine and evaluate participants' proficiency.

4.3.9 Pilot participants' productive vocabulary

In order to explore the productive vocabulary of pilot participants, samples of their L2 French writing were collected, in the form of essays written specially for the pilot, as a response to the instruction "Describe your academic year". The original idea was to evaluate the productive vocabulary of the participants by applying Meara's P_Lex program to these writing samples. The P_Lex program divides short English and French texts into segments and assesses their "difficulty", in terms of the proportion of words from the 3K frequency band and beyond. However, running this program has proven more difficult than expected, we therefore turned to the V- word program (Meara, 2014) to calculate the tokens and types of each UG1 and UG2 pilot participant. Results are presented in Tables 4.7 and 4.8.

Participants	Tokens	Types	X-Lex score
UG1 one	456	255	4750
UG1 two	598	272	3300
UG1 three	315	177	3250
UG1 four	185	112	3000
UG1 five	301	166	2700

Table 4.7: Number of tokens and Types in UG1 essays, pilot study

Participants	Tokens	Types	X-Lex score
UG2 one	573	251	4950
UG2 two	277	177	4300
UG2 three	269	159	4000
UG2 four	301	186	3750
UG2 five	337	203	3550
UG2 six	222	141	3500
UG2 seven	365	195	3500

UG2 eight	355	201	3400
UG2 nine	310	181	3350

Table 4.8: Number of tokens and Types in UG2 essays

There was a good correlation between X-Lex results and the number of types in student L2 scripts, 0.52 for UG1 and 0.62 for UG2, however, more information about the frequency of types would have been useful in order to determine whether the participants' productive vocabulary frequency profiles matched the passive vocabulary profiles. In the main study, the essay type question has not been replicated though and productive vocabulary was assessed using a new specially developed test (the Flex15 test) inspired by Meara and Fitzpatrick's (2000) Lex30 test. The reason for this change of instrumentation was to elicit more low frequency words without requiring participants to write lengthy essays. The new test is described in Section 4.5.2 below.

4.3.10 Conclusion

Having tested a small number of undergraduate students of French at the end of their first and second year at university, the first conclusion we drew from the pilot study was that the variation in vocabulary size amongst L2 learners of French observed by David, (2008) and Milton (2009) was confirmed. There was even an overlap between year groups, with some first year learners having a much larger receptive vocabulary knowledge than that of second year students.

Furthermore, the average vocabulary size at the end of UG1 was almost double the average of year 12 vocabulary size. This spurt, like the one that happens between GCSE and A-level, seems to coincide with the time when learners freely elect to take their study of French further. However, to explain this spurt, further questions need to be asked:

> Is the spurt the result of L2 learners' motivation, hence taking every opportunity to learn new vocabulary?
>
> Is it due to a change of topics for which students have limited vocabulary to talk or write about and for which they must build a vocabulary bank quickly?

From the answers participants have given about their language learning, it seems likely that they are motivated language learners and that when this motivation leads to intensive, regular and varied opportunities to practise the language we can notice a larger vocabulary breadth for these

students. These observations suggested the need to pursue more fully in the main study both the issue of motivation, and the type of L2 learning practices in which students engage.

After such an impressive spurt between A-level and UG1 vocabulary breadth, a relative plateau was found between UG1 and UG2 with an increase of only 487 on average. However, as already pointed out, it would be useful to test students' 5 to 10K frequency bands in order to find out whether there is indeed a plateau or whether students are in fact learning more low frequency band words. The main study has taken this into account and in the second X-Lex test; participants were also tested on 6K and 7K bands.

Interestingly, at UG1 and UG2, learners showed level 2 deficits but apart from that, they display a regular lexical profile. However, this says nothing about what their lexical profile might be like between 5K and 10K. Would the frequency theory still apply? In addition to testing the receptive knowledge of low frequency words, it was decided to test the active knowledge of less frequent words using for instance an adapted version of the Lex30 test alongside the X-Lex test.

Furthermore, even though there were some strong correlations between UG1 exam grades and vocabulary size, especially for listening and speaking, there were no correlations at all between exam grades at UG2 and vocabulary size.

4.3.11 Concluding implications of the pilot for the main study

The pilot study confirmed that the objectives set for the main study are worth pursuing in more details as some findings from previous studies were confirmed but others were not. The need to explore factors underlying individual variation is also an important objective, both to understand advanced second language acquisition and to inform pedagogy.

The pilot also had some direct implications for the methods used to collect data in the main study. First of all, it confirmed the usefulness of the X-Lex test to study the development of receptive vocabulary. The paper version of this test was used because even though it is more time consuming to calculate the words per frequency bands using this version, it did allow all of the students to take the test at the same time and without risk of losing data. Secondly, as some of the scores on passive vocabulary were very high, for instance 4950 out of 5000, the main study takes into account that some participants' vocabulary uptake may be from lower frequency bands. Therefore the second receptive test given to participants at the end of the academic year samples the first seven rather than 5 frequency bands to get a full picture of their lexical profile and to verify whether it replicates the lexical profile in Aizawa's study (2006), in which it was

found that the frequency band difference from 4000 to 8000 seem to be more even than that of the bands from 1000 to 3000 (quoted in Milton, 2009: 28).

As a result of the pilot, motivation theories have also taken a more predominant place in the theoretical framework of this study which led to the development of the motivation questionnaire, adapted from Taguchi et al (2009), in order to determine if and to what extent motivation factors can shade some lights on lexical variation. Although using the V-Lex tool to calculate the number of tokens was useful, in the main study, participants' depth knowledge is tested using a variant of the Lex30 test. Finally, we have seen, using exam grades to evaluate lexical development and proficiency is not always reliable and for this reason, the Elicited Imitation test (EI) has been adopted as a proficiency measure for the main study.

4.4 Participants and setting

The participants in the main study are 40 undergraduate students including 32 females and 8 males (see Table 4.9) aged between 19 and 23 at a Russell group university in Britain. They study French either as a single subject or in combination with another language (89%) or another subject (11%).

	UG1	**UG2**	**UG3**
Female	10	10	12
Male	2	3	3
Total	12	13	15

Table 4.9 Breakdown of participants per gender and per year group

All participants have joined the university having previously studied French at school and college and given the high entry requirements for this university, all students gained top grades in A-level French, mostly A* and A, or in other qualifications such as the International Baccalaureate (IB). Despite the declining A-level entries in languages in the UK, the French cohorts at the university are healthy. In 2014/2015, there were 111 first year students, 114 second year students and 76 finalists who had spent their intercalary year in a French-speaking country either studying in a partner university, being an English language assistant in schools or completing an internship.

The university has a system of language stages mapped onto a number of national qualifications as well as the Common European Framework of Reference for Languages (CEFR). For instance, post-A-level students of French usually embark on Stage 4 French and the institution states that "successful completion of Stage 4 is approximately equivalent to reaching Levels B2/C1 of the Common European Framework and Level 3 of the National Language Standards" (University of Southampton, 2016). So typically, undergraduate students study French at Stage 5 in their second year (equivalent to reaching Level C1 of the CEFR and Level 3/4 of the National Language Standards). In their final year, students who did not go to a French-speaking country for their year abroad typically study French at Stage 6 and those who did complete their intercalary year in a French-speaking country study French language at Stage 7 (equivalent to reaching Level C2 of the CEFR and Levels 4/5 of the National Language Standard).

Students are placed in language stages based on their language knowledge rather than on their year of studies. However, in this research project, all participating first years are Stage 4, all second year students are Stage 5 and all finalists are Stage 7. For this reason, we will refer to the three cross-sectional groups as year 1 (UG1), year 2 (UG2) and finalists (UG3). In that sense, we can say that the participants are representative of the larger cohort too as this is the most common distribution of students across the three stages. Ab-initio students have not been included because the study aims at investigating the development and variability in L2 French learning at university amongst instructed students of French who share a very similar language learning history. Finalists who did not complete their year abroad in a French speaking country have not been included either, so as to minimise variability among the UG3 sample.

4.5 Procedure for main study

The research project was reviewed and approved by the Faculty ethics committee in June 2014 (ERGO submission number 7759) and an announcement was sent to all the students in UG1, UG2 and UG3 in October 2014, via Blackboard, the Virtual Learning Environment (VLE) used by the university, informing them that volunteers were needed for a study on French lexical development. The message stipulated that participants would be required to undertake various short tasks three times over the academic year. By the end of October 2014, over 10 participants for each year group were gathered and consequently a meeting was organised with students who had expressed interest in order to provide them with further information regarding the aims of the study, the instruments being used and the timescale. In this meeting, students were also

made aware that they would be very welcome to get in touch with myself, the study researcher, if they wished to obtain some feedback on their vocabulary profile as well as to talk about vocabulary learning strategies, which was the only compensation offered to take part in the study. All participants were informed that the data collected would be saved anonymously and that would not be discussed with other members of the Department. Participants who were also the researcher's students were informed that the research data would not, in any way, be taken into account in the assessment of their academic work. Students then had the opportunity to raise further questions. During this meeting, students were given an information sheet detailing the study and its procedure (Appendix C) as well as a consent form (Appendix D). At the end of this information session, students who were still willing to participate in the study completed the consent form.

In November 2014, the first round of tests was organised. Participants were contacted by email and a group meeting at a time and place that suited all was organised on campus. On the day, participants took the paper versions of the X-Lex receptive test and the Flex15 productive test, that are described below, and which are referred to in the study, as X-Lex Nov and Flex15 Nov. Before leaving this first test session, participants were also asked to select a timeslot for the next test, namely the Elicited Imitation (EI) test. This proficiency test requires participants to listen to various sentences in French and to repeat them aloud as accurately as possible, therefore individual times had to be organised. The researcher, with the help of the French language assistant met the 40 participants on one-to-one basis in December 2014 to administer the EI test. Participants' answers were recorded on a Dictaphone. The audio-files were then anonymised and saved onto the researcher's computer. In March 2015, a second group meeting was organised on campus on a day and at a time that suited all, so that participants could complete the motivation questionnaire and the language learning background questionnaire. Instructions were given and participants were encouraged to develop as much as possible their answers to open-ended questions concerning their language learning history.

Finally, a last group meeting was organised in May 2015 for participants to take the second versions of the X-Lex test and of the Flex15 test referred to in the study as X-Lex May and Flex15 May. Participants were familiar with the format of the tests, identical to that used for the November tests, the only difference being that the May tests contained different stimulus words. Nevertheless, participants were once again given clear instructions about the tests before completing them. Upon completing these last tests, participants were thanked for their

participation and were encouraged to get in touch if they had any questions about the study or if they wished to have feedback on their lexical development over the academic year.

4.6 The research design

4.6.1 Receptive vocabulary measurement

Most L2 vocabulary research involves estimating the number of words that learners know that is finding out the breadth of vocabulary they have at their disposal. Given the large size and idiosyncratic nature of advanced learners' lexical knowledge, measuring learners' vocabulary breadth or size comprehensively and directly is effectively impossible. Sampling methods are therefore typically used, and total vocabulary size is estimated from these. There are two main methods which are used to estimate vocabulary size, both of them based on existing word lists in the target language: "One is based on a sampling from a dictionary and the other is based on a corpus or a frequency list derived from a corpus" (Nation, 2010: 363). For the first method, a sample of words is extracted from a dictionary and learners are tested on these words. Corpus-based studies, on the other hand, work with lists of words of known frequency and are useful to identify and test learners' knowledge of both high frequency and lower frequency words. Such tests can be biased toward particular frequency bands, so that for example tests of high frequency words have been extensively used with learners of English at lower proficiency levels.

To test the receptive vocabulary of learners, recognition of the form-meaning relationship is often tested, which, as we have seen in the previous chapter, is the first step towards multifaceted lexical knowledge. Many receptive vocabulary tests that rely on learners' recognition provide learners with translations or explanations to choose from (Milton, 2009). One of the most widely used of such tests, which was considered for use in the present study but ultimately abandoned, is the Vocabulary Levels Tests (Nation, 1990). The VLT was designed to provide an estimate of vocabulary size for second language learners of general academic English (Schmitt et al, 2001: 55). An example taken from Nation (2010: 416) is given in Figure 4.4 below. Participants' knowledge of the words on the left is tested as they have to match three of them with the definitions provided. Words for the VLT were taken from the frequency lists available at the time: Thorndike and Lorge (1944), Kučera and Francis (1967) and the General Service list (West, 1953). Nation's Levels Test (revised by Schmitt in 2001) is considered by Schmitt as the closest thing to the most appropriate receptive vocabulary test available in the field of L2 vocabulary pedagogy and research. However,

it is also a complex test as it requires participants to know not only the words on the left that are being tested but also those that make up the definitions. For this reason, this test format was not used, as it does not only test participants' knowledge of targeted words but it also to some extent evaluates their knowledge of the words in the definitions.

A Vocabulary Levels test: Version 1

This is a vocabulary test. You must choose the right word to go with each meaning. Write the number of that word next to its meaning.

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

Figure 4.4 Nation's Levels Test, sample item

An alternative approach to testing vocabulary breadth is to use checklist tests which are deceptively simple. Participants are presented with a list of words and they have to tick the ones they recognise. One of the most common of such tests, and the one selected to test the receptive vocabulary of the participants in the present study, is the Yes/No test developed by Meara and Milton (2003), the X-Lex test.

This meaning recall Yes/No test is made up of 20 randomly selected words from each of the first five 1000 word frequency bands for the particular target language, and 20 non-words (120 items in total). Participants tick the words they believe they recognise, and they are awarded 50 points for each real word selected to obtain the raw score. For each non-word selected, 250 points are then subtracted, to produce the adjusted score. The presence of non-words in this type of Yes/No test "allows an estimate to be made of the degree of over-estimation which a learner is making and scores can be adjusted on the basis of this" (Milton, 2009: 73). For the two French versions of X-Lex created for the main study, the frequency list used is Baudot's (1992) lemmatised French vocabulary list which compiles words from a variety of literary, media and other sources. The non-words were selected from Milton's French X-Lex Vocabulary Test 1 and French X-Lex Vocabulary Test 2 (2009: 257-258). Even though there is a computerised version of the test, a paper version

was used for the X-Lex test in November, sampling the five thousand most frequent words in French (Appendix E) and for the X-Lex test in May, sampling the seven thousand most frequent words (Appendix F). The scoring criteria explained above were applied to both tests and results will be discussed in the next chapter.

One of the reasons why this test model was selected is because it is straightforward to create as the template is readily available (Milton, 2009: 258). It is equally easy to administer to the participants and it is usually completed in around five minutes. The instructions are very simple and are "Please look at these words. Some of these words are real French words and some are invented but are made to look like real words. Please tick the words that you know or can use". Secondly, as opposed to other recognition tests such as Nation's Levels Test, X-Lex does not require students to know other words than the ones that are being tested. Finally, there is a great amount of background research on the Yes/No method that X-Lex exploits. This method was first used by Higgins (1977), though it was only in 1983 with the work by Anderson and Freebody with L1 speakers that it started to gain more recognition and importance. Meara and Buxton (1987) and Meara and Jones (1988) subsequently developed the technique for L2 speakers, particularly speakers of English as a second language. Since these early papers, many vocabulary research studies have been conducted and the majority of them advocate the Yes/No methodology. The proven robustness of the method is therefore one more reason why it was chosen for the present study.

4.6.2 Productive vocabulary measurement

As already discussed in the previous chapter, vocabulary knowledge is multi-dimensional and the different elements that comprise it cannot all be tested at the same time so the relationships between these elements are still not very clear. Some feel that productive vocabulary knowledge is so complex and little understood that valid standardised tests are difficult to devise. Schmitt comments: "I feel that global measures of lexis (e.g. type-token ratios) will generally be less informative than measures of how appropriately the individual items are used but at the moment, there is no recognized measure of the appropriacy of written lexis in compositions" (2009: 39). And Milton notes that "attempting to measure depth of knowledge appears to have asked more questions than it has answered about the way words are acquired" (2009: 168). In these circumstances, it is not an easy task to test productive vocabulary.

For this research project, we were inspired by Meara and Fitzpatrick's (2000) Lex30 test which is a free productive test first created to estimate the productive knowledge of L2 English learners. One of the reasons for considering the Lex30 test as an inspiration for the Flex15 tests is that Lex30 is "underpinned by a model of vocabulary acquisition whereby the order in which learners acquire words aligns with the frequency of those words in language use" (Fitzpatrick and Clenton, 2017: 850). This means that the productive and the receptive tests in this research project are based on the same assumption concerning vocabulary acquisition.

A Lex30 test also tends to elicit more low frequency words than discursive tests like the Lexical Frequency Profile (LFP). In the latter, the format of the test and the number of function words that participants have to produce might impede the production of low frequency words. In the Lex30 test on the other hand, there are multiple lexical activations, that is to say, for each stimulus word, learners need to dip four times in their lexicon to find four responses, a task that tend to elicit content words and more low frequency words. Some might argue though that Lex30 type tests are not particularly useful to gather information about the quality of vocabulary produced. For instance, it is not possible to know whether the elicited vocabulary could be used "with semantic appropriateness and grammatical accuracy in context" (Fitzpatrick and Clenton, 2017: 860) which is the highest level of quality of a learner's word knowledge on the four level scale mapping of the Capture Zones of Productive Vocabulary developed by Fitzpatrick and Clenton (2017: see Figure 4.5 below).

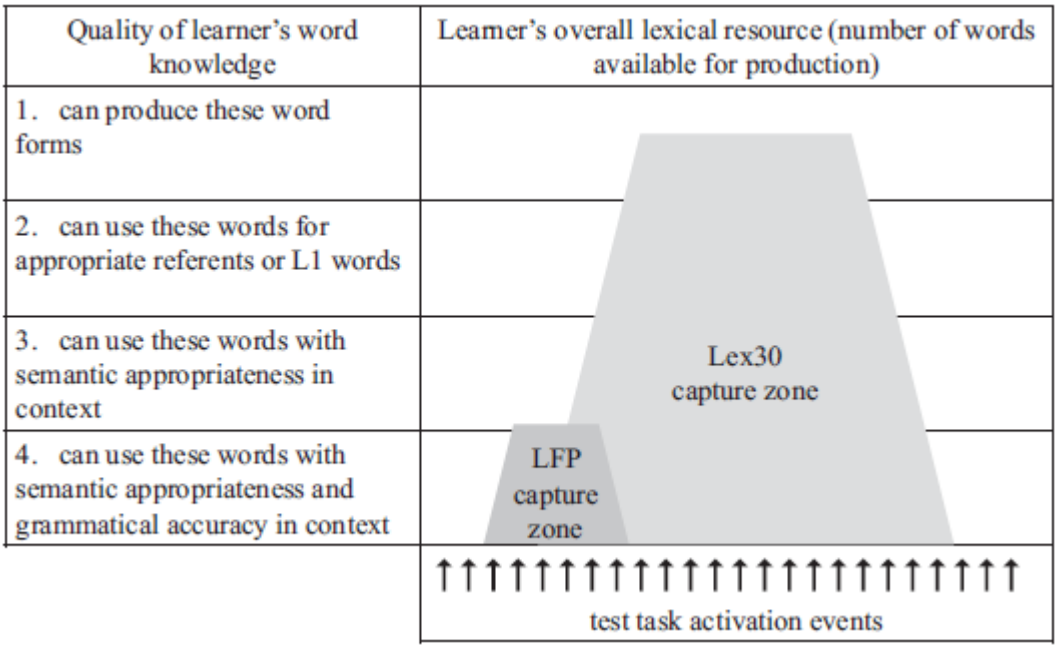


Figure 4.5 Vocabulary test capture: Lex30 and LFP

However, Lex30 types of test are more appropriate for the present research project because one of the aims is to measure the size, that is to say the quantity of productive vocabulary knowledge and to consider it in relation with receptive vocabulary knowledge and overall proficiency. The quality of vocabulary will nevertheless be considered when analysing the productive vocabulary results. Another reason why Lex30 was a source of inspiration for the Flex15 tests is that it is a very simple and quick test to administer. Finally, its validity and reliability have been tested (Fitzpatrick and Clenton, 2010).

The original Lex30 test consists of thirty stimulus words in English. Participants are required to read these words and to write four associated words they can think of for all thirty stimuli, hence leading to a maximum of 120 answers per participant. All thirty stimulus words are taken from Nation's 1000 most common words to ensure participants can recognise the stimuli and can produce four words as requested. In order to elicit a variety of responses from participants and therefore to differentiate between them, words that tend to elicit high frequency responses, such as "salt" which very often elicits "pepper" are not included. In selecting stimulus words that are not obviously associated to a response, Meara and Fitzpatrick aim at tapping into participants' low frequency vocabulary. All Lex30 responses are categorised per frequency band and infrequent answers, that is to say answers that fall outside the 1K band, are given a point.

The Flex15 November and Flex15 May tests designed for the present research share some similarities with Meara and Fitzpatrick's Lex30 test. First, Flex15 tests and the Lex30 test are

comprehensive measures of productive vocabulary because they do not focus on testing specific vocabulary, in other words, "the tester does not have a predetermined list of items that the testee must produce" (Fitzpatrick and Clenton, 2017: 850). Both Flex15 and Lex30 tests also take a *discrete* approach to measuring productive vocabulary meaning, as we have just seen, that they do not embed the test within an assessment such as a reading or a writing task.

For the present study, the Flex15 tests consist of fifteen stimulus words instead of 30, first for practical reasons as the time and resources available to analyse data are limited. Selecting half of the number of stimuli compared to the Lex30 test will still enable us to calculate for instance the percentage of low frequency answers and to compare it with prior findings. Low frequency words tend to be elicited by "the multiple dip activation events" that participants go through every time they have to produce four responses to one given stimulus. So if infrequent vocabulary is elicited by the number of responses requested rather than by the total number of stimulus words, reducing the total number of stimuli should not mean reducing participants' occasions to access their infrequent vocabulary.

The fifteen stimulus words for Flex15 November and for Flex15 May were selected from Baudot's (1992) French frequency list and include nouns, verbs and adjectives. As for Lex30, the stimuli were selected on the basis that participants would recognise the words but one important difference though is that in the Flex15 tests some of the stimulus words were selected from low frequency bands. One reason for selecting some infrequent words was to investigate whether the frequency of a stimulus has an impact on the lexical activation of infrequent vocabulary. In other words, does a low frequency stimulus word tend to elicit low frequency answers? The details of the words and their frequency will be detailed in the next chapter. Even though we did not use a French equivalent to the Edinburgh Associative Thesaurus to verify the word associations of the stimulus words, care was taken to avoid stimulus words that have obvious associations. Retrospectively though, fewer cognates could have been included in Flex15 November and Flex15 May.

A paper version of the Flex15 November test (Appendix G) and the Flex15 May test (Appendix H) was distributed to participants who had about 10 minutes each time to complete the tests. The format of the two tests was identical but the words were not. The answers given by participants were then lemmatized and the frequency of each lemma was then checked against Baudot's list. Every answer was then scored one point and non-words, proper nouns and numbers were labelled level 0 and were not scored. Words that were not found in either Baudot's frequency list or in Lonsdale and Le Bras (2009)'s frequency dictionary were categorized as *not found*. Whereas

in Lex30 tests only responses beyond the 1K band are given a point, in the Flex 15 tests all responses are given a point and they are organised in 5 frequency bands: 1K, 2K, 3K, 4K and 5+ K. This way we expected to determine the lexical profile of participants as far as their productive vocabulary is concerned in the same way as for their receptive vocabulary. This also enabled us to observe how the lexical profile of participants might change over a period of three years of instruction at university and to observe differences or similarities in receptive and productive profiles. As we will see in the next chapter, correlations between the different tests, receptive, productive and proficiency ones, have been calculated using SPSS. However, given the small totals for some bands in Flex15 tests, we decided to group responses from 1K and 2K bands together to reach more meaningful statistical results. This grouping does differ from the way results are organised and analysed in Lex30 test. On the other hand, these two bands are often grouped and labelled as high frequency bands.

4.6.3 Proficiency measurement

The global oral proficiency of the participants in this study was tested using an Elicited Imitation (EI) test in which participants listen to increasingly more complex sentences once and then have to repeat them aloud. The test is based on Hulstijn's (2012) view of proficiency as based on two kinds of abilities: basic and higher language cognition (BLC and HLC). BLC is the language ability shared by most adult speakers; it is this general proficiency that the EI test is aimed at. For some critics such as Vinther (2002), elicited imitation is merely a parroting exercise. However, as we have seen in Chapter 3, proponents of an EI test argue that it reflects underlying syntactic and lexical capacity. That is, if the length of a stimulus sentence exceeds working memory capacity, a participant can only repeat the sentence if he/she "has grammatically parsed and decoded the message and formed a mental representation of it" (Tracy-Ventura, McManus, Norris and Ortega, 2014: 5). Even though EI has been used in SLA research since the 1980s (Hameyer, 1980 and Savignon, 1982 in Tracy-Ventura et al, 2014: 4) and despite the fact that it has been used increasingly more in the last few years, the method remains scarce in L2 French studies. Prior to the study reported by Tracy-Ventura et al (2014), only three studies had used the Elicited Imitation method to investigate L2 French proficiency (Markman, Splika and Tucker, 1975; Erlam and Loewen, 2010; Burger and Chrétien, 2001).

The EI test used in this study was developed by Tracy-Ventura et al (2014). As reported in Chapter 3, it is based on the model developed by Ortega (1999) and consists of 30 sentences ranging from 7 to 19 syllables (Appendix I) with a variety of grammatical structures and vocabulary that students have to repeat. The EI audio was presented via a laptop and participants' answers were

recorded and then saved. The results were scored according to Ortega (1999)'s scoring guidelines (Appendix J). A score of 4 out of 4 was awarded for exact repetitions, that is to say when a repeated string matched the stimulus exactly, both in form and meaning. Strings which were ungrammatical but which preserved the original meaning were scored 3 out of 4. Strings preserving at least more than half the content of the original stimulus were scored 2 and strings which left out more than half of the content were scored 1. Non repetitions or unintelligible repetitions were scored 0. The researcher scored all the EI tests and her supervisor a sample of them. An inter-rater agreement rate of 88.9% was consequently achieved and any disagreements in scoring were each time resolved though discussions.

In addition to being available in French, other benefits of the Elicited Imitation Test (EIT) are that it is quick to administer, and it has versions in different languages which can be useful for cross-linguistic comparisons. Studies using EI in different languages have shown positive correlations with other proficiency measures, such as end-of-year grades, oral interview performance, and speech rate in oral picture-based narrative. For instance, in L2 Chinese Mandarin, higher EI scores correlated positively with oral performance which tapped elements of Complexity, Accuracy and Fluency (CAF) (Wu and Ortega, 2013). In L2 Korean, Kim, Tracy-Ventura and Jung (2016) reported a strong correlation between EI scores and CAF measures, the strongest of the three being between EI scores and fluency. Finally, the high reliability of Elicited Imitation scores has been confirmed in many recent studies with Cronbach's coefficients of .92 (Tracy-Ventura et al, 2014), .96 (Kim et al, 2016) and .97 (Wu and Ortega, 2013), making it a reliable method to test L2 proficiency.

4.6.4 Motivation measurement and language learning background questionnaire

To measure the motivation of participants, a questionnaire adapted from Taguchi et al (2009) was used, itself adapted from Dörnyei (2005). The questionnaire contains six-point Likert-scales and question-type items assessed from "not at all" on the left to "very much" on the right of the six-point rating scales. Taguchi et al's (2009) three versions in Japanese, Chinese and Iranian were designed for 1,586 learners in Japan, 1,328 in China and 2,029 in Iran. The total number of questionnaire items was 67 for the Japanese and Chinese versions and 76 for the Iranian version. In the version that was designed for the present study (Appendix K), 37 questionnaire items were used, investigating aspects of the L2 Self Motivation system described in the previous chapter such as: criterion measures, Ideal L2 Self, Ought-to L2 self, parental encouragement/family

influence, linguistic confidence, instrumentality, attitudes towards learning French, travel orientation interest in the French language, French anxiety, integrativeness, cultural interest and attitudes toward L2 community.

In addition to the motivation questionnaire, an expanded version of the background questionnaire was designed to find out more about the participants themselves and their L2 French learning history (Appendix L), asking a series of factual, behavioural, and attitudinal questions that are described below. First of all, the factual questions concerned participants' age, their nationality and their degree programme. They were also asked whether a language other than English was spoken at home and if yes, whether they had taken an exam in this language, the type of activities they would undertake in this language, such as reading, watching TV and how often they engaged in these types of activities. Participants were also asked whether they studied another foreign language and to what level. Further questions then dealt with participants' L2 learning history, asking them to provide information about the amount of time they had learned L2 French for, and their GCSE and A-level grades in French as well as some information about whether they had spent any time in the L2 environment and if so how long for, when and in what circumstances. This background information was included to interpret the findings of the motivation questionnaire as well as the vocabulary and proficiency tests.

Two behavioural questions were then asked, one about the activities participants usually undertake to practise their L2 and the other one about the activities participants undertake to learn L2 vocabulary in particular. It is known that open-ended questions have their limitations in L2 research (Dörnyei and Taguchi, 2010: 36) for instance, participants can be less engaged and give short and not very useful answers. To minimise this issue, examples of language and vocabulary development activities such as watching DVDs, participating in tandem exchanges with a native speaker etc. were listed. On the day participants completed the questionnaire, they were also reminded and encouraged to develop their answers to the open-ended questions. These qualitative data about independent activities to develop their language in general and vocabulary in particular were then coded according to Nation's (2007) four strands: meaning-focused input, meaning-focused output, language-focused learning, and fluency development. Each of these strands provide different learning opportunities and even though there is no evidence that activities should be divided equally across these four strands, it is widely accepted that each one of them is important for L2 development (Webb and Nation, 2017: 191). As we will see in Chapter 5, this categorisation is not perfect but it has the advantage to yield some information about the participants' exposure to L2 French outside the classroom. Also, "the principle of the four strands

states that, for a vocabulary learning programme to be well balanced, the time spent both inside and outside the classroom should be divided equally between the strands" (Webb and Nation, 2017: 179). Therefore, this categorisation will enable us to investigate the type of input and to determine whether there are some noticeable imbalance between the strands.

Four further short open-ended questions related to the participants' experience, likes, and dislikes about L2 French at university concluded the questionnaire. The aims in adding these questions were, firstly to allow participants to express personal views because "respondents often like to have an opportunity to express their opinions more freely and may find it frustrating to be completely limited to choosing from ready-made option" (Dörnyei and Taguchi, 2010, 37). Secondly, the questions were planned in combination with the motivation questionnaire and added to seek possible exploratory analysis for the motivation quantitative data.

4.7 Conclusion

This chapter has outlined the research design and described the research procedure. A triangulation approach was adopted in order to fill a gap in the literature as lexical development is not empirically studied in relation to both proficiency and motivation. The validity and reliability of the X-Lex, Flex15 and EI tests have been established in prior studies and will be reviewed in the next chapter on data analysis.

Chapter 5 Results

5.1 Introduction

In this chapter, an analysis of the research data is presented in relation to the research questions which are reiterated. Quantitative data collected from the X-Lex, Flex15 and Elicited Imitation tests are examined, followed by the data from the motivation questionnaire. The qualitative results about exposure to L2 French are then analysed.

5.2 X-Lex tests results

5.2.1 Objectives

In this section the results of X-Lex tests, the Yes/No receptive vocabulary test will be described and analysed. The objectives in administrating these tests to first year undergraduate students (UG1), second year students (UG2) and finalists (UG3) both at the beginning and at the end of the academic year are to explore the first two research questions posed in Chapter 1: what are the rate and the variation of lexical development in L2 French over a 3 year period of instructed learning? And is the frequency factor observable in L2 learners' lexical development? With X-Lex test, it is the receptive vocabulary development which is the focus of this section and as such, the following sub questions are also considered:

- What is the students' vocabulary size? How does it compare to vocabulary size findings from previous studies?
- What is the lexical profile of undergraduate students of French?
- How does the receptive vocabulary of students develop over an academic year?
- How do the three groups of students compare to each other in terms of vocabulary size, vocabulary development and lexical profile?

5.2.2 Participants and methods

40 students completed the X-Lex November test (UG1*,* *n*= 12; UG2, *n*= 13; UG3, *n*=15) and 35 of them completed the X-Lex May test (UG1, *n*= *9*; UG2, *n*= 12; UG3, *n*=14). As a reminder, the X-Lex tests consisted of 100 words each across the 1000-5000 word frequency bands, from Baudot's (1992) frequency list, as well as 20 non-words. Students ticked words they recognised and 50

points were awarded for each real word checked, hence producing the raw score. 250 points were then deducted every time a non-word was checked, giving the adjusted score, reported in the results below. The first observation that can be made is that there was very hardly any over guessing from participants as most of them ticked less than 5 non-words every time they took the test. In fact, only participant UG307 for the May test ticked over 5 non-words. This performance is therefore considered as unreliable and was not included in the results below.

5.2.3 Overall receptive vocabulary size development

Using SPSS Statistics 24, descriptive statistics have been calculated for the complete group of students for the two tests and it was found that over the November to May period, vocabulary size has increased overall from a mean of 3147.06 words in November to a mean of 3397.06 in May. A paired-samples t-test was conducted to evaluate the increase in receptive vocabulary and the reported increase was significant from X-Lex November (M= 3147.06, SD=734.63) to X-Lex May (M=3397.06, SD=652.26), *t* (33) = 3.67, p < .001 (two-tailed). The mean increase from X-Lex November to X-Lex May is 250 with a 95% confidence interval ranging from 388.23 to 111.76. The eta-squared statistic (0.28) indicated a large effect size.

					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	X-LexNov adjusted score - X-LexMay adjusted score	250.00	396.194	67.947	388.239	111.761	3.679	33	.001

Table 5.1 X-Lex Nov and X-Lex May paired samples test

5.2.4 Receptive vocabulary size development per year group

Let's now turn to the three groups to investigate in more detail their lexical development over the course of the academic year.

Looking at Table 5.2, the first conclusion that we can draw is that, as expected, each groups' receptive vocabulary has developed from November to May. We can also notice that there is a continuous development from one year group to the other, that is to say the receptive vocabulary of UG1 is smaller than the vocabulary size of UG2 and the UG3 group has the biggest receptive vocabulary size. Table 5.2 also shows that the standard deviation for UG3 at the end of the academic year is smaller than for UG2 and UG3, showing that the receptive vocabulary of finalists is more homogenous than other groups'. We can hypothesise that this homogeneity is the result of instruction.

	X-Lex Nov			**X-Lex May**		
	UG1	**UG2**	**UG3**	**UG1**	**UG2**	**UG3**
Mean	2570.83	2996.15	3610.00	3044.44	3329.17	3703.85
Minimum	1750	1650	2650	1900	2400	3150
Maximum	3700	4450	4450	4250	4800	4200
Standard Deviation	646.83	787.78	577.00	763.39	704.03	356.75

Table 5.2 Means, minimums, maximums and standard deviations of X-Lex Nov and May tests per year group

If all groups' receptive vocabulary has developed over the academic year, T-tests conducted to evaluate the difference in means per year group show that this development has been different for each group. Indeed, UG2 is the only group displaying a statistically significant difference from

X-Lex November (M= 3037.50, SD = 807.95) to X-Lex May (M= 3329.17, SD = 704.03), t (11)= 2.57, p < .05. The mean increase in scores was 291.66 with a 95% confidence interval ranging from 541.01 to 42.31. The eta squared statistic (0.37) indicates a large effect size. Development between November and May for both UG1 and UG3 just fails to reach statistical significance, with a probability of 0.064 and 0.140 receptively. However, this may be a function of group size, i.e it is harder to show significance the smaller the group as UG1 have made the largest absolute gain (about 470 words).

5.2.5 Variance between groups for each X-Lex tests

Looking at minimums, maximums and mean scores from the November test (Table 5.2), we can notice that there are some similarities between groups, in particular between UG2 and UG1 on one the hand and between UG2 and UG3 on the other. For instance, with a minimum score of 1750, UG1 is similar to UG2 which has a minimum score of 1650; and UG2 and UG3 have the same maximum score of 4450. Hence the most significant difference seems to be found between UG1 and UG3, a conclusion which is confirmed by a one-way between-groups analysis of variance conducted to explore the impact of year groups (UG1, UG2 and UG3) on receptive vocabulary size as measured by the X-Lex November test. There was a statistically significant difference in the X-Lex November tests for the three groups: *F* (2, 37) = 8.19, p = .001. The effect size, calculated using eta squared, was 0.30 indicating a large effect size and post-hoc comparisons using the Tukey HSD test indicated that the mean score for UG1 (M = 2570.83, SD = 646.83) was significantly different from UG3 (M = 3610, SD = 577). UG2 (M = 2996.15, SD = 787.78) did not differ significantly from either UG1 or UG3. The one-way between-groups analysis of variance was also conducted for the X-Lex May test but no significant difference was found (p = 0.55). This reflects the greater gains made by the lower groups, reducing the overall spread of mean scores.

5.2.6 Lexical profiles

Let's now explore the X-Lex November and May test scores of the three year groups by frequency bands in order to determine their lexical profile.

Figure 5.1 below illustrates students' scores per frequency band for the November test. As explained in Chapter 4, 1K represents the first 1000 most common words in French from Baudot

(1992)'s frequency list, 2K, the first 2000 words and so on until the 5k band. In Figure 5.1, the Y axis is therefore out of 1000. In the X-Lex test, words were ordered in five columns, each column representing a frequency band. The first column was therefore made up of words from the 1K frequency band, followed by the column containing words from the 2K frequency band and so on until the last column containing words from the 5K frequency band. As we can see below, a frequency factor can be observed as proportionally more high frequency words are recognised than low frequency words. In fact, all but the 2K band follow a regular frequency pattern. The three year groups seem to be "level 2 deficit learners" a term given by Milton to students whose knowledge of the second most frequent 1000 words is less high than anticipated (Milton, 2009). This apparent deficit is particularly noticeable for UG1 and UG3.

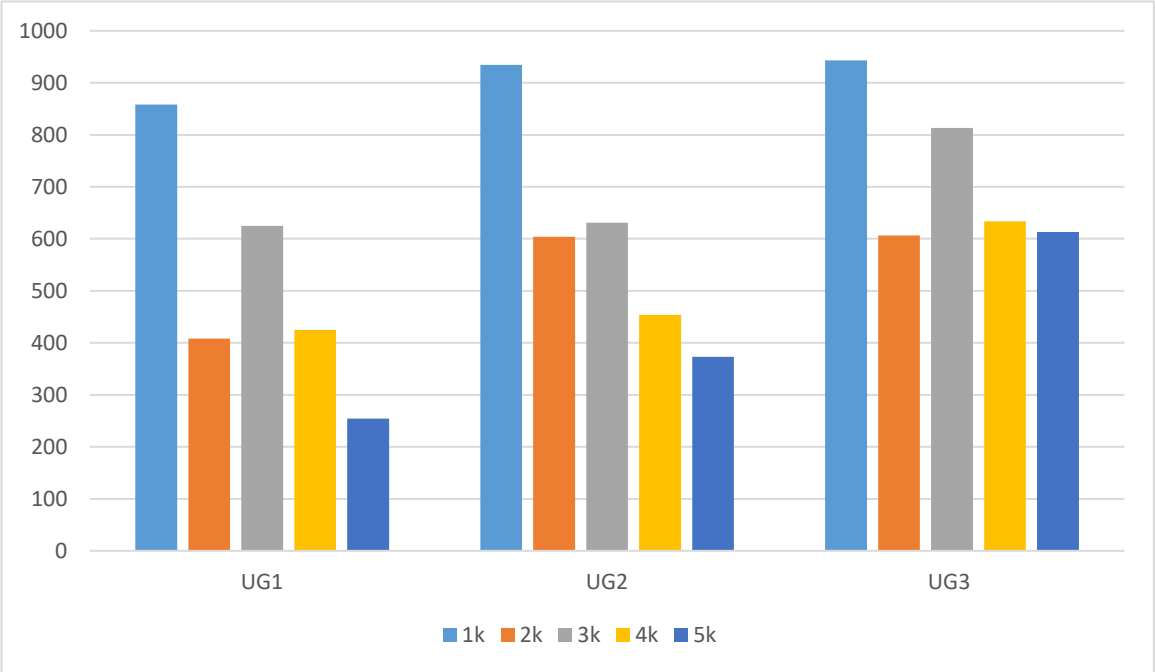


Figure 5.1 UG1, UG2 and UG3 lexical profiles for X-Lex Nov test

Figure 5.2 represents the X-Lex May scores per frequency band and we can notice that by the end of the academic year the lexical profile of the three year groups is more regular with the means of words in each bands decreasing as the frequency of bands themselves increase, and the effective disappearance of any "level 2 deficit" can be noted. Therefore, the X-Lex test provide valuable evidence that the lexical profile of students can change over a short period.

For first year students, the mean of all bands has increased except for the 3K band that went down from 625 to 544.44.

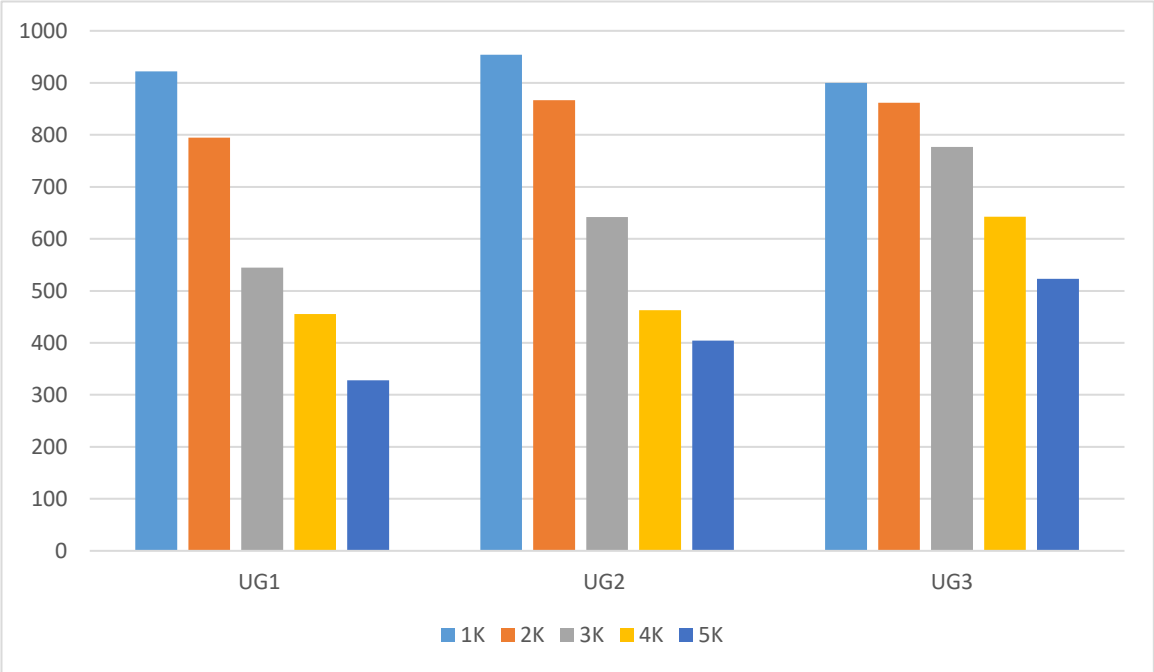


Figure 5.2 UG1, UG2 and UG3 lexical profiles for X-Lex May test

For second year students, the mean of all bands without exception increased from November to May. As for final year students, they are the group with the most irregular change in lexical profile with an increase in mean for 2K and 4K bands from November to May but a decrease in 1K, 3K and 5K bands. Despite such minor variations in lexical profile developments over the academic year, as previously mentioned, the total of means from November to May has increased for UG1, UG2 and UG3 and there is also continuous progress between UG1, UG2 and UG3 that is noticeable in both tests.

5.2.7 Lexical profiles beyond the 5K band

Given that students are high achievers in the language and that some of their X-Lex November scores were at ceiling level, it seemed appropriate to investigate their receptive vocabulary size beyond the limit of the 5K frequency band. Therefore, in the second X-Lex test, students were also tested on the 6K and 7K frequency bands. For some reasons, not all students completed this part of the test and one, UG307, was disregarded because, as previously explained, the student had over guessed. Full data for X-Lex May are available for therefore slightly smaller numbers of participants (UG1, *n*= 9 ; UG2, *n*= 12; UG3, *n*=11). All three groups seem to know more words in the 7K band than in the 6K band so the frequency factor is not observable between those higher frequency bands. We can note too that the UG1's knowledge of 7K words is better than that of the UG2 group. But the difference between these two bands is small for all groups and an ANOVA test of between-group differences was non-significant.

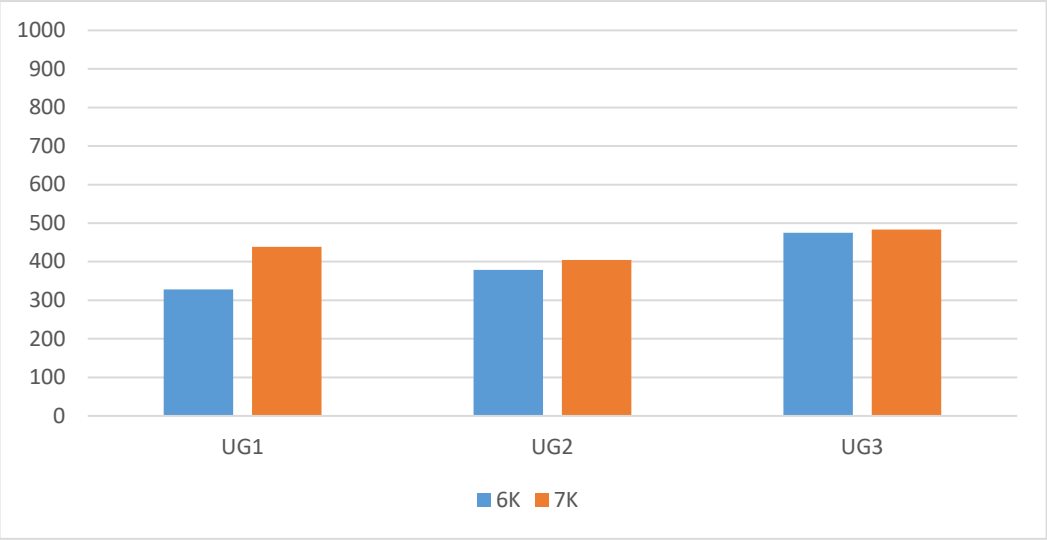


Figure 5.3 UG1, UG2 and UG3 means for 6K and 7K bands

	Mean for 6K frequency band	Mean for 7K frequency band	Total
UG1	327.77	438.88	766.65
UG2	379.16	404.16	783.32
UG3	475	483.33	958.33

Table 5.3 UG1, UG2 and UG3 means for 6K and 7K bands

In Figure 5.4, year groups' mean scores for all seven frequency bands included in the X-Lex May test are presented. As previously mentioned, the frequency factor is observable up to the 5K band only, for all three groups.

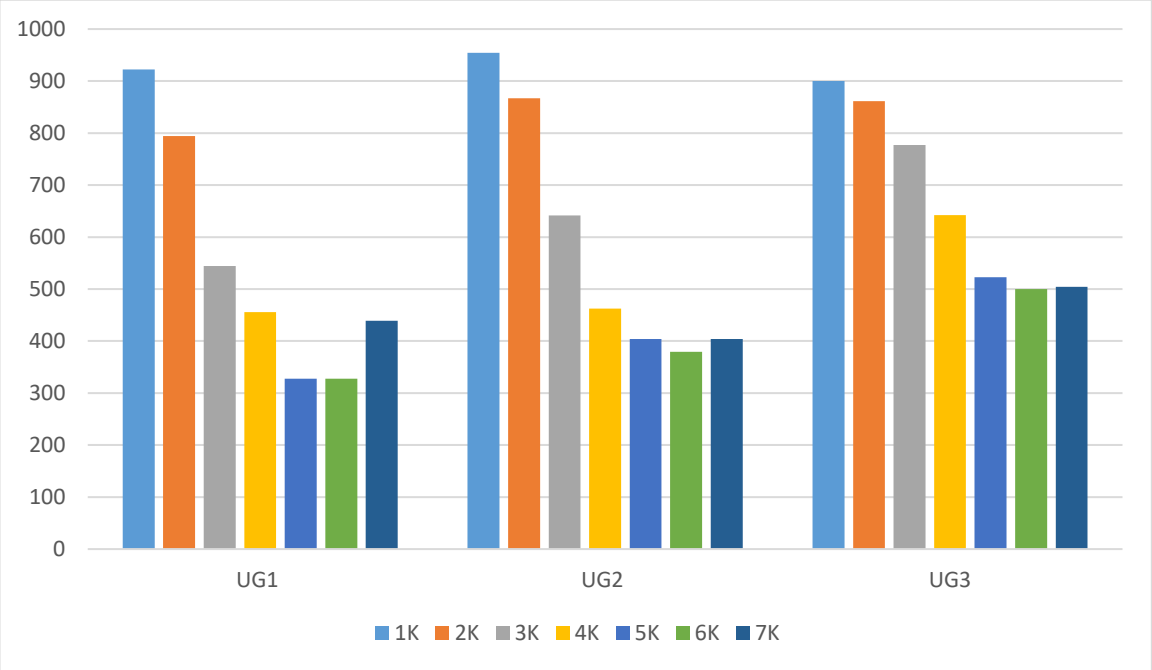


Figure 5.4 UG1, UG2 and UG3 lexical profile for 1K-7K bands

5.2.8 Summary of X-Lex tests results

The results of the X-Lex tests show that the vocabulary size of the students is generally in line with that found in previous studies. The November results in particular are quite close to David's results which she obtained in the middle of the academic year as shown in Table 5.4 below.

	UG1	**UG2**	**UG3**
David (2008)	2524	2854	3359
Pignot-Shahov (2018)	2570.83	2996.15	3610

Table 5.4 L2 French undergraduates' receptive vocabulary scores

This demonstrates that to some extent vocabulary size can be useful to differentiate different learning levels such as different years of studies in a degree programme. Nevertheless, as found in other studies when measuring receptive vocabulary, this study also shows that there is a great deal of individual variation within groups. For instance in the X-Lex May test, a UG1 student, with a mean score of 4250 has a marginally larger receptive vocabulary knowledge than a UG3 student who scored a maximum score of 4200. So if the receptive vocabulary size can be an indicator of the level of students by no means it should be considered as a unique and reliable placement test for instance.

What the present study also shows is that the vocabulary of students does develop over the course of an academic year and that the X-Lex test is sensitive enough to illustrate this change. Only the UG2 group gains statistically significantly more receptive vocabulary knowledge from November to May, but it seems that this was perhaps a function of group size (the UG1 gain of c450 points is the largest in absolute terms). The study also shows that the groups became more similar over time, so that in the X-Lex May test, there was no statistically significant difference at all between the groups. Indeed, with a less than 100 word gain between November and May the UG3 display hardly any receptive vocabulary development over the academic year. So

administering X-Lex tests twice over one academic year is useful to demonstrate that the vocabulary size does develop over a relatively short period of time, but at different rates depending on group.

The X-Lex results in relation to the lexical profile of students over the academic year, provide further useful information on vocabulary growth. Indeed, as we have reported above, in the X-Lex November test, all groups seem to show a level 2 deficit, that is to say that their knowledge of words in the 2K band is lower than anticipated. For Milton (2009), who also reports some level 2 deficit data from 21 Greek learners (Milton, 2006), these results could be due to differences in aptitude such as memory. Indeed, in his study, Milton found that the Greek students who were level 2 deficit tended to score higher on the LAT_B memory test. But the X-Lex May results show that by the end of the academic year, the apparent deficit has disappeared and all groups have a regular lexical profile. Whatever may be influencing vocabulary gain from different frequency bands, whether it is incidental exposure from L2 input, memory as per Milton or confidence as investigated by Kamimoto (2005), or motivation as we will investigate in this study, it can lead to fast changes in students' lexical profile.

In many studies measuring L2 receptive vocabulary, participants are school pupils and as such are not tested beyond the limit of the 5k band (Milton, 2006; David 2008; Richards and Malvern, 2007). And as scoring between 4500-5000 in the X-Lex test corresponds to level CEFR C2, it is very unlikely school pupils would gain such a score. However, more advanced learners such as our participants, should be able to score around 4500 and it seemed therefore appropriate to extend the test to the 6K and 7K frequency bands. The two points we can make about the data collected about the 6K and 7K bands is that, first, the frequency factor is not observable at these higher frequency levels and second, the two levels are almost flattened. This corroborates Aizawa's (2006) findings who conclude in the study of 363 Japanese learners of English at university in Tokyo that "after about the 5000 word frequency level, the variation between levels is not only slightly inconsistent, but also too small to be statistically significant" (in Milton, 2009: 28). But even if the frequency factor is not observable beyond the 5K bands, it is nevertheless a positive result as it shows that advanced learners' vocabulary develops into low frequency bands, perhaps as a result of intentional learning of specialised vocabulary domains.

Also, according to Milton the frequency factor is "a tendency and not an absolutely rigid rule" because "a more frequent word will not inevitably be learned before a less frequent one" as we have seen in the results of the two X-Lex tests. What is interesting though is that the lexical profile of the three groups is similar when taking the tests. In November, they all show a level 2

deficit and in May they all have a regular lexical profile. Yet, the uptake within each frequency band differs, UG1 gains in all but the 3K bands, UG2 gains in all bands and UG3 increases in 2K and 4K bands but decreases in 1K, 3K and 5K bands. Once again, the relatively small sample sizes within each year group may explain this apparent variability.

The UG3 group is the only one in which a participant's test was disregarded because more than 5 non-words were ticked. This is somehow intriguing as the test was taken seriously and it is unlikely the student was trying to guess. David also notices "the final-year undergraduate group was the one with the largest proportion of high false alarms"(2008: 178) and that there was no obvious reason why this should be.

5.3 Flex15 tests results

5.3.1 Objectives

In this section, the results of the free productive task Flex15 inspired by Meara and Fitzpatrick's (2000) Lex30 productive vocabulary test, are described and analysed. The aims in administering this productive vocabulary test both at the beginning and at the end of the academic year are the same as the aims stated in the previous section. The two main questions that are investigated are: what are the rate and the variation of lexical development in L2 French over a 3 year period of instructed learning? And is the frequency factor observable in L2 learners' lexical development? In trying to address these two points, the objectives are also to:

- Measure the L2 productive vocabulary of students using word frequency data;
- Determine whether any changes can be observed in their productive vocabulary in terms of frequency profiles, over one academic year;
- Determine whether the three groups of students differ from each other in terms of productive vocabulary frequency profiles;
- Determine if there is any relationship between the participants' receptive and productive vocabulary.

5.3.2 Participants and methods

40 students completed the productive vocabulary test in November, which will be referred to as Flex15 Nov (UG1*, n*= 12; UG2, *n*= 13; UG3, *n*=15) and 34 students completed the Flex15 May test (UG1, *n*= *9*; UG2, *n*= 11; UG3, *n*=14). The Flex15 Nov task consisted of 15 stimulus words (13

nouns, 1 verb and 1 adjective). These fifteen stimulus words were selected from Baudot's frequency list, and are mostly high frequency words. (Three were taken from the first 1K frequency band, eight from the 2K frequency band, one from the 3K band, one from the 4K band and 2 from the 5K band). All words were chosen carefully so as to avoid stimuli which would trigger one obvious answer. In selecting the few mid-frequency and low frequency words, efforts were made to select lexical items that would be known to participants despite being non frequent words. However, as we can see in Tables 5.5 and 5.6, there are quite a few cognates in both Flex15 Nov and Flex15 May tests. This was not a deliberate decision but a random result of selection factors of lower frequency words that might still be known to participants and that can elicit a variety of responses.

Stimulus words	**Position in Baudot's frequency list**
Voiture (car)	538
Demander (to ask)	128
Ciel (sky)	934
Photo (picture)	1730
Fête (party)	1187
Jardin (garden)	1403
Espoir (hope)	1246
Carte (card)	1260
Mémoire (memory)	1880
Vacances (holidays)	1677
Futur (future)	1363
Assiette (plate)	2818
Stupide (stupid)	3651
Ordinateur (computer)	4428
Crayon (pencil)	4939

Table 5.5. Frequency of stimulus words for Flex15 Nov test

For each stimulus words, participants were asked to write four words that they would associate with them. The task took 10 minutes to complete and was repeated in May with different stimulus words. The 15 stimulus words for the Flex15 May test included 11 nouns, 2 verbs and 2 adjectives and were also selected from different frequency bands based on Baudot's list. The majority of them were also high frequency words, with 8 taken from the first 1K frequency band, 3 from the 2K frequency band, 1 from the 3K band, 2 from the 4K band and 1 from the 5K band.

Stimulus words	**Position in Baudot's frequency list**
Pouvoir (to be able to, power)	37 (v), 435 (n)
Savoir (to know, knowledge)	65 (v), 4160 (n)
Maison (house)	201
Journal (newspaper)	403
Argent (money, silver)	487
Mer (sea)	492
Radio (radio)	787
Rapide (fast)	818
Oreille (ear)	1113
Tradition (tradition)	1137
Pluie (rain)	1768
Lundi (Monday)	2430
Tristesse (sadness)	3371
Football (football)	3600
Loyal (loyal)	6135

Table 5.6 Frequency of stimulus words for Flex15 May test

Each participants' responses were lemmatised according to the French lemmatization list produced by Měchura and then available on lexiconista.com. (The resource is unfortunately not available anymore.) The list itself is based on "Lexique" by New and Pallier (2003), a database of 135 000 French words for which information about frequencies, lemmas, grammatical categories is given. Once lemmatized, the frequency of each answers was then checked against Baudot's list. Every answer was then scored one point and points were added up for level 1, level 2, level 3, level 4 and level 5 and beyond frequency bands. Non-words, proper nouns and numbers were labelled level 0 and were not scored. Words that were not found in either Baudot's frequency list or in Lonsdale and Le Bras (2009)'s frequency dictionary were categorized as *not found*.

5.3.3 Omissions

Before analysing participants' responses to the productive tests, let's first turn to their omissions. For the Flex15 Nov test, out of a possible total number of responses of 720, there are 35 omissions for UG1, that is to say 4.86% of their overall responses. For UG2, omissions represent 2.69% with 21 omissions out of 780 possible answers. The omission rate of UG3 is drastically smaller with only two omissions out of a possible 900 answers, representing 0.22% of UG3 overall answers. The decline in omission rates across the three year groups seems to suggest that more proficient learners have more productive vocabulary at their disposal than less proficient learners. This supports Levelt's model in which words are central to proficiency.

Interestingly though, by May, UG1 and UG2's omission rates fall to 0.74% and 0.55% respectively. Not only do the two groups have by then a similar negligible omission rate but they are also producing more words than UG3 which, with a total of 14 omissions out of a possible 840 possible responses, has an omission rate of 1.66%. Having said that, 13 out of the 14 UG3 omissions are attributed to one student only. If this participant's responses were dismissed for having such a high rate of omissions, the UG3 omission rate would then fall to 0.11% hence showing again a small proficiency effect in participants' ability to produce words. But the difference in omissions between the three groups is much less important by the end of the academic year than for the Flex15 Nov test.

5.3.4 Overall productive vocabulary development

Table 5.7 and Table 5.8 provide descriptive statistics for the November and May tests, showing mean numbers of responses produced and standard deviations, for all groups.

		N	Minimum	Maximum	Mean	Std. Deviation
1 UG1	Flex15NovL1-L5+	12	41	59	54.58	5.195
	Valid N (listwise)	12				
2 UG2	Flex15NovL1-L5+	13	48	60	56.77	3.444
	Valid N (listwise)	13				
3 UG3	Flex15NovL1-L5+	15	56	60	58.33	1.047
	Valid N (listwise)	15				

Table 5.7 Means, minimums, maximums and standard deviations of Flex15 Nov test

Year group		N	Minimum	Maximum	Mean	Std. Deviation
1 UG1	Flex15MayL1-L5+	9	52	60	57.67	2.646
	Valid N (listwise)	9				
2 UG2	Flex15MayL1-L5+	11	55	59	57.36	1.286
	Valid N (listwise)	11				
3 UG3	Flex15MayL1-L5+	14	46	60	57.21	3.556
	Valid N (listwise)	14				

Table 5.8 Means, minimums, maximums and standard deviations of Flex15 Nov test

Paired-samples t-tests were also conducted to evaluate the difference between the Flex15 Nov test and the Flex15 May test (see Table 5.9). This showed that there was indeed a statistical increase in the number of words produced by UG1 from November (*M*=56.11, *SD* =2.89) to May (*M* = 57.67, *SD* = 2.64), *t* (8) = 3.27, p < .001 (two-tailed). The mean increase in Flex15 scores for UG1 was 1.55 with a 95% confidence interval ranging from 2.65 to .46. There was no statistically

significant increase in the number of words produced from November to May for the UG2 and UG3 groups.

Year group			Paired Differences					t	df	Sig. (2-tailed)
			Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
						Lower	Upper			
1 UG1	Pair 1	Flex15NovL1-L5+ - Flex15MayL1-L5+	1.556	1.424	.475	2.650	.461	3.277	8	.011
2 UG2	Pair 1	Flex15NovL1-L5+ - Flex15MayL1-L5+	.818	4.119	1.242	3.585	1.949	.659	10	.525
3 UG3	Pair 1	Flex15NovL1-L5+ - Flex15MayL1-L5+	1.071	3.385	.905	.883	3.026	1.184	13	.257

Table 5.9 UG1, UG2 and UG3 paired differences between Flex15 Nov and Flex15 May tests

5.3.5 Differences in productive vocabulary responses between year groups

A one-way between-groups analysis of variance was also conducted to explore productive statistical differences between the three groups for the Flex15 Nov test, and it showed that there was a statistically significant difference at the $p < .05$ level. F (2, 37) = 3.81, p = .031. Despite reaching statistical significance, the actual difference in mean scores between the groups was quite small. The effect size, calculated using eta squared, was 0.17. Post-hoc comparisons using the Tukey HSD test indicated that the mean score of UG1 (M = 54.58, S = 5.19) was significantly different from UG3 (M = 58.33 , S =1.04). UG2 (M = 56.77, S = 3.44) did not differ significantly from either UG1 or UG3. It has to be noted though that with a significance value of .031, the assumption of homogeneity of variance is violated. Looking at standard deviation, it seems that UG3 has a different variance 1.04 compared to UG1, 5.19 and UG2, 3.44. So a Welch test was run p (.040).

Flex15NovL1-L5+

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	93.842	2	46.921	3.819	.031
Within Groups	454.558	37	12.285		
Total	548.400	39			

Table 5.10. Flex15 Nov one-way between-groups analysis of variance

A one-way between-groups analysis of variance was conducted to explore productive vocabulary as measured by the Flex15 May test but as already anticipated from the descriptive statistics analysis, no statistically important differences were found between any of the three groups.

5.3.6 Changing profiles

Let's now turn to the frequency of participants' responses. As already mentioned, words falling into the Level 0 category (non-words, proper nouns, numbers and not found words) were excluded. All other words were checked against Baudot's frequency list and allocated to a

frequency level or to a “Not found” category, as shown in Table 5.11. below for participant UG201 at Flex15 Nov entry.

ID	omit	L0	L1	L2	L3	L4	L5	Not found	Total
UG 201	1	2	29	11	7	5	3	2	60

Table 5.11 Flex15 Nov frequencies for UG201

In order to investigate changing frequency patterns over time, level 1 and level 2 scores have been grouped together as it is traditionally agreed that these first two bands represent high frequency words. Scores for Levels 3, 4, 5 and beyond have also been grouped and represent middle and low frequency words. T-tests were conducted to evaluate the relative frequency of L1/L2 and L3/L4/L5 + from November to May across the three-year groups. Looking at the descriptive statistics first (Table 5.7 and Table 5.8), we can see the number of L1/L2 responses produced increased from November to May whereas the L3/L4/L5 + responses decreased over the same period. Statistically, the increase in L1/L2 responses is significant (Table 5.12) from November (M = 39.85, S =3.27) to May (M = 42.68, S =4.38), t (33) = 3.68, p < .001 (two-tailed). The mean increase in score was 2.82 with 95% confidence intervals ranging from 4.38 to 1.26. The eta squared statistic (0.29) indicated a large effect size.

Paired Samples Test									
		Paired Differences					T	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Flex15Nov L1, L2 – Flex15May L1, L2	2.824	4.469	.766	4.383	1.264	3.684	33	.001

Table 5.12 Flex15 Nov and Flex15 May paired samples test for Level 1 and Level 2

The decrease in productive L3/L4/L5+ vocabulary (Table 5.13) is also statistically significant from November (*M* =17.29, *SD* = 3.77) to May (*M* = 14.71, *SD* = 3.61), *t* (33) = 3.61, *p* < .001 (two-tailed). The mean decrease in scores was 2.58 with a 95% confidence interval ranging from 1.13 to 4.04. The eta squared statistic (0.28) indicated a large effect size.

Paired Samples Test									
		Paired Differences					T	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Flex15Nov L3, L4, L5 + - Flex15May L3, L4, L5 +	2.588	4.179	.717	1.130	4.046	3.612	33	.001

Table 5.13 Flex15 Nov and Flex15 May paired samples test for Levels 3, 4, 5 +

Proportionally more L1/L2 words were produced in May by the cohort overall, but the increase is driven by UG2, with an increase in L1/L2 scores from November (*M* = 40.09, *SD* = 3.08) to May (*M* = 43.55, *SD* = 4.27), *t* (10) = 3.05, p < .001 (two-tailed). The mean increase was 3.45 with a 95% confidence interval ranging from 5.97 to .93. The eta squared statistic (0.48) indicated a large effect size. Increases for the other two groups were non-significant. However, looking more closely at the descriptive statistics (Table 5.14), we can also notice that the increase from November to May in L1/L2 productive vocabulary for UG1 is very similar to the increase for UG2. If the UG1 increase is not statistically significant, it is probably due to the small number of participants in this group.

Year group			Mean	N	Std. Deviation	Std. Error Mean
1 UG1	Pair 1	Flex15Nov L1, L2	40.44	9	3.909	1.303
		Flex15May L1, L2	43.44	9	3.812	1.271
2 UG2	Pair 1	Flex15Nov L1, L2	40.09	11	3.081	.929
		Flex15May L1, L2	43.55	11	4.275	1.289
3 UG3	Pair 1	Flex15Nov L1, L2	39.29	14	3.148	.841
		Flex15May L1, L2	41.50	14	4.832	1.291

Table 5.14 Level 1 and level 2 means and standard deviations for Flex15 Nov and Flex15 May per year group

Year group			Paired Differences					T	df	Sig. (2-tailed)
			Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
						Lower	Upper			
1 UG 1	Pair 1	Flex15Nov L1, L2 – Flex15 May L1, L2	-3.000	4.416	1.472	-6.394	.394	-2.038	8	.076
2 UG 2	Pair 1	Flex15Nov L1, L2 – Flex15 May L1, L2	-3.455	3.751	1.131	-5.975	-.934	-3.054	10	.012
3 UG 3	Pair 1	Flex15Nov L1, L2 – Flex15 May L1, L2	-2.214	5.206	1.391	-5.220	.792	-1.591	13	.136

Table 5.15 Flex15 Nov and Flex15 May paired samples test for level 1 and level 2 per year group

All three groups produce less L3/L4/L5+ vocabulary in May compared to November but this decrease is statistically significant for UG3 only with scores from November (*M* = 19, *SD* = 3.01) to May (*M* = 15.71, *SD* = 2.70), *t* (13) = 3.53, p < .001 (two-tailed). The mean decrease in scores was 3.28 with a 95% confidence interval ranging from 1.28 to 5.29. The eta squared statistic (0.48) indicated a large effect size.

Year group			Mean	N	Std. Deviation	Std. Error Mean
1 UG1	Pair 1	Flex15Nov L3, L4, L5 +	15.67	9	4.690	1.563
		Flex15May L3, L4, L5 +	14.22	9	3.598	1.199
2 UG2	Pair 1	Flex15Nov L3, L4, L5 +	16.45	11	3.236	.976
		Flex15May L3, L4, L5 +	13.82	11	4.557	1.374
3 UG3	Pair 1	Flex15Nov L3, L4, L5 +	19.00	14	3.013	.805
		Flex15May L3, L4, L5 +	15.71	14	2.701	.722

Table 5.16 Levels 3, 4, 5 + means and standard deviations for Flex15 Nov and Flex15 May per year group

Year group			Paired Differences					t	df	Sig. (2-tailed)
			Mean	Std. Dev	Std. Error Mean	95% Confidence Interval of the Difference				
						Lower	Upper			
1 UG1	Pair 1	Flex15Nov L3, L4, L5 + - Lex 30May L3, L4, L5 +	1.444	4.558	1.519	-2.059	4.948	.951	8	.370
2 UG2	Pair 1	Flex15Nov L3, L4, L5 + - Lex 30May L3, L4, L5 +	2.636	4.843	1.460	-.617	5.890	1.805	10	.101
3 UG3	Pair 1	Flex15Nov L3, L4, L5 + - Lex 30May L3, L4, L5 +	3.286	3.474	.928	1.280	5.291	3.539	13	.004

Table 5.17 Flex15 Nov and Flex15 May paired samples test for levels 3, 4 and 5+ per year group

5.3.7 Correlation between productive and receptive vocabulary

In order to determine whether a correlation can be established between the two receptive vocabulary tests and the two productive vocabulary tests, in terms of the number of items recognised/ produced, Pearson r was calculated and it seems that there is a medium correlation between receptive and productive vocabulary in November (r= .435) but not in May (Table 5.18 and Table 5.19). The lack of correlation between the two tests in May is certainly due to the fact that the cohort is at ceiling level in terms of number of responses produced in the Flex15 May test.

		X-LexNov adjusted score	Flex15NovL1-L5+
X-LexNov adjusted score	Pearson Correlation	1	.435**
	Sig. (2-tailed)		.005
	N	40	40
Flex15NovL1-L5+	Pearson Correlation	.435**	1
	Sig. (2-tailed)	.005	
	N	40	40
**. Correlation is significant at the 0.01 level (2-tailed).			

Table 5.18 X-Lex Nov and Flex15 Nov correlations

		X-LexMay adjusted score	Flex15MayL1-L5+
X-LexMay adjusted score	Pearson Correlation	1	.213
	Sig. (2-tailed)		.234
	N	34	33
Flex15MayL1-L5+	Pearson Correlation	.213	1
	Sig. (2-tailed)	.234	
	N	33	34

Table 5.19 X-Lex May and Flex15 May correlations

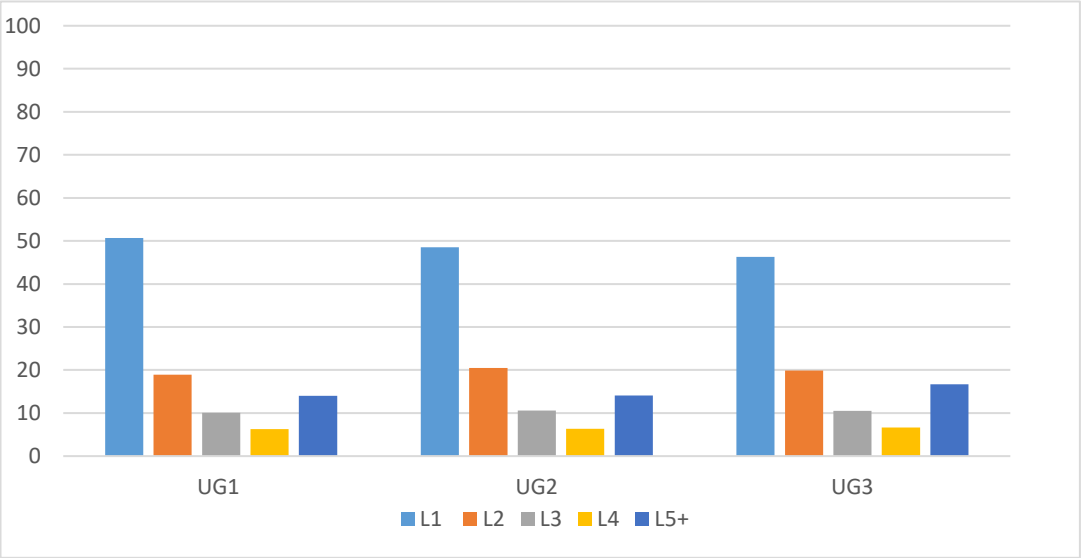
5.3.8 Descriptive analysis of Flex15 responses per frequency band

Let's now turn to a more detailed descriptive analysis of responses per individual frequency band. The first observation that can be made is that in both tests and for all three groups, responses are influenced by word frequency. Indeed, the percentage of responses tend to decrease as the frequency bands increase, as shown in Tables 5.20 and 5.21, and Graphs 5.1 and 5.2.

	L1	**L2**	**L3**	**L4**	**L5+**
UG1	50.68%	18.93%	10.07%	6.25%	14.04%
UG2	48.50%	20.46%	10.56%	6.36%	14.09%
UG3	46.28%	19.88%	10.51%	6.62%	16.68%

Table 5.20 Percentage of responses by frequency levels for Flex15 Nov

In Flex15 Nov, the percentages of L3 and L4 answers are very similar between the groups and L5 response rates are again similar between UG1 and UG2 (Table 5.20). However, UG3 produce slightly more L5 answers than the two groups.

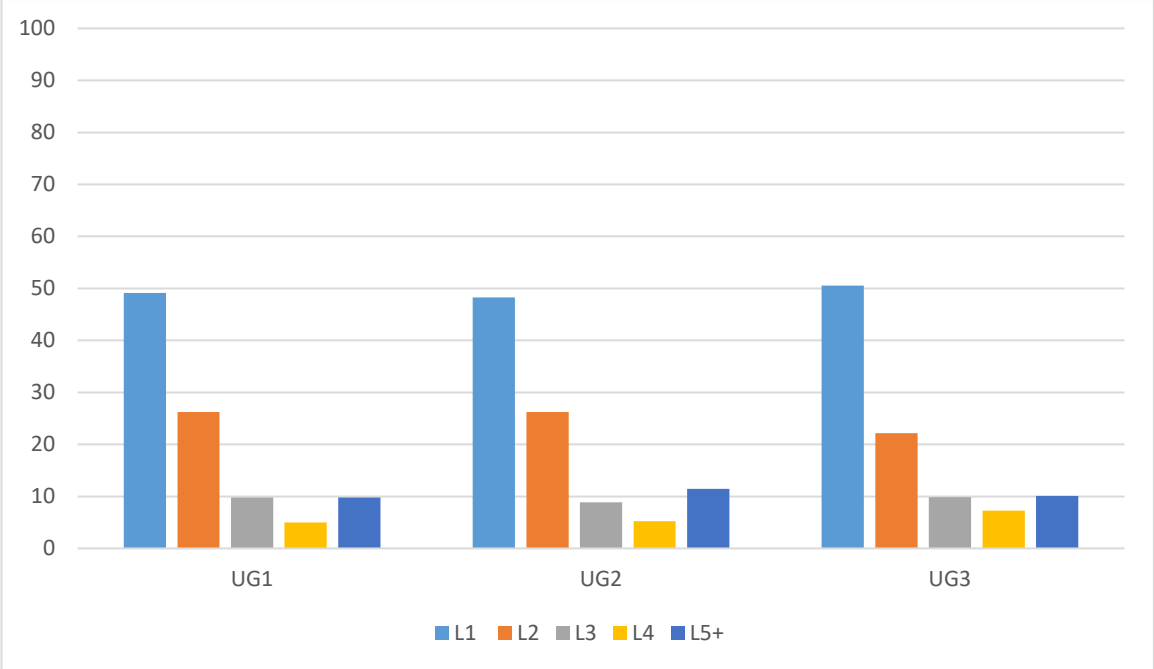


Graph 5.1 Percentage of responses by frequency levels for Flex15 Nov

At the end of the academic year, even though the frequency factor follows the same broad trend for all three groups, with a growth in Level 2 responses compensating for a decrease in responses at higher Levels, there is somewhat more variation between groups than observed in November (Table 5.21 and Graph 5.2).

	L1	**L2**	**L3**	**L4**	**L5+**
UG1	49.13%	26.20%	9.82%	5%	9.82%
UG2	48.26%	26.23%	8.84%	5.21%	11.44%
UG3	50.56%	22.15%	9.88%	7.25%	10.13%

Table 5.21. Percentage of responses by frequency levels for Flex15 May



Graph 5.2 Percentage of responses by frequency levels for Flex15 May

If globally most of UG1, UG2 and UG3's responses are high frequency words, what about the frequency of their first answers to the stimulus words? Analysis of Flex15 Nov shows that UG1 and UG2's frequency profiles for their first answers are very similar with 70.23% of UG1 answers falling in high frequency bands (1K and 2K combined) and 70.20% for UG2 (Table 5.22). However, only 65.68% of UG3 first answers fall in the L1 and L2 bands, suggesting that UG3 participants tend to produce slightly more low frequency answers as first answers. In some cases, 100% of first answers are high frequency words for UG1 and UG2. For example, the stimuli *jardin* (garden), *espoir* (hope) and *mémoire* (memory) have high frequency responses only for UG1, and *jardin* (garden), *futur* (future) and *demander* (to ask) have similarly high frequency responses for UG2. All these stimuli themselves are high frequency words. On the other hand, none of the stimulus words triggered 100% high frequency first responses for UG3.

	L1	**L2**	**L1/L2**
UG1	53.57%	16.66%	70.23%
UG2	54.25%	15.95%	70.20%
UG3	50.98%	14.70 %	65.68%

Table 5.22 First answers frequency bands in percentages for Flex15Nov

Most low frequency stimuli tend to elicit more low frequency words than high frequency stimuli do as shown in Table 5.23. However, not all low frequency stimuli trigger low frequency responses. For instance, the relatively low frequency cognate *footbal*l (3600), elicits 82.53% of high frequency answers (UG1= 100%; UG2= 83.33%; UG3 = 64.28%), as shown in Table 5.23.

Low frequency stimulus	**L1/L2 percentage responses overall**	**UG1 L1/L2 responses**	**UG2 L1/L2 responses**	**UG3 L1/L2 responses**
Lundi (2430)	56.42%	55.00%	50.00%	64.28%
Assiette (2818)	38.88%	30.00%	40.00%	46.66%
Tristesse (3371)	63.09%	66.66%	58.33%	64.28%
Stupide (3651)	47.11%	36.36%	58.33%	46.66%
Football (3600)	82.53%	100.00%	83.33%	64.28%
Ordinateur (4428)	16.66%	16.66%	33.32%	0.00%
Crayon (4939)	54.05%	75.00%	53.84%	33.33%

Table 5.23 Low frequency stimuli and their high frequency responses per year group

5.3.9 Summary of Flex15 test results

The Flex15 test results discussed above show the three year groups' productive knowledge has increased from November to May in terms of their ability to produce 4 response items for each test stimulus (though from an already high level in November, effectively to ceiling on this test in May). It is possible that this increase in the number of words produced in May compared to November could be attributed to a practice effect. However, Fitzpatrick and Clenton (2010) demonstrated the validity and the reliability of the Flex15 while testing 103 learners of English, from low to advanced level, and concluded that "Lex30 task elicits a data set which is broadly representative of the individual's lexicon" (2010: 541). If this is the case, we can then assume that all UG participants' productive vocabulary did improve over the year and particularly that of UG1, where the ceiling effect was less pronounced.

Flex15 is sensitive enough to show increase in productive vocabulary over a short period of time but it seems that the data can also be affected by participants who, for a reason or another, do not seem to engage with the word association task like UG309 who is responsible for 13 omissions out of 14 in the Flex15 May test. This raises the issue of whether a system of penalisation should be introduced. Similarly to the "over-estimation" penalty used for the receptive vocabulary X-Lex test, an "under-engagement" one could be put in place for participants with abnormal rate of omission.

We noted a medium correlation between the Flex15 November and the X-Lex November tests (r= .435) which is smaller than Meara and Fitzpatrick (2000)'s, who noted a strong correlation between the participants' receptive Yes/No vocabulary test and the Flex15: 0.841 (p <0.01). In the present study, there was no correlation at all between the two tests at the end of the academic year. As mentioned, this could be due to the cohort being at ceiling level in the Flex15 May test. Other possible reasons for the lack of correlation will be discussed in Chapter 6. A strong frequency effect is also observable. Not only do many first responses to fall into the most frequent 2,000 words, but the majority of words produced by all participants are high frequency words.

The present study also seems to confirm that Flex15 is an adequate method to tap into rarer words in an economic way. Indeed, Fitzpatrick and Clenton (2017) report three studies in which the Lex30 test was used and note that the proportion of infrequent words produced in these three studies (words beyond the 1K band) were 44%, 38% and 33%. These rates were considered successful as they were higher than other productive tests had reported. In the present study, the proportions of infrequent words (words beyond the 1K band) produced for the Flex15 May were 50.84% for UG1, 51.74% for UG2 and 49.44% for UG3. These results show that not only Flex15 is an adequate tool to elicit low frequency words but also that the participants' lexicon has developed into low frequency bands.

Interestingly too, the lexical frequency of responses changes in the same way over time for all groups: with more L1 and L2 words and fewer L3, L4 and L5+ word produced in May than in November. It is very unlikely that participants know fewer low frequency words in May than in November. The decline in rarer words might be the result of the choice of stimulus words (there are more Level 1 stimulus words in the May test), or it might be that the processing ability and/or the productive vocabulary of this cohort has developed in ways that are not captured by the Flex15 test. This point will be further discussed in Chapter 6.

5.4 Elicited Imitation test results

5.4.1 Objectives

In this section, the results of the Elicited Imitation Test are analysed. The aims in administering this L2 French proficiency test developed by Tracy-Ventura, McManus, Norris and Ortega, (2014) are to investigate Research Question 4 (3.5), i.e.

> To what extent does overall L2 proficiency explain the pattern of L2 vocabulary development over time?

Sub questions which will be considered are:

- What is the general level of French proficiency of the different year groups, and what is the relationship between these?
- What is the relationship between EI as a measure of general proficiency and other indices of language proficiency (X-Lex scores, Flex15 scores and end-of-year grades)?

5.4.2 Participants and methods

Forty students completed the Elicited Imitation test in the middle of the academic year (UG1, *n*= 12; UG2, *n*=13; UG3, *n*= 15). As a reminder, the test included 30 sentences ranging from 7 to 19 syllables in length, and ranked from the lowest to the highest number of syllables. Participants were played a recording of these sentences read at normal speech rate by a French native speaker and encouraged to repeat as much as they could. Each sentence was followed by a 2-second pause and then a beep signalling when the repetition should begin. Tests took about 9 minutes to complete. Participants' answers were then scored based on the five-point scoring system used by Tracy-Ventura et al (2014) (see Appendix M for an example).

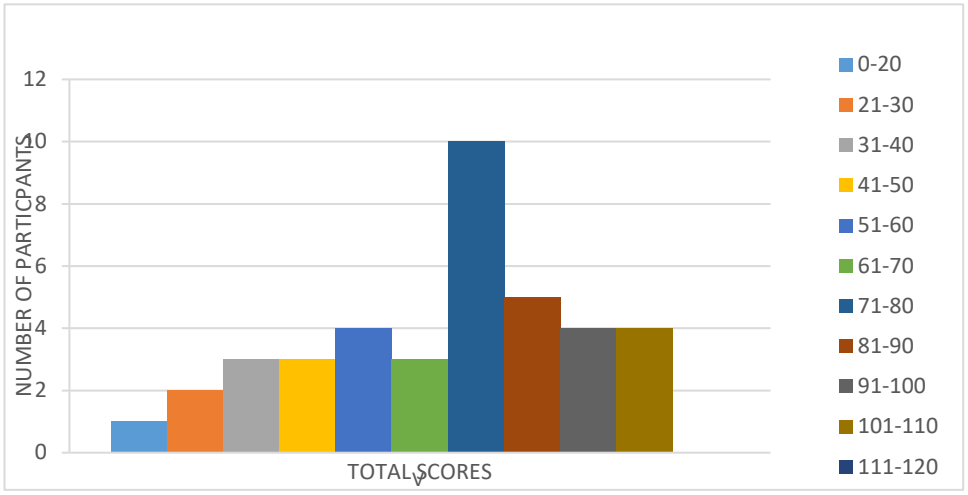
5.4.3 Overall EI test results

As we can see in Table 5.24 below, the EI mean score from the 40 L2 French participants is 69.70 (out of a possible 120 maximum score) with a standard deviation of 23.37. This 69.70 average is close to Tracy-Ventura et al's mean of 62.90 when testing 29 second year university students of L2 French (2014: 152). Therefore the EI tests seems to be a reliable tool to assess proficiency of advanced L2 French learners.

	N	Minimum	Maximum	Mean	Std. Deviation
EI TOTAL	40	16	107.00	69.70	23.37

Table 5.24 Mean, minimum, maximum and standard deviation for EI scores

The scores range from 16.00 to 107.00, showing not only that there are great individual differences between participants but also that with a score of 107.00, some participants' proficiency level is approaching native speakers' proficiency level with an average of 118.40 and range from 115.00 to 120.00 in the study of Tracy-Ventura et al (2014: 152). With 10 out 40 students having an EI mean score between 71.00 and 80.00, this is the most common band as we can see in Graph 5.3 below representing the score distribution. After this peak, there is a drastic decrease in the number of participants scoring in the higher bands 81-90 (5 students), 91-100 and 101-110 (4 students respectively).



Graph 5.3 EI scores distribution

5.4.4 EI test results per year group

Let's now turn to the EI scores per year group. As we can see in Table 5.25, the mean scores per year group increase from 56.00 for UG1, to 64.70 for UG2 and finally to 85.00 for UG3, suggesting, as one would expect, that proficiency improves with time and that second year students are more proficient than first year students but less proficient in their L2 than final year students, on the whole.

Looking at Table 5.25, we can also notice that the standard deviation for UG1 and UG2 is very similar at 23.06 and 23.64 but much lower for UG3: 13.77. As for the range of scores, for UG1 the minimum score is 16.00, 29.00 for UG2 and a much greater minimum of 66.00 for UG3. In terms of maximum scores, UG1 is 87.00, UG2 105.00 and UG3 107.00. So looking at descriptive statistics, it is clear that UG3 proficiency is better than that of UG1 and UG2. There seems to be more homogeneity as well within this UG3 group with a smaller range of scores and lower standard deviation. The difference between UG1 and UG3 in particular is striking with UG1 mean score (56.00) smaller than UG3 minimum (66.00). However, some UG2 participants score almost as highly on the EI test as some UG3 participants.

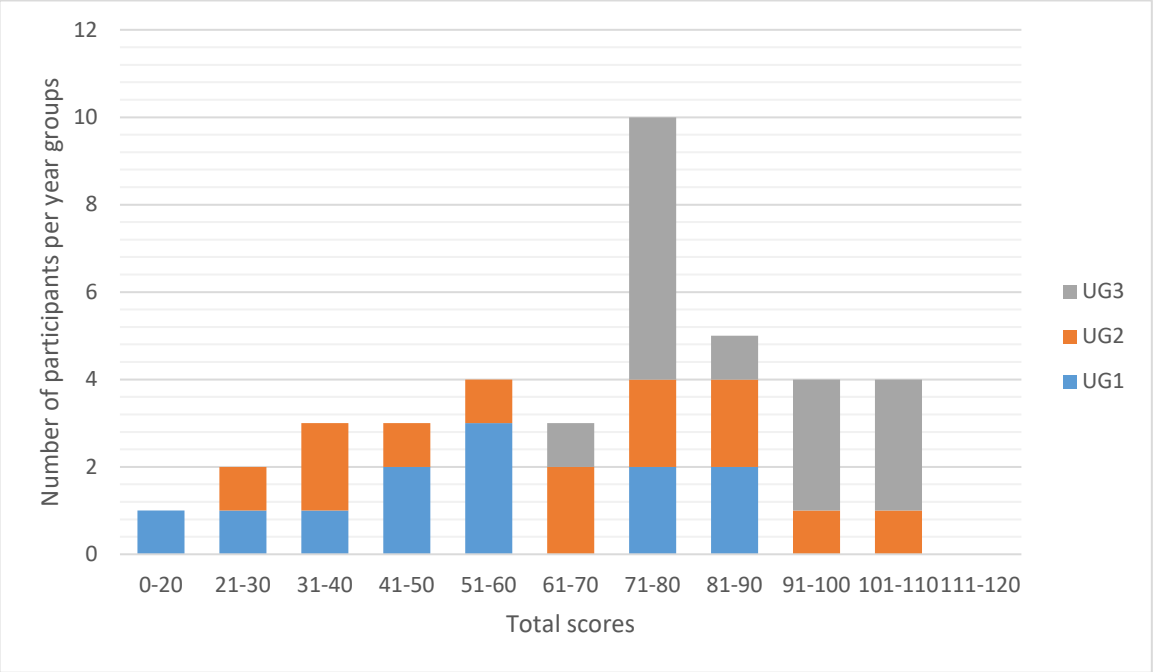
Year group		N	Minimum	Maximum	Mean	Std. Deviation
UG1		12	16	87	56.00	23.06
UG2		13	29	105	64.69	23.64
UG3		15	66	107	85.00	13.77

Table 5.25 Mean, minimum, maximum and standard deviation for EI scores per year group

The difference in proficiency between UG3 on one side and UG1 and UG2 on the other that can be observed from the descriptive statistics is also statistically significant. A one-way between groups analysis of variance was conducted to explore the EI results per year group and there was a statistically significant difference at the $p < .05$ level for the three year groups: F (2, 37) = 7.4, p = .002. The effect size (eta squared) was .28, showing a large effect size. Post-hoc comparisons using the Tukey HSD test indicated that the mean score of UG3 (M= 85, SD= 13.77) was

significantly different from UG1 (*M*= 56, *SD*= 23.06) and from UG2 (*M*=64.7, *SD*= 23.64). These results confirm that the proficiency level of the final year students who have spent the previous academic year in a French-speaking country, and likely exposed to large amounts of L2 input, is much higher than first and second year students'.

We saw earlier that the 71-80 score band was the most predominant one with 10 participants out of 40 falling into it. However we can see in Graph 5.4 below that most of these students are actually final year students. The graph also illustrates well what has been discussed when investigating the descriptive data, such as the fact that UG3 students do not score below 60 out of 120. The graph also shows that UG1 and UG2 tend to share the same score bands at lower levels, for instance 21-30, 31-40 and 51-60, while UG2 shares higher scores with UG3 91-100, 101-110 but also 61-70, a band which no UG1 participants fell into. The 71-80 and 81-90 bands are shared by all three groups and as for the 0-20 band, only one UG1 participant scored at this level.



Graph 5.4 EI scores distribution per year group

The graph also highlights the great individual differences in proficiency amongst these learners, especially at the lower levels, as we have seen. For instance, while one UG1 participant scored between 0-20, at the lower end of the scale, two others scored between 81-90, that is to say beyond the UG1 mean score.

5.4.5 EIT scores and other variables

Let's now turn to other variables and investigate how they relate to the EI scores described above. To do so, the relationship between EI scores, end-of-year exam grades, X-Lex scores and Flex15 scores have been investigated using Pearson product-moment correlation coefficient. As we can see in Table 5.26 below, that there is a strong positive correlation between EI scores and end-of-year exam grades *r =.* 64 and between EI scores and X-Lex *r* = .69; both of these correlations are statistically significant. However there is a weaker correlation between EI scores and Flex15 scores *r*= .29, which just fails to achieve statistical significance. A strong correlation between EI scores and end-of-year exam grades has already been observed before; for instance Tracy-Ventura et al (2014) report a correlation of *r*= .78 between these two variables. However, the correlation between X-Lex and EI scores in this study is much stronger than in Tracy-Ventura et al (r= .12). The correlation between receptive vocabulary size and proficiency in the present study supports the link between vocabulary and proficiency in Levelt's model as he places vocabulary/knowledge of lemmas at the centre of the linguistic system, with grammar deriving from this. As Nation's puts it, "there is undoubtedly a relationship between frequency of occurrence of individual items and the role of system knowledge in their use" (2010: 59). This also shows that Elicited Imitation test does not tap only into speaking-related knowledge. In both this study and Tracy-Ventura's though, the correlation between active vocabulary tasks and EI scores is relatively low with *r*= .30 and *r* =. 32 respectively.

		end of year exam grade	EI scores	X-LexNov	Flex15Nov
end of year exam grade	Pearson Corr.	1	.639**	.465**	.332*
	Sig. (2-tailed)		.000	.002	.036
	N	40	40	40	40
EI scores	Pearson Corr.	.639**	1	.693**	.296
	Sig. (2-tailed)	.000		.000	.064
	N	40	40	40	40
XlexNov	Pearson Correlation	.465**	.693**	1	.341*
	Sig. (2-tailed)	.002	.000		.031
	N	40	40	40	40
Flex15Nov	Pearson Corr.	.332*	.296	.341*	1
	Sig. (2-tailed)	.036	.064	.031	
	N	40	40	40	40

Table 5.26 Correlations between EI scores, end-of-year exam scores, X-Lex scores and Flex15 scores

When looking at these correlations again but per year group, we can see that the correlations between EI scores and end-of-year exam grades and EI scores and X-Lex are less strong for UG3 than for UG1 and UG2, but this might be due to the small group sizes or perhaps advanced proficiency is a bit less vocabulary-centred than lower level proficiency.

	End-of-year exam grades	**X-LexNov scores**	**Flex15 Nov scores**
UG1 EI scores	.722**	.586*	.386
UG2 EI scores	.837**	.774**	.058
UG3 EI scores	.492	.361	.327

Table 5.27 Correlations between EI scores, end-of-year exams, X-Lex scores and Flex15 scores per year group

5.4.6 Item difficulty

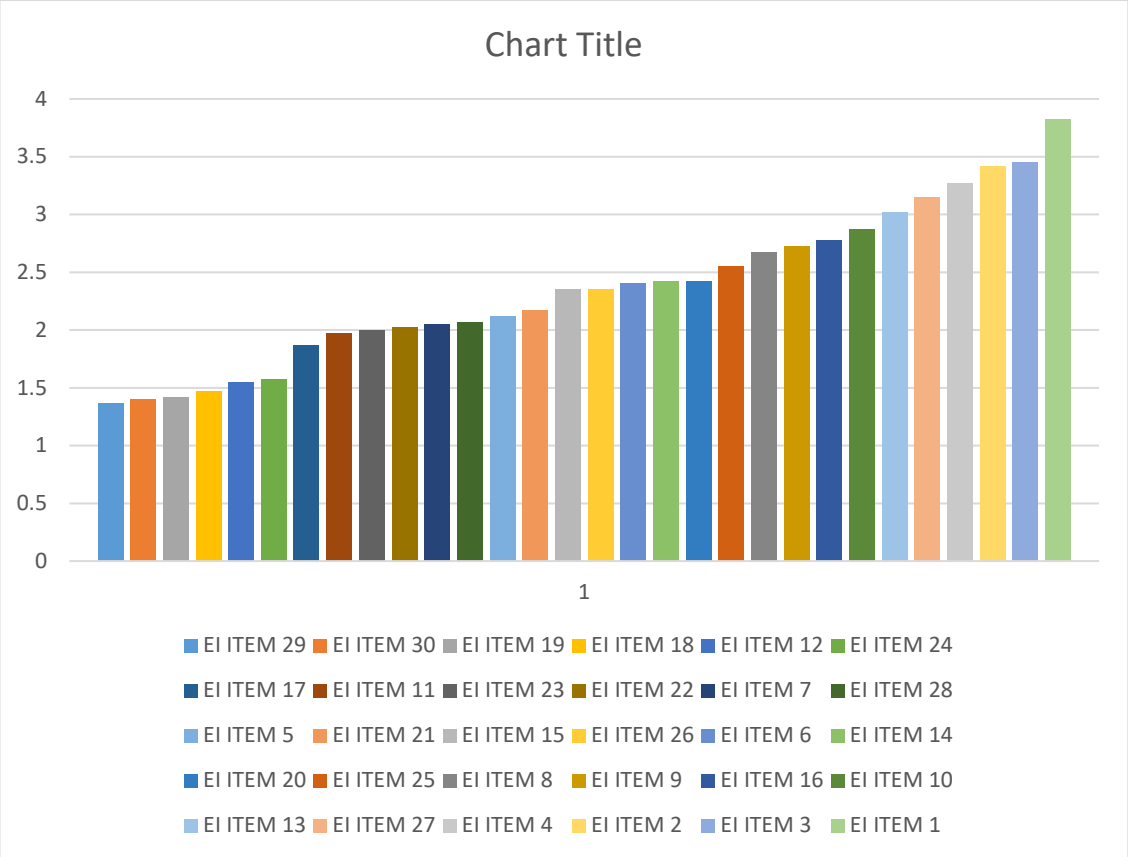
As already mentioned, the 30 sentences making up the EI test are ranked from the shortest (7 syllables) to the longest (19 syllables). Because it is thought that the number of syllables could be a source of difficulty, when it comes to repeating stimuli, participants are therefore expected to score higher on the first few sentences and their score is predicted to decrease as the number of syllables increases. Graph 5.5 below represents the 30 items of the EI test in order of difficulty. The most difficult item, item 29 on the left hand-side with an average repetition score of 1.37 out of four, and the easiest item is the last on the right-hand side, item 1 with an average of 3.82 out of 4. As a reminder, a score of 4 was awarded for perfect repetition. So at both extremes of the graph, the number of syllables seem to be a good indicator of item difficulty. Indeed, item 1 is the shortest stimulus of the test with 7 syllables: *je dois aller au coiffeur* (I have to go to the hairdresser's) and item 29, with 18 syllables: *l'examen n'était pas aussi difficile que ce que vous m'aviez dit* (The exam was not as difficult as you had said) is one of the longest sentences. Only item 30 with 19 syllables is longer: *Il y a énormément d'individus qui ne mangent rien du tout le matin* (There are lots of people who do not eat anything in the morning). Item 1 and item 29 were also reported by Tracy-Ventura et al (2014: 153) as the easiest and the most difficult items for their 29 second year university students of L2 French. The means of these two items are also very similar in these two studies with, in Tracy-Ventura et al, 1.07 out of 4 (compared to 1.37 in the present study) for item 29 and 3.97 out of 4, in Tracy-Ventura et al, (compared to the present study finding of 3.82) for item 1.

However, syllable count is not the only predicator of difficulty. For instance, item 27: *le nombre de fumeurs de cigare augmente chaque année* (the number of cigar smokers goes up each year) and item 29: *L'examen n'était pas aussi difficile que ce que vous m'aviez dit* (The exam was not as difficulty as you had said) have a similar number of syllables, (17/18) and yet their mean score is very different with a mean of 3.15 for item 27 and 1.37 for item 29. This difference between these two particular items was also noted by Tracy-Ventura et al with a mean of 2.76 for item 27 and 1.07 for item 29. As mentioned by Tracy-Ventura et al, the difference between these two items is in their syntactic complexity. Item 27 is a monoclause whereas item 29 has two clauses and while item 27 contains 9 morphemes, item 29 on the other hand contains 14 morphemes.

In addition to syntactic complexity, and "interaction between phonology and syntax, prosody and register" (Tracy-Ventura, 2014: 153), lexical frequency could also be a difficulty predictor. Item 12: *le restaurant est censé servir de très bons plats* (the restaurant is supposed to serve very good food) has the same number of syllables as item 13: *je veux une maison où mes animaux peuvent*

habiter (I want a home where my pets can live) and they both consist of one clause and yet the mean of item 12 is 1.55 whereas the mean of item 13 is 3.02. But in item 12, there are two low frequency words (*censé: 9832; plat: 4817)* whereas most content words in item 13 fall within the 2K frequency band (*maison, animal, habiter*).

Lexical complexity could also explain the relatively small mean of 2.05 for item 7: *Après le repas, j'ai fait une paisible sieste* (After dinner, I had a long, peaceful nap). Despite being short, this sentence was reproduced with some difficulty as only 1/8 (5 participants out of 40) scored 4 points and this was mostly due to them being unable to repeat correctly the last two low frequency words: 4K band for *paisible* and 8K band for *sieste*.



Graph 5.5 Average repetition scores per EI item

If short sentences can be more difficult to repeat because of their lexical, phonological and syntactic features, the opposite is also observable, that is to say, most participants can repeat some sentences with a fair amount of syllables quite easily. As we can see from Graph 5.5, this is the case for items 27 (17 syllables) and 25 (16 syllables) with means of 3.15 and 2.55 respectively. Interestingly, in item 27 *Le nombre de fumeurs de cigares augmente chaque année* (The number of cigar smokers goes up every year), all the content words are from the 1K frequency band apart

from *cigare* (cigar) which is in the 6K band but is also a cognate. Therefore, despite having 17 syllables, this sentence does not present any lexical difficulty. In item 25 *Le voleur que la police a arrêté était très grand et mince* (The thief that the police arrested was very tall and thin), again all content words are from the first 2K frequency band apart from *voleur* (thief) from the 4K band. But in this last case, we can argue that *la police* (police) primes the word *voleur* (thief).

5.4.7 Summary of Elicited Imitation test results

The Elicited Imitation test seems to be a valid tool to measure general L2 proficiency. Indeed, the results above correlate well with Tracy-Ventura et al (2014) who tested students from the same institution, while their UG2 mean score was 62.90, the UG2 mean score of this study is 64.69. EIT mean scores increase from UG1, to UG2 and to UG3, but we have also noticed some individual differences. UG1 and UG2 in particular can display some great differences in proficiency levels, probably due to individual learning biographies prior to university entry. UG3 EI mean is significantly greater from UG1 and UG2 means showing that the impact of an academic year in a French-speaking country and its associated rich input have had a tremendously positive effect on most UG3 participants, and has also reduced individual variation, i.e. the lower scoring UG1 and UG2 students seem to have an opportunity to "catch up" with their higher scoring peers during the year abroad.

Finally, EI scores correlate significantly with end-of-year exam grades as well as with X-Lex scores but not with scores on the active vocabulary test Flex15. It also seems that EI scores themselves can be affected by the frequency of lexical items included in individual sentences, as well as by other morphosyntactic features. In fact, syllable count might be testing participants' memory while lexical, phonological and syntactic factors could be considered to be tapping the more implicit language competencies modelled by Levelt.

5.5 Motivation and language learning background questionnaire results

5.5.1 Objectives

In this section, the results of the motivation questionnaire based on Taguchi, Magid and Papi (2009) and the results of the language learning background questionnaire are analysed. The aim in administering the motivation questionnaire in particular was to investigate the following question:

- To what extent can the individual difference factor of motivation explain learners' variation in L2 vocabulary and L2 proficiency?

5.5.2 Participants and methods

35 participants completed the motivation and language background questionnaires (UG1, *n* = 10; UG2, *n* = 11; UG3, *n* = 14), early in the second semester of the academic year. The motivation questionnaire is made up of 37 statement-type items investigating twelve motivation factors which are detailed in Table 5.28 alongside the distribution of item numbers for each motivation factor.

Factor name	**Item number**
Criterion measures	1, 2, 3
Ideal Self	4, 5, 6
Ought-to L2 Self	7, 8, 9, 10, 11
Instrumentality-promotion	12, 13, 14
Instrumentality-prevention	15, 16
Attitudes towards learning French	17, 18, 19
Interest in the French language	20, 21, 22, 23
French anxiety	24, 25
Integrativeness	26, 27
Linguistic self-confidence	28, 29, 30
Culture interest	31, 32, 33, 34
Attitudes towards L2 community	35, 36, 37

Table 5.28 Item numbers per motivation factor

The 35 participants who completed the motivation and language background questionnaires are 7 males and 28 females; this represents well the (im)balance between men and women in language classes in general at this university and in British universities in general (Lanvers, 2017). The average age of the cohort is 20.5 years and participants have studied French for 10 years on average. The majority of the participants are L1 English speakers, apart from four who speak additional languages at home. UG105 speaks German and considers himself/herself advanced in

the language; Polish is spoken at UG211's home but he/she notes that he/she is of beginner level; UG312 speaks both Punjabi and German fluently and UG314 has Spanish as an additional language and considers himself/herself to be fluent. 26 out of the 35 participants study French along with another language for their degree, while 9 study French language only.

5.5.3 Motivation results

The group mean score (on a scale of 1-6) for each motivational component was calculated, merging scores for relevant individual items, and components were then ranked from the highest to the lowest. Table 5.29 below shows the motivational components ranked from the most highly scored to the least highly scored. The top three motivation factors for the 35 participants are: attitudes towards L2 community, interest in the French language and attitudes towards learning. The whole cohort is motivated by the prospect of travelling to French-speaking countries, meeting French speakers and learning more about them. Participants are also interested in the language, they find it motivating to hear French spoken, and they want to know more about the way French is used in conversation as well as to learn about the differences between French and English. These findings confirm what we might have expected, that is to say that these participants are highly interested in the French language and culture. After all, they have been up to then successful learners of the language who have chosen to study French for their undergraduate studies at university.

Factor	**Average**
Attitudes towards L2 community	5.39
Interest in the French language	5.08
Attitudes towards learning French	5.01
Instrumentality-prevention	4.86
Instrumentality-promotion	4.79
Ideal Self	4.73
Culture interest	4.68
Integrativeness	4.42
Linguistic self-confidence	4.39

Criterion measures	4.15
French anxiety	3.98
Ought-to L2 Self	2.58

Table 5.29 Motivation factor ranking

These three top motivational components are followed by instrumentality-prevention, meaning that students are motivated to learn the language because they do not want to be seen as weaker students and because they want to gain good grades. Again, it is interesting to see how highly important this factor is and yet it could again have been expected as all the participants have had to gain high A-level grades to study at this university. For at least two years before starting their studies at university, they have had to work towards obtaining and maintaining good grades, a factor that has stayed primordial for many of them.

Instrumentality-promotion is the next factor of motivation, that is to say that participants feel that studying French will be useful to get a job, to study abroad for a long period of time and that studying French also offers new challenges in their life. The promotion component is closely followed by the ideal self factor which depicts the way participants see themselves in the future. They can imagine themselves speaking French as native speakers, using it in their future career and while living abroad. This prospect is quite motivating for them and is ranked as the 6th most important motivation factor to learn French. With an average of 4.68 for culture interest and 4.42 for integrativeness, participants are also motivated about learning more about French culture.

Finally, the linguistic self-confidence and criterion measures components are still highly rated but come closer to the bottom of the list of motivational statements. That is to say that participants do, on the whole, consider themselves good at French. They are doing the best they can in French, spending time learning it and doing extra tasks to improve.

The bottom two factors are, French anxiety and Ought-to Self. French anxiety is low relative to other factors, meaning that participants are not nervous on the whole about using French in the classroom or about making errors when speaking French. The fact that the Ought-to Self component is the last factor demonstrates that participants learn the language primarily for themselves and not for their parents, their tutors or their peers. They learn French to fulfil the vision they have of themselves as French speakers who could then live, study or work abroad. The low importance of the Ought-to Self in the present study supports similar findings in recent studies on motivation to learn languages other than English, also referred to as LOTEs. In other

words, L1 English speakers often rank the Ought-to Self as one of the least important motivational factor for learning foreign languages. This has recently lead Dörnyei and Al-Hoorie (2017) to wonder whether this motivational component, as it currently stands, was even useful anymore when investigating LOTEs learning. This point will be discussed further in Chapter 6.

In Table 5.30 below, mean scores for the different component factors have been broken down per year group. We notice that the top and the bottom motivation factors are the same for all three groups and that they are actually the same as the whole cohort's top and last components too. Indeed, attitudes towards L2 community is the most important motivational component for participants regardless of their year of study. The last component for all year groups is Ought-to Self and it shows that participants do not feel pressured by either parents or tutors to study the language and that their learning is driven by personal ambitions.

Factor	**UG1**	**UG2**	**UG3**
Attitudes towards L2 community	5.63 (1)	5.18 (1)	5.4 (1)
Interest in the French language	5.2	4.9	5.14
Attitudes towards learning French	5.03	4.84	5.13
Instrumentality-prevention	4	4.4	4.85
Instrumentality-promotion	5.23	4.63	4.61
Ideal Self	4.93	4.36	4.87
Culture interest	4.35	4.75	4.86
Integrativeness	4.15	4.63	4.46
Linguistic self-confidence	4.13	4.11	4.8
Criterion measures	4.16	3.75	4.47
French anxiety	3.65	4.4	3.89
Ought-to L2 Self	2.56 (12)	2.79 (12)	2.43 (12)

Table 5.30 Ranking of motivation factors per year group

5.5.4 Relationship between motivation and vocabulary size

But what is the relationship between motivation and vocabulary development? To explore how far motivation can explain variation in vocabulary, even within such a generally highly motivated group, we have selected the participants who scored the lowest and the participants who scored the highest on the second receptive vocabulary test, X-Lex May, which tested their knowledge of the seven thousand most common words in French. Table 5.31 below shows participants' numbers and their individual X-Lex May scores.

	X-Lex test May scores (out of 7K)
	Lowest
UG108	2100
UG101	2900
UG208	2900
UG210	3050
UG205	3100
UG211	3100
	Highest
UG303	5000
UG306	5200
UG301	5200
UG105	5700
UG303	5800
UG202	6500

Table 5.31 Lowest and highest X-Lex May scores

Let's start with the analysis of the motivation factors of the lower group. As we can see in Table 5.32 below, the two first most important factors are attitudes towards L2 community and interest in the French language, the same top factors as for the whole group ranking.

Factor	**Average (Lowest X-Lex score group)**
1. Attitudes towards L2 community	5.38
2. Interest in the French language	5.12
3. Instrumentality-promotion	5.05
4. Attitudes towards learning French	5
5. Culture interest	4.74
6. Instrumentality-prevention	4.58
7. Ideal Self	4.49
8. French anxiety	4.49
9. Integrativeness	4.16
10. Linguistic self-confidence	4.05
11. Criterion measures	3.94
12. Ought-to L2 Self	2.03

Table 5.32 Ranking of motivation factors for lowest X-Lex May scores

At the third place instrumental-promotion is ranked slightly higher than for the whole group, that is to say that participants with the lowest X-Lex scores value the language as important to live abroad for a period of time, that French does offer them a new challenge, and that it will be useful to get a job in the future. Instrumental-prevention on the other hand comes at the 6th place, compared to the 4th place for the whole group. So this group declares themselves to be

working hard at French but slightly less so than the average of the whole group. Another difference between this group and the whole group is that participants with smaller vocabulary size declare French anxiety as a more important factor, ranking at the 8th place compared to the 11th place for the whole group. But the ranking of other motivational factors is otherwise quite similar to the ranking observed for the whole group.

Let's now turn to the group of participants with the largest receptive vocabulary, who scored between 5000 and 6500 out of 7000 on X-Lex. The ranking of their motivation factors is presented in Table 5.33 below.

Factor	**Average (Highest X-Lex score group)**
1. Instrumentality-prevention	5.08
2. Interest in the French language	5.08
3. Culture interest	5.03
4. Attitudes towards learning French	4.88
5. Attitudes towards L2 community	4.77
6. Linguistic self-confidence	4.72
7. Ideal Self	4.60
8. Instrumentality-promotion	4.33
9. Criterion measures	4.16
10. French anxiety	3.91
11. Integrativeness	3.83
12. Ought-to L2 Self	3.13

Table 5.33 Ranking of motivation factors for highest X-Lex May scores

Three motivational factors for this group are ranked at higher places than for the whole group. The first one is instrumentality-prevention which is the most important factor for this group. That is to say that participants who have the largest passive vocabulary size, declare that studying hard for French to gain good grades and studying as to not being considered a weaker student is the most motivating factor. In comparison, instrumentality-prevention is ranked at the 4th place for the whole group. The other higher ranked motivation factor is culture interest at the third place compared to the 7th for the whole group and linguistic self-confidence, at 6th place here as opposed to 9th for the whole group.

Three motivation factors on the other hand are ranked lower when focusing on participants that have large passive vocabulary. They are: attitudes towards L2 community found at the 5th place for this group whereas it was the top motivation factor for the whole group. Instrumental-promotion is ranked 8th as opposed to 5th place for the whole group and integrativeness ranked as 11th position for this group as opposed to 8th for the whole group.

To sum up, the motivational factor ranking of participants who have the largest vocabulary size is more different to the whole group profiling than that of participants who have a lower vocabulary size. The motivation of participants with large vocabulary is very much more driven by instrumentality-prevention, and they also tend to be more linguistically confident.

If we now compare the lower and higher groups, they principally differ on four motivational factors, as we can see in Table 5.34: attitudes towards L2 community, instrumentality-promotion, instrumentality-prevention and linguistic self-confidence.

	Ranking for lowest achieving participants	Ranking for highest achieving participants
attitudes towards L2 community	1	5
instrumentality-promotion	3	8
instrumentality-prevention	6	1
linguistic self-confidence	10	6

Table 5.34 Major motivational factors differences between lowest and highest scores in vocabulary test

So the picture that seems to be emerging from these rankings is that participants who have less receptive vocabulary knowledge are motivated by the prospect of traveling to French-speaking countries, of learning about French speaking countries and of living and studying abroad. On the other hand, participants with good receptive vocabulary knowledge are firstly motivated to study hard for French so as to gain good grades and to avoid being considered as weaker students. So it seems that weaker students put the emphasis on the vision of themselves using French whereas stronger participants have less distant and future goals. While also envisaging themselves living in a French-speaking country, stronger participants' top goals are to do well in their assignments and in any activities assessing their performance.

These findings seem to suggest that motivational factors can to some extent explain some of the variations in receptive vocabulary knowledge. Indeed, to develop a large vocabulary size, one benefits from having an instrument-prevention approach, which implies intentional vocabulary learning. Interestingly, one consequence of this can be envisaged to be that these participants therefore spend more time and effort on developing their language skills. If this is the case, one can anticipate that there will be a correlation between vocabulary size and exposure to L2 too, and this will be investigated in Section 5.6. In any case, this finding is also positive from a pedagogical point of view as it suggests that a motivated and focused approach to tasks and assessments set by the institution, leads participants to develop their vocabulary knowledge in quantity and in quality (across the first seven frequency bands in French) that will benefit their L2 development.

A conclusion concerning the motivational factors of students whose receptive vocabulary is less developed is not easy to draw. Indeed, it would seem hasty to conclude that advanced learners that are very well motivated about the prospect to use French in the future will develop their vocabulary less well. What this might suggest though is that these students might not have the strategies to achieve this vision and therefore might need help in developing strategies or skills to achieve the vision of themselves as L2 French speakers. But the investigation of their exposure to L2 French is also crucial as less developed vocabulary might also be the result of less exposure.

5.5.5 Relationship between motivation and proficiency

Extensive learning and processing vocabulary leads to vocabulary development and in turn to better L2 proficiency. Therefore it is interesting to investigate motivational factors linked to proficiency too.

To what extent do very proficient students and less proficient differ in their motivation? Are there clues in their motivation questionnaire that could help understand why some participants have particularly low levels of proficiency? To find out, two distinct proficiency groups were identified, regardless of year group, i.e. the five participants who scored the lowest on the EI test were selected as well as eleven participants who scored the highest. The reason why the more proficient group is larger is because we wanted to include proficient learners from all three year groups, not just the most proficient from the final year of study. The participants' numbers as well as their EIT scores are given in Table 5.35.

	EIT scores (out of 120)
	Lowest EIT scores
UG106	16
UG110	27
UG211*	29
UG210*	32
UG206	40
	Highest EIT scores
UG207	82
UG101	83
UG105*	87
UG310	88
UG205	91

UG314	92
UG313	94
UG315	94
UG202*	105
UG301*	105
UG306*	105

*Are also found in the lowest and highest vocabulary score group

Table 5.35 Highest and lowest proficiency scores

As we can see in Table 5.35, there is a limited overlap with the high/low X-Lex groups and high/low proficiency groups. In Section 5.4.5, a correlation between X-Lex May and EIT scores was reported but, looking at the 40% overlap with low receptive vocabulary/low proficiency and the 36% overlap between high receptive vocabulary/high proficiency, it is clear that both are associated but receptive vocabulary knowledge is not the main predictor of proficiency.

Table 5.36 below shows mean scores for each motivation factor for the whole cohort, for the low proficiency subgroup, and the high proficiency subgroup. The numbers in brackets correspond to the ranking positions of the components. For the highest proficiency group, there were 2 equal ranks, as shown in the Table.

Factor	**Average (all three year groups)**	**Average (Lowest proficiency group)**	**Average (Highest proficiency group)**
Attitudes towards L2 community	5.39 (1)	5.4 (1)	5.27 (2)
Interest in the French language	5.08 (2)	4.9 (3)	5.36 (1)
Attitudes towards learning French	5.01 (3)	4.6 (5)	4.93 (3)
Instrumentality-prevention	4.86 (4)	4.1 (9)	4.36 (8)
Instrumentality-promotion	4.79 (5)	5.13 (2)	4.31 (9)
Ideal Self	4.73 (6)	4.66 (4)	4.81 (5)

Culture interest	4.68 (7)	4.25 (8)	4.93 (3)
Integrativeness	4.42 (8)	4.3 (7)	4.5 (5)
Linguistic self-confidence	4.39 (9)	3.66 (11)	4.81 (5)
Criterion measures	4.15 (10)	4.06 (10)	4.21 (10)
French anxiety	3.98 (11)	4.5 (6)	3.86 (11)
Ought-to L2 Self	2.58 (12)	2.60 (12)	2.7 (12)

Table 5.36 Motivation factors ranking by proficiency levels

Starting with the low proficiency group, most scores are high, and therefore differences in rank order need to be interpreted with caution. The first noticeable difference between this group and the whole cohort is that instrumentality-promotion is higher, being placed at the second place, as opposed to the fifth place for the whole cohort. This means that the prospect of what could happen in the future, e.g. live abroad, find a job, when one learns French, is very motivating for these students. (It has to be noted that promotion was also the second most important factor for the UG1 group.) Another component which is ranked higher within the lower proficiency group is French anxiety. Participants with lower proficiency are more nervous to speak French in class and are more worried to make mistakes when speaking French. Anxiety therefore might impact negatively on the meaning-focused output activities that participants could engage in both inside and outside the classroom. With less interaction, the uptake of vocabulary could be negatively impacted on too. In addition, the motivation factor of attitudes towards learning French is a little lower compared to the whole cohort, meaning that participants seem to like the atmosphere of their French classes a bit less, and they find learning French a little less interesting that the rest of the cohort.

Turning to the more proficient participants who scored between 82 and 105 on the EIT, in first position as the most motivational factor is interest in the French language. Even though this factor is not far behind for the whole cohort, at the second place, it is still interesting to note that it is the most important factor for this group only. This also means that participants might look more proactively for interaction as they agree that they are interested in the way French is used in

conversation and they are feel excited when hearing French spoken. If that is the case, this, in turn, could increase the uptake of vocabulary both incidentally and explicitly.

The second noteworthy difference between the high proficiency group and the whole cohort is that the factor of culture interest is higher for the former. This component is ranked at the third place compared to the 8th place for the whole group. This means that participants may be motivated to access authentic resources via internet for instance, that they enjoy learning about French culture in their content modules and that they would like more modules to be taught in French. They therefore demonstrate a strong desire for even more opportunities to develop not only their cultural knowledge but also their language knowledge through more in-class input. The final noticeable difference lays in the "linguistic self-confidence" factor. This component is in fourth place for the more proficient group, compared to 9th place for the whole cohort and 11th place for the low proficient group. In other words, high proficiency learners believe they are good at French and at languages in general and many of them feel French is their favourite language. It is possible that this self-confidence is a personality trait too but more realistically, their linguistic self-confidence might result from good results in assignments for instance or enjoyable and successful use of the language in native-like environment. This will be discussed further in Chapter 6. In any case, this again should have a positive effect on their learning and on the development of their fluency as these participants are not afraid to access a wide variety of resources and to tackle different activities to consolidate or expand their knowledge.

To sum up briefly this results section, we can say that the motivation factors that are the most and the least influential are reasonably stable across the year groups as we have the same top and bottom ones for UG1, UG2, UG3 and the whole cohort. However, low and high proficiency groups display somewhat different motivation traits that could explain to some extent their varied proficiency scores. As we have seen, lower proficiency participants are more likely to be affected by anxiety, which may negatively affect their opportunity to practice and learn, whereas high proficient learners are confident in their learning and do not seem to hold back from practising the language.

5.5.6 The motivational factors of the most advanced learners

As we can see in Table 5.35, four participants have achieved both the highest X-Lex vocabulary scores and the highest EI scores: two finalists (UG301 and UG306), one second year participant

(UG202) and a first year (UG105). The ranking of this small subgroup's motivational factors is given below in Table 5.37 and compared to the ranking of these factors for the three year groups.

Factors and ranking for most advanced learners	**Average (highest X-Lex and EI test achievers)**	**Ranking for all three year groups**
1. Interest in French language	5.43	2
2. Culture interest	5.18	7
3. Attitudes towards L2 community	5.08	1
4. Linguistic self-confidence	5.00	9
5. Attitudes towards learning French	4.75	3
6. Instrumentality-prevention	4.62	4
7. Ideal Self	4.58	6
8. Instrumentality-promotion	4.33	5
9. Integrativeness	4.12	8
10. French anxiety	4.10	11
11. Criterion measures	4.08	10
12. Ought-to Self	3.55	12

Table 5.37 Motivational factors ranking for the four high achievers in vocabulary and proficiency and comparison with the ranking for all three year groups.

Once more, mean scores are generally high. The main differences of interest between the two groups are that the top achievers in vocabulary and proficiency have higher culture interest and higher self-confidence compared to the whole group. Instrumentality-promotion is also lower for them. However, instrumentality-prevention which was the most important motivational factor as far as vocabulary size was concerned, does not seem to have a distinct position here as it is ranked 6th for the high achievers as opposed to 4th for the whole group.

Finally in Table 5.38, the top three motivational factors for the low and high achievers for the second X-Lex vocabulary test and for the low and high achievers for the Elicited Imitation test are presented for comparison. The common top motivational factor for all these four groups is interest in the French language, while the distinctive priority motivational factor for high achievers' groups is culture interest. Indeed, this could support further the supposition that efficient vocabulary development is linked to the wish to find out more about the culture and therefore to access more resources. In turn, this exposure to L2 French leads to vocabulary uptake that leads to higher proficiency level.

The fact that lower achieving groups share the same top three motivational factors is also interesting, suggesting that a vision of an L2 self speaking the language and using to travel is not enough by itself to trigger learning habits to develop an advanced vocabulary size and a good level of proficiency. If this is the case, then we would expect the same groups to display less exposure to L2 than the high groups.

	Vocabulary	**Proficiency**
Low group	1. Attitudes towards L2 community 2. Interest in the French language 3. Instrumentality-promotion	1. Attitudes towards L2 community 2. Instrumentality-promotion 3. Interest in the French language
High group	1. Instrumentality-prevention 2. Interest in the French language 3. Culture interest	1. Interest in the French language 2. Attitudes towards L2 community 3. Culture interest

Table 5.38 Top three motivational factors for the low and high achievers for X-Lex May

5.5.7 Personal reasons and motivation for learning L2 French at university

In addition to completing the motivation section of the questionnaire, participants were also asked about the reason(s) why they have chosen to study L2 French post A-level at university. Responses given support the findings from the motivation questionnaire, as the most common reasons listed were linked to the wish to become more fluent in the language, to meet French speakers and learn more about the culture. For instance UG314 notes "I find the language attractive and I enjoy French literature and film" and UG304 that he/she "was interested in French culture and literature and wanted to learn more" and to add "as I visited France so much as a child I was also very intrigued by the language itself". The desire to be proficient as a major motivational factor does relates not only with our motivation questionnaire data but also with prior findings from Oakes (2013) and Busse and Williams (2010) who also identify proficiency as the highest motivator.

Instrumentality-promotion types are the second most common reasons given by participants for wanting to pursue learning French at university. 43% of the participants explicitly say that they study French because they want to work abroad or because speaking French will be beneficial for their careers. UG201 declares "I want to have a job in a language field or to move abroad", UG106 for instance states "I think speaking other languages is a useful skill for getting a job". When learning a language is given as a career boosting option, participants do not usually specify though which type(s) of career, if any, they have in mind. There is only one example of very specific career path already considered that pushed UG103 to take French as part of his/her degree, as he/she declares "aspirations of working in the Armed Forces or with the EU as an interpreter or a translator in the future". The lack of clear career goals might be because participants do not feel it is important to mention these, or it might be that as Lanvers states "language degrees are considered to have less obvious follow-on career paths compared to some other degrees" (Lanvers, 2017: 520).

The third most common reason given for studying French at university is that participants enjoyed learning it at school. Prior findings very often point to low levels of motivation in secondary school (from 11 to 16 years old) modern languages classrooms due to a range of issues (Mitchell, 2003; Erler and Macaro, 2012), and so this might be surprising. It could be that their enjoyment was found mostly after the compulsory age of learning languages, that is beyond 14, as some like UG308 state that "I really enjoyed studying French at school especially at A-level and I did not want to stop". But many seem to have enjoyed French all the way from the first year of school like UG310 for instance who declares "I knew I wanted to study languages when I started learning

French in year 7" and many declare that they always enjoyed French at school (UG212, 301, 206, 101). UG101 also seems to have not only enjoyed languages at school but also understands the importance of enjoyment in the modern languages classrooms and has "the aim of completing a PGCE afterwards to spread the love of French to the next generation of linguists as my teachers did to me". These findings show that enjoyment is a very powerful motivator and in some cases like in UG101's, an aspiration trigger. According to the Department for Education (DfE) statistics as reported in Lanvers (2017) students who enjoy language classes at school are more likely to be "privately educated or from selective state schools" (2017: 524). As this kind of information was not asked of participants, we cannot comment on it.

The next most common reason to study L2 French at university given by students is the fact that they were good at it at school. Almost half of the students who declare this also mentioned that they enjoyed languages at school. Whether they enjoyed languages because they were good at them or whether they were good at them because they enjoyed languages cannot be determined, but the fact remains that the link between the two is strong. The only one for whom the link between enjoyment and performance is not particularly present is UG108 who declares that "French was never my best subject at school, but I always enjoyed the classes". This also correlates with the high ranking of the instrumentality-prevention factor at the 4th place when analysing the motivation questionnaires of all the participants across the three year groups. These two findings correspond also to Higgins's (2014) prevention-focused orientation, which is "typically found in students who seek security, and who are keen to comply with external demands and prevent negative outcomes" (Lanvers, 2017: 525), like obtaining low grades. And as we have seen in the analysis of the motivation questionnaires, prevention-focused orientation was particularly important in relation to vocabulary size.

Less often mentioned reasons for studying French at university included the fact that some students, like UG302 and UG313 did not know what else to study, supporting again the idea that languages might not lead to a clear career path but that they are still considered to be useful for the future. Only one student declares that speaking English is not enough and explains that he/she "wanted to speak foreign languages rather than just English" (UG206). Apart from this example, the notion of a "rebellious profile" (Lanvers, 2017: 524) which challenges the fact that there are many British monoglotes is not an explicit motivator for any of the participants. Of course, they are likely to know that they are going against the grain by studying languages. Indeed, they are probably asked why they are studying languages on a fairly regular basis and the small size of their classroom post compulsory period will also have shown them that learning a

language makes them stand out. They do not explicitly say though that this is a motivating factor. On the other hand, one can argue that this is an underlying factor for many motivation factors such instrumentality-promotion. Indeed, when participants declare that studying French will help with their career, what is implied is that a person who speaks L2 French in the UK has a quality that the majority of people do not have. It is not quite a rebellious state but it is a state of going against the flow. Given that instrumentality-promotion is an important factor for all the students, we could say that participant's Self is being built on the rejection of being like the norm, on the rejection of being monolingual. But whether one can talk of a state of rebellion or of an opportunity to be different (and to make the most of it to access different careers for instance) will be discussed further in Chapter 6.

5.5.8 Summary

Overall, we have seen that first, second and final year languages students are all greatly motivated by learning the language itself and by learning more about the places where the language is spoken as well as about native speakers. It is also clear that all participants have chosen to study French at university for personal reasons as the Ought-to Self factor is always ranked as the least important one. As far as vocabulary learning is concerned, it would seem that motivational factors may to some extent explain some of the variation in vocabulary size. The motivational factors of instrumentality-prevention and culture interest in particular are linked with the development of larger vocabulary. This implies that participants who aim at obtaining good grades and accessing a variety of authentic resources to learn more about the L2 culture, and therefore likely to be undertaking regular, intentional vocabulary study, are more likely to know more vocabulary that participants who are most motivated by the general challenge of learning the language and the prospect to use it to travel, to study or to work in countries where L2 French is spoken. And the same goes for proficiency, where a more global and distant vision of the benefits of speaking the language for work or travel is linked to a lower level of proficiency compared to participants who rank culture interest as one of the main motivations for learning the language. The analysis of the participants' personal reasons for learning L2 French connect with some of the findings from the motivation questionnaire but it also shows that some aspects of motivation for learning LOTE (languages other than English) are distinct from the assumptions underlying traditional L2 motivation questionnaires. In the next section, the exposure to L2 French and its link to vocabulary and proficiency development is analysed.

5.6 L2 Exposure and language learning background questionnaire results

5.6.1 Objectives

In this section, the responses to the open-ended questions concerning participants' L2 French exposure out of the classroom are analysed. The main objective is to investigate what is the relationship between the extent of exposure to French and vocabulary learning?

5.6.2 Categorising L2 exposure

Information about exposure to French language during independent study time, is investigated through two open-ended questions in the language learning background questionnaire that was completed by participants in March. As a reminder, the first one aims at finding out about activities participants take part in to improve their French overall, by asking: "Outside French language classes, and related coursework, what activities do you do to practise, consolidate or further your language knowledge and skills?" A list of potential activities were given as examples, and participants were also asked to provide information about the frequency with which they undertake these activities. The second open-ended question regarding independent learning activities focuses particularly on vocabulary development activities, asking: "What strategies do you make use of to build your French vocabulary?"

Exposure to French was therefore primarily investigated through the number and the type of activities listed by participants for language development and vocabulary development. Each learning activity that was mentioned was then categorised following Nation's (Nation, 2007) four strands: meaning-focused input, meaning-focused output, language-focused learning and fluency development, as shown in the example in Table 5.39 below. Activities listed in black relate to general French language development and the ones in orange, to vocabulary development in particular.

Whenever possible, the frequency with which students undertake these activities was also indicated (see Appendix N). In the example below for instance, UG1 reports chatting with native speakers for one hour every week. But more often than not, the frequency with which the activities listed by participants are completed is not provided. For example, UG1 does list that he/she makes vocabulary lists but the frequency of the activity per week or per month is not

given. In other cases, frequency references are vague, for example, UG2 states he/she reads books “sometimes”, UG108 reads the news “once or twice a week”.

	Meaning focused input (learning through reading and listening)	**Meaning-focused output (learning through speaking and writing)**	**Language-focused learning** **(deliberate attention to language features)**	**Fluency development** **(use of familiar materials as quickly as possible in listening, speaking, reading and writing)**
UG101	Subscribing to email newsletters Le Monde « à la une » and about.com (2h-3h/week).	Chatting online with native speakers (1h/week). Integration of vocabulary into sentences.	Studying grammar (1h-2h/week). Proof-reading GCSE/A-level essays for friends. Quizlet.com, online flashcards Writing my own vocabulary lists on paper.	

Table 5.39 UG101 Independent French learning activities in the four strands

Webb and Nation propose that “for a vocabulary learning programme to be well balanced, the time spent both inside and outside the classroom should be divided equally” (Webb and Nation, 2017: 179) between the four strands. Given the limited information about activity types and the proportion of time committed to these available in this study, we cannot formally test this suggestion. However the strands are a helpful way of systematizing the qualitative questionnaire data. Therefore, before analysing the results, we will define the four strands briefly.

The first strand, *Meaning-focused input,* involves incidental learning through listening and reading activities (Webb and Nation, 2017: 179). Examples of meaning-focused input listed by participants include “watching films” (UG102), “reading magazines” (UG104), and “listening to French/francophone music” (UG202). *Meaning-focused output* involves learning though production, i.e. speaking and writing. Activities that fall into this category are for instance

"conversing with my French Erasmus buddy" (UG207) and "emailing my French friends" (UG212). *Language-focused learning* involves deliberately studying different language features and includes activities such as "making my own lists of words to learn" (UG306) and "creating flashcards with new vocabulary" (UG309).

The fourth strand, *Fluency development* involves using the language knowledge one already has but under time pressure. The more general term *fluency development* is usually understood by teachers and learners as opportunities to speak and write in L2. Activities that are related to communicating, such as "conversing with my French Erasmus buddy", could thus have been included in fluency development. However, the definition of *fluency development* within the four strands does include an element of "pressure to process the material as quickly as possible" (Webb and Nation: 2017: 181). Given this definition, no activities listed by the 35 participants could be coded with confidence as belonging to this fourth strand, as self-report could not capture this information.

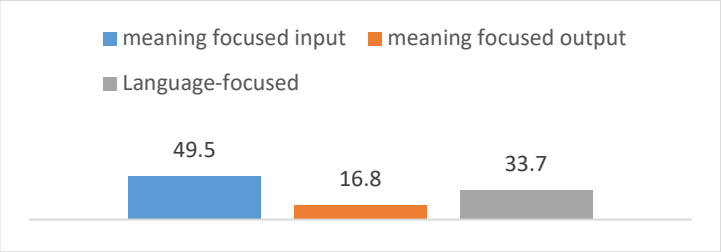
It is also evident that some activities can actually be linked to more than one strand. For instance, the activity "speaking to my Erasmus friend" (UG111) is classified for the purpose of this study as a meaning-focused output activity because the participant stresses the fact that he/she practises speaking. However, is also a meaning-focused input as the participant who converses with a friend also needs to listen to the friend's comments. As Nation (2007) proposes, the divisions between strands should be seen as flexible and so does their precise distribution.

5.6.3 L2 exposure results

5.6.3.1 Description of out of class learning activities

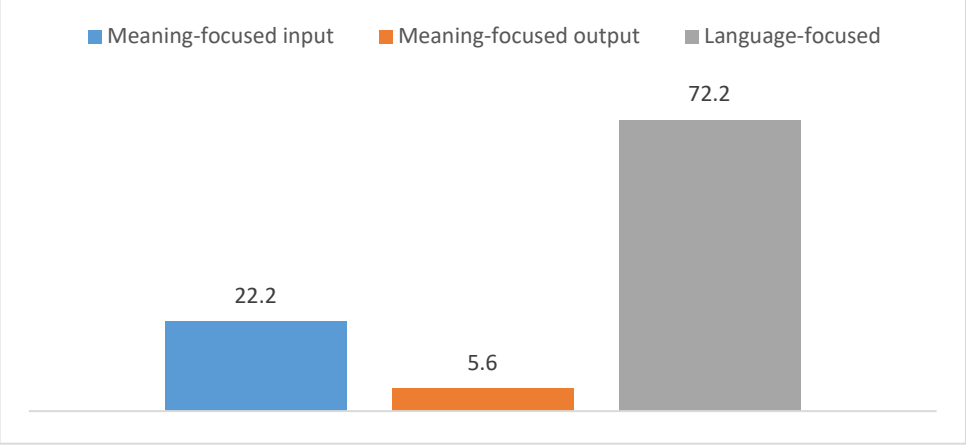
Results show that the 35 participants listed 190 independent activities that expose them fairly regularly to French language outside the classroom, which is an average of 5 activities listed per participant (Graph 5.6). 49.5% of these mentions refer to activities categorised as meaning-focused input such as "reading newspapers and magazine from the LLC" (UG103), "reading parallel books and news articles" (UG201) and "watching films" (UG205). With 33.7%, language-focused activities are the second most common types reported and include "keeping vocabulary and idioms lists" (UG308), "studying grammar" (UG312) or "testing myself on verbs" (UG208). Finally, 16.8% of language development activities can be classified as meaning-focused output; these include "speaking with Erasmus students" (UG104), "speaking French at café parler"

(UG207) and “writing an explanation in French of words I don’t know” (UG303). As already mentioned the frequency of occurrence for these activities is not always provided and when reported, frequencies may vary from “once or twice a week” (UG209), “frequently” and “very infrequently” (UG203) to “once every few months” (UG202).



Graph 5.6 Percentage of independent learning activities per strand

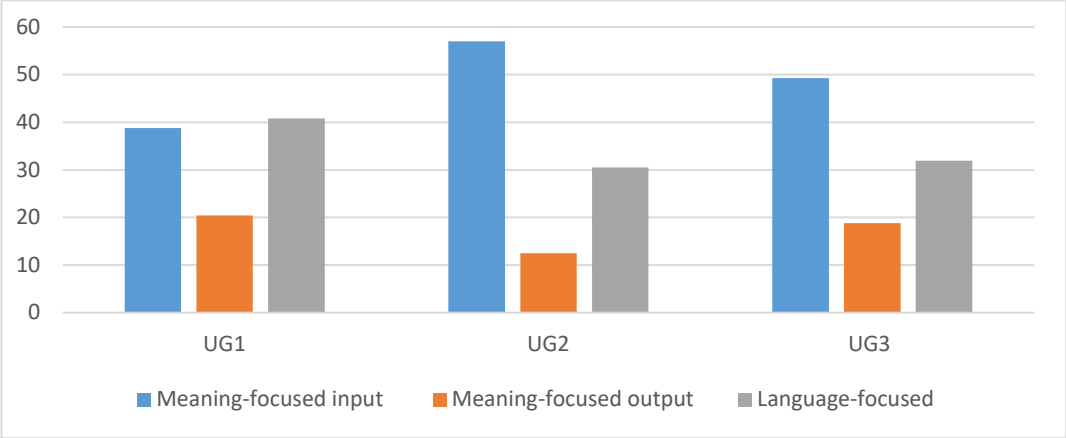
Out of the 190 language learning activities listed, 72 (37.9%) were identified as directly concerning the development of French vocabulary outside of the classroom. Of this total, the majority are language focused (72.2%), followed by meaning-focused input (22.2%) and meaning-focused output (5.6%) as shown below in Graph 5.7.



Graph 5.7 Percentage of independent vocabulary learning activities per strand

In the distribution of overall reported activities within the three strands, UG2 and UG3 display a similar pattern; the highest percentage of activities named by these two groups are meaning-focused input with 57% and 49.3% respectively; followed by language-focused with 30.5% and

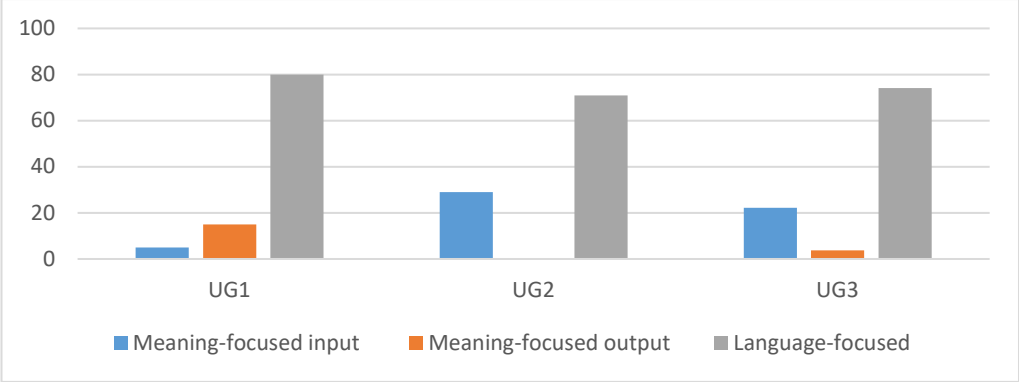
31.9%, and meaning-focused output with 12.5% and 18.8%. UG1 as a group however, favours language-focused activities (40.8%), meaning-focused input (38.8%) and meaning-focused output (20.4%) as shown in Graph 5.8 below.



Graph 5.8 Percentage of French learning activities reported per year group and per strand

We can notice that all participants have fewer meaning-focused output activities than other types of activities, which is so say speaking and writing activities are lagging behind and are constituting less than one fourth of the total out of class activities.

In terms of the distribution of activities specifically aimed at L2 French vocabulary development, all three year groups display a strong tendency to favour language-focused ones, representing 80.0% of all vocabulary learning activities for UG1, 71.0% for UG2 and 74.1% for UG3. For UG2 and UG3, the second most favoured strand is meaning-focused input with 29% and 22.2% respectively. Meaning-focused output activities make up only 3.7% of UG3's reported vocabulary activities and 0% of UG2's. UG1 on the other hand reports somewhat more meaning-focused output with 15% of activities and 5% of input activities, as shown in Graph 5.9 below.



Graph 5.9 Distributions of reported French vocabulary learning activities per year group and per strand

5.6.3.2 Relationship between reported out of class exposure and vocabulary size

The exposure to L2 French outside the classroom has been analysed for participants who scored the lowest and the participants who scored the highest in the X-Lex May test to investigate whether there is a link between the participants' vocabulary size and their reported out of classroom exposure.

Let's start with the lowest group of six participants who have already been identified in section 5.5 (UG108, UG101, UG208, UG210, UG205, UG211) and whose vocabulary scores are between 2100 and 3100 out of 7000. These six participants report a similar range of activities to other participants, but on the other hand, there are some issues with the reported frequency of these activities. In some cases, no frequency is given and when it is provided, some low-attaining participants clearly spend little time on developing French outside the classroom and much less than they are advised to spend by their institution. For instance, UG108, who scored the lowest on the passive vocabulary test, reports that he/she watches DVDs once a term, reads once or twice a week, chats online to native speakers once every three or four weeks and writes grammar notes once every two weeks. This information suggests a probable link between the participant's low exposure to French and his/her low vocabulary knowledge. Similarly UG208 "sometimes" reads online articles, watches Youtube videos and occasionally watches French films with English subtitles.

The other low-attaining participants also display a low frequency level of exposure to French outside the classroom, when reporting general learning activities. UG210 watches French films once a month. He/she undertakes some other activities more often, like "listening to French radio

two to three times a week" or "listening to Michel Thomas' *Teach yourself French* three to four times a week". But there is no clear evidence of active learning involved in these activities. UG211 does not specify the frequency of the activities he/she lists but again his/her engagement in those does not appear great as he/she "skims" through the news. UG205 lists "watching films, listening to music and looking up unknown words in the dictionary" without specifying how often these activities are taking place.

As for activities that target vocabulary development in particular, participants with low X-Lex scores tend to say they make lists of words they encounter they do not know (UG101, UG108, UG208, UG205, U210). Making online flashcards is also popular (UG101, UG211) and UG101 makes sentences using new learned words. UG208 listens to songs to learn new words. If five out of six participants look up words they do not know on the other hand, the quantity of new words they actually look up and learn is not specified. And given that, as we have seen, these participants tend to have fairly little exposure to French, the amount of new vocabulary or expressions or even the amount of occasions to encounter known vocabulary but in different contexts is likely to be low.

Let's now turn to the six participants who scored the highest in the X-Lex May test, with scores ranking from 5000 to 6500 out of 7000 (UG302, UG306, UG301, UG105, UG303, UG202). UG202, who achieved the highest X-Lex May score with 6500, lists eight different language learning activities that he/she undertakes outside the classroom, including "listening to French/Francophone music, reading blogs, newspaper articles online, reading books etc". Reading activities tend to happen on a regular basis from weekly to a few times per week, while other activities such as watching a French film occur less often, once or twice a month. UG202 as well as UG303 and UG301 also list reading as an activity they explicitly undertake to develop their vocabulary. Others like UG306 who do not explicitly say they read to develop vocabulary do however undertake this activity most days.

So regular reading and reading specifically for learning new vocabulary are reported regularly by those participants who have the largest vocabulary size. This reinforces the idea that a lot of vocabulary uptake at this advanced level does take place outside the classroom including when students are explicitly engaged and seek to make vocabulary uptake their goal. Of course, this may also be an example of how "accumulation of lexical knowledge may reduce the learning burden of unknown words" (Webb and Nation, 2017: 57). In that case, it is possible to imagine that these participants learn more from reading because they already know many more words

than participants who have less vocabulary knowledge and who, therefore read much less efficiently.

While this qualitative analysis has revealed some apparent links between exposure and vocabulary uptake, care must be taken not to generalise too hastily. For instance, UG105, who is a participant with both one of the largest vocabulary size of the whole cohort and also with one of the highest scores for EIT, mentions very limited out of class activities: these are “watching DVDs once a week and making online flashcards using Quizlet”. UG302 who has a X-Lex May score of 5000 mentions “watching the odd French film occasionally, speaking with native speakers on tandem exchanges and learning vocabulary with the help of the app AnkiDroid”. UG105 has developed a vocabulary size of 5700 and UG302 of 5000, so it is very unlikely that they have reached those figures by doing only the types of activities listed. So participants’ reporting of exposure does vary and needs to be treated with caution when being analysed.

Looking at the types of activities this group undertake to develop their vocabulary, we can notice that some like UG105 and UG302 use the same online tools as participants in the other group, such as the online flashcard generating website Quizlet and the app Ankidoroid. But one reported difference is that participants who have a larger vocabulary not only look up words they do not know; they also report actively engaging in re-using them. UG301 for instance, imagines sentences with the new vocabulary he/she has learned and searches for different uses of new words on webcorp.org. UG303 when looking up new words also writes explanations for them in French and UG306 makes an effort to reuse new words in his/her speech and assignments. Even UG202, who does not specify re-using new words in any way, shows that he/she does pay particular attention to the type of vocabulary he/she wants to learn, evaluating whether new words he/she encounters are relevant for him/her before looking them up. Participants with smaller vocabulary size did not list any such evaluation or extra activities when looking up new words, with the exception of UG101 who makes sentences recycling new words.

To sum up this section briefly, it seems that participants who have a lower vocabulary knowledge tend to undertake fewer activities in L2 out of the classroom, and when they do, they do not seem to be as focused or actively learning from them. Participants with higher vocabulary knowledge seem to be spending more time learning French out of the classroom and have a stronger intentional approach to their learning.

5.6.3.3 Relationship between reported out of class exposure and overall proficiency

To explore relationships between out of class activities and overall proficiency, the language learning activities of the low proficiency group identified in section 5.5 and made up of five participants (UG106, UG110, UG211*, UG210*, UG206) have also been analysed. (Participants marked with an asterisk were also in the low passive vocabulary group discussed in the last section.)

In most cases, the quantity of meaning-focused input activities is quite low for these participants, indeed, they do not seem to read or to listen to much French outside the classroom. UG106 for instance, does not mention any reading in French activities taking place during his/her free time. The only meaning-focused output mentioned is “watching French vlogs” but even this activity seems to take place infrequently as the participant adds “when I have time”. The participant notes that he/she does not like reading in English either.

UG110 has a slightly different language learning background as he/she did not study French at A-level but took French up again at university. In his/her final year, he/she has reached post-A-level stage and is studying the language with mainly first year modern languages students. In this case again, opportunities to encounter French in contexts outside the classroom seem to be treated casually. UG110 states that he/she reads books and watches films in French “to relax”. There is no stated frequency of occurrence for these activities but being defined as relaxing activities, it is possible that the attention of the student to form-meaning relations may not be great, and that vocabulary learning and fluency development are not happening because of this potentially low commitment level.

For UG210, the reported listening and reading activities also seem casual as he/she mentions practising listening skills through YouTube videos and “skimming” through the news. UG211 also shows signs of low focus when listening and reading outside the classroom as he/she listens to French radio, quite regularly: two or three times a week but “usually on in the background whilst getting ready in the morning”. UG211 also watches DVDs, but with a frequency of only once a month. From all the examples given so far, it seems that participants know that they have to vary the resources they use outside the classroom to develop their French language. They also are aware that different skills need to be practised but it seems that not enough time and attention are directed towards these activities for good quality language uptake to happen.

Out of the five participants in the low proficiency group, one offers a possible counter-example. UG206 lists reading French newspapers online as well as reading books and watching YouTube videos of French vloggers and yet is still in the bottom five. As the frequency of these activities is not given, however, it is possible, as with the above examples, that they do not occur often enough for him/her to strengthen his/her knowledge of the vocabulary he/she already has neither to expand his/her vocabulary nor his/her fluency. These five participants have all listed at least one meaning-focused out of class activity, writing or speaking activities, usually "emailing a French friend" or "speaking French to other people" within a language tandem or "café parler". But when the frequency of occurrence is given, these activities seem to happen only once a week.

Let's now turn to exposure to French for the high proficiency group made of eleven participants as identified in section 5.5 (UG207, UG101, UG105*, UG310, UG205, UG314, UG313, UG315, UG202*, UG301*, UG306*; asterisk indicates participants who were in the top-scoring group for X-Lex May). The first interesting point is that the majority of them, that is seven participants out of eleven, give frequencies for their language learning activities, and the activities listed clearly seem to happen regularly. UG306, watches films, TV games or series once a week and reads news most days; UG313 listens to French news every day and often watches films or series. UG101 spends two to three hours a week reading French newsletters online and speaks online with native speakers for an hour every week.

Five out of eleven high proficiency participants also report reading novels in French, some explaining clearly what type of books they prefer; for example, UG310 reads French versions of books he/she has already read in English. This is an effective way to develop fluency and to consolidate knowledge of vocabulary already known as one sees it in context, while at the same time learning new vocabulary too. But novels contain so many different words that for them to be effective sources of language learning, participants must already have a good vocabulary size. And indeed, the five novel readers have an average X-Lex May score of 3720. With this solid vocabulary knowledge already in place, reading novels in French should therefore enable them to cope well with unknown vocabulary and make the most of this input activity.

Interestingly, within the low proficiency group, reading novels was also mentioned by some participants, but there was an occasional feel to this activity. UG210 has read one novel for his/her content module and UG110 reads books but does not provide any information about the type of books or the number of books read. And with a receptive vocabulary knowledge of 1650 and 1900 respectively, it is possible that for UG210 and UG110 novels contain too many unknown words and are therefore not the most appropriate resource to develop language knowledge and

vocabulary in particular. In other cases, such as UG107 (who is neither in the lower proficiency group nor the higher), reading a novel is listed as an activity but he/she mentions the quantity of one book a year. It is likely that the rarity of this occurrence will have little effect on his/her vocabulary and proficiency development.

In brief, it seems that more frequent exposure is a trait of more proficient learners as it is a trait of participants who have a wider vocabulary knowledge too.

5.6.4 Summary

The investigation of the participants' exposure to French outside of the classroom seems to demonstrate that it can, to some extent, explain differences in vocabulary size and in proficiency. It highlights that there are three main differences between low and high proficiency groups: frequency of exposure, participants' engagement with materials, and relevance of materials. For participants who have a low vocabulary size and a low proficiency level, the frequency of exposure to French seems relatively limited; the participants' attention to the language from these input activities is too low and the resources are not always adequate for their level. On the other hand, participants with larger vocabulary size and higher proficiency level tend to be more engaged in their L2 French activities out the classroom (which also happen more often), in particular when focusing on vocabulary development. Other research tells us that "incidental vocabulary learning is probably enhanced by support from deliberate learning techniques and resources that help to develop lexical knowledge" (Webb and Nation, 2017: 57). The high attaining group seem to understand this and to be implementing it, unlike the low attaining group, who paradoxically have greatest need to follow such advice. Reading is also a much more common and frequent activity among participants of the higher proficiency level group.

The information provided by the participants about their exposure to French outside the classroom also seems to suggest that they are all aware that they should practise different skills (listening, reading, writing, speaking). But the low level of engagement in those skills by some first and second year participants (who make up 100% of the low proficiency group), can impact negatively on their incidental learning. Hence, the need to identify low proficiency and low vocabulary level participants, to reinforce the importance of deliberate learning, and to advise them on those particular strategies which will help students at their level to improve their language learning experience. It is a limitation of qualitative research of this kind, however, that the way exposure is reported can vary greatly from one participant to another. For instance we

saw that the L2 activities listed by UG105 who scored in the top group for both vocabulary and proficiency do not reflect at all his/her language ability.

In conclusion, exposure to French can to some extent and with this cohort explain some of the difference in proficiency and vocabulary learning. It also has some potential implications for the teaching of French at this advanced level that we will discuss in the next chapter.

5.7 Conclusion

In this chapter, the main research questions of the study have been addressed. The data from X-Lex, Flex15 and the Elicited Imitation tests as well as the results from the motivation questionnaire and about exposure to L2 French have been analysed and related to each other. Some references to other studies or to the literature have been made where necessary. However, a full discussion of the data follows in Chapter 6.

Chapter 6 Discussion

6.1 Introduction

This chapter provides an analysis of the key research findings presented in chapter 5, with reference to each of the research questions and in relation to previous research studies. As a reminder, the main questions that this research aims at answering are:

- What are the rate and the variation of lexical development in L2 French over a 3-year period of instructed learning?
- Is the frequency factor observable in L2 learners' lexical development?
- To what extent can the individual difference factor of motivation explain learners' variation?
- To what extent does overall L2 proficiency explain the pattern of L2 vocabulary development over time?
- What is the relationship between the extent of exposure to French and vocabulary learning?

The first section (6.2) discusses the rate and the variation of lexical development in L2 French over a 3-year period of instructed learning and whether the frequency factor is observable. The role of overall L2 proficiency and the pattern of L2 vocabulary development over time are analysed in section 6.3. Differences in lexical development in relation to the extent of exposure to French as well as the relationship with the individual difference factor of motivation are discussed in section 6.4.

6.2 Lexical development over a 3-year period and the frequency factor

6.2.1 Receptive vocabulary development

The first research question investigates the rate of L2 French lexical development of UG1, UG2 and UG3 participants in a UK university. The results of the two X-Lex tests conducted at the beginning and at the end of the academic year show that each year groups' receptive vocabulary has increased within one academic year. As we can see in Table 6.1, there is also an instruction

effect in the sense that final year students have a larger receptive vocabulary size than second year students who in turn have a larger receptive vocabulary size than first year students.

	X-Lex Nov			X-Lex May		
	UG1	**UG2**	**UG3**	**UG1**	**UG2**	**UG3**
Mean score	2571	2996	3610	3044	3329	3704
Minimum Score	1750	1650	2650	1900	2400	3150
Maximum Score	3700	4450	4450	4250	4800	4200
Standard Deviation	647	788	577	763	704	357

Table 6.1 Means, minimums, maximums and standard deviations of X-Lex Nov and May tests per year group

However, UG2 is the only group displaying a statistically significant within-group difference between the X-Lex November test and the May test. These results show that, within the three year of instruction, there is no vocabulary spurt, like those documented in the UK education system at GCSE or A-level for instance (Milton, 2009).

By the end of the academic year, the difference between all three year groups' receptive vocabulary size is small, with a difference in vocabulary between UG1 and UG2 of only 285 words and between UG2 and UG3 of 375 words. The similarity of these descriptive results is confirmed statistically as, using SPSS, it was found that the only statistically significant difference was between UG1 and UG3 on the X-Lex November test. David (2008) also reported a small difference

of 330 words, compared to 285 in the present study, in receptive vocabulary between her UG1 and UG2 participants. In a study of L2 French vocabulary growth through UK school and university, Milton (2006) also reported a small difference of 400 words, again very close to the present study's result of 375 words, between the receptive vocabulary scores of second year university students and final year students, who had completed their intercalary year abroad. These small uptakes in receptive vocabulary over an academic year do not necessarily mean that participants have not improved. Instead, as Milton points out it is possible that "very advanced learners, who already have large vocabularies, might develop depth and fluency of use at the expense of breadth, in order to make gains and become more proficient" (2009: 77).

As Milton (2009) points out "the instructional setting plays an important role not only in the amount of words that is learned for each level but also the number of words that need to be learned for each level". Therefore, the small difference in receptive vocabulary between UG1, UG2 and UG3 also raises the question as to whether or not current targets of receptive vocabulary to match certain education levels are appropriate. In Meara and Milton (2003), approximate vocabulary size scores associated with CEFR levels are proposed, and a 3250-3750 vocabulary size corresponds to CEFR B2 level. This implies that, based on their receptive vocabulary size alone, for the first five thousand most frequent words in French, UG2 and UG3 are all at B2 level. For UG2, described by the university as CEFR C1, the expected approximate vocabulary size should be 3750-4500 and for UG3, CEFR C2, the vocabulary size should be 4500-5000. However, we saw in chapter 4, that the receptive vocabulary of the three groups has also developed over the academic year beyond the 5K bands and into the 6K and 7K ones. Not only does this mean then that advanced learners are learning rarer words but it also means that when taking into account the first 7,000 most frequent words in French, their receptive vocabulary size exceeds, in the case of UG1, the vocabulary size prescribed by the CEFR, with a total score of 3811. With receptive vocabulary scores of 4112 for UG2 and 4662 for UG3, both groups are within the CEFR vocabulary size range prescribed for these levels.

These data therefore suggest that advanced learners' receptive vocabulary should be tested beyond the 5K band, firstly because their lexical knowledge expands beyond this limit and not taking this into account is not getting a full representation of their development. This also suggests that a vocabulary size of 4000 words for instance is not the equivalent of knowing the first four thousand most frequent words. It also seems that a receptive vocabulary size of about

2570 words, the UG1 mean at the beginning of the academic year, is enough to be able to learn rarer words. This supports the figure of 2000 words which is the number of words that are considered as essential in language teaching as they contribute a lot to coverage in general texts (Nation, 2001; Milton, 2009). As expected, we also observed a wide range of scores within the three year groups, supporting the well-known fact in language learning that the development of individuals' vocabulary can vary greatly. As Milton suggests (2009: 243) "while we understand something about how groups of learners grow vocabulary in a foreign language, we really do not understand and cannot yet predict how individuals will behave". For instance, the receptive vocabulary score of one UG1 participant is higher than the receptive mean score for UG3. Reasons for individual variations are numerous and we will review motivation and exposure effects on vocabulary and proficiency development below.

6.2.2 Receptive vocabulary development and the frequency factor

When analysing the frequency of responses for all 3 year groups in the X-Lex November test, the influence of the frequency factor was observed for all bands except for the 2K band, that is to say all groups displayed a 2K deficit. Their knowledge of 2K band vocabulary is not as high as expected. This could be either because of the test itself, that is, maybe the 2K words selected for the November test produced this effect, or it might be because the three groups are 2K deficient, a common enough profile observed by Milton 2007 (Milton, 2009: 244). As we have just seen, the first two thousand words are seen as a threshold to gain autonomy in language learning, therefore this requires attention. It might mean that teaching materials do not teach and do not recycle enough the vocabulary found in the 2K frequency band. If this is the case, this should be addressed as many studies have proven that the first two thousand words constitute a threshold for intermediate learners to become more autonomous learners and to be able to deal with authentic materials, supporting their incidental learning which is key, as we will see, to lexical development at these advanced stages.

If in the X-Lex Nov test, there was a 2K deficiency, the frequency factor is however observable for all three groups in the X-Lex May test. That is to say, the number of words known for each frequency bands decreases as the frequency band gets higher. While the three groups are sensitive to the frequency factor, when we turn to individual profiles we notice that 59% of participants, 20 participants out of 34, display regular frequency profiles. This is in line with

Milton's study that found that 60% of learners, 132 of 227, have the same frequency profile displayed by the group as a whole (Milton, 2009: 34). The percentage of regular profile varies according to the year group: 67% of UG1 participants have a regular frequency profile, compared to 75% of UG2s and 38% of UG3s. As reported by Milton, our data suggest that for more advanced students, a ceiling has been reached for high frequency words and that for this reason the frequency effect is not noticeable for the lowest bands for some of these students.

If all the UG1, UG2 and UG3's profiles are 2K deficient in November and then regular in May, the way each frequency band develops over an academic year varies from band to band and these developments are also different for each year group. UG2 for instance displays the most straightforward change with all bands increasing from November to May; UG1 shows an increase in all the bands apart from 3K; and UG3 shows an increase in mean scores for the 2K and 4K bands and a decrease in the 1K and 5K bands. Beyond the 5K band, with a means of 767, 783 and 958, for UG1, UG2 and UG3 respectively, the results for vocabulary known at 6K and 7K frequency levels show that university students also know a number of rarer words and that their knowledge of those rarer words is larger as they progress through their degree. The frequency factor is not visible at 6K and 7K level as participants seem to know more 7K words than 6K ones, but as we can see in Table 6.2 below, the difference between these two bands can be very small. It is clear from investigating receptive vocabulary development of the three groups that there are some individual differences. Later, we will discuss to what extent exposure to French and motivation could potentially explain these differences.

	Mean for 6K frequency band	**Mean for 7K frequency band**	**Total**
UG1	328	439	767
UG2	379	404	783
UG3	475	483	958

Table 6.2 UG1, UG2 and UG3 means for 6K and 7K bands

Going back to the research questions related to receptive vocabulary, we can conclude that the rate of lexical development within an academic year is mostly not statistically significant. The only statistically significant increase in receptive vocabulary from November to May is for UG2 and even then, the absolute increase is only of 291 words. By the end of the academic year, all three

groups have similar receptive vocabulary size. The frequency factor is observable for all the bands except for the 2K one in November, but across all bands without exception in May. Beyond the 5K band, the frequency factor is not observable anymore. We also observed great individual variation within groups with for instance 67% of UG1 participants having a regular frequent profile, compared to 75% of UG2s and 38% of UG3s. The receptive vocabulary sizes and uptakes at these levels are similar to results from prior studies (David, 2008; Milton 2006).

6.2.3 Productive vocabulary development

The second research question regarding the development of lexical knowledge was to investigate the development of productive vocabulary knowledge of the three groups between the beginning and the end of the academic year; and to determine if the frequency factor plays a role in this productive vocabulary development and whether receptive and productive vocabulary knowledge are related.

First of all, the Flex15 data seems to validate the rule of thumb according to which as learners become more proficient, they generally can produce more words (Milton, 2009: 122). It is mostly noticeable for the UG1 group for whom the number of omissions in the November test is 4.86%, falling to 0.74% in May. The UG1 group is the only one to show a statistically significant increase from November to May in its productive vocabulary. It cannot be ruled out of course that there is a task effect for the decrease in omission of answers from UG1. The percentage of omissions in the November test is 2.69% for UG2, falling to 0.74% in May. Only UG3 increases its omission rate from 0.22% in November to 1.66% in May, with a total of 14 omissions out of a possible 840 possible responses. Having said that, 13 out of the 14 UG3 omissions are attributed to one student only. If this participant's responses were dismissed for having such a high rate of omissions, the rate would then fall to 0.11% hence showing again a proficiency effect in participants' ability to produce words. As expected from having such similar omission rates, no statistically significant differences were found between the groups in the number of words produced. This might have been different if the test instructions had stated that students should write as many words as possible, as in Waring's (2007) study for instance, where participants produced up to 18 words per stimulus. The fact that the omission rates are quite low in the present study is however encouraging and shows that advanced learners can conceptualise and

activate a fair amount of words for a range of semantic fields triggered by the different cue words.

6.2.4 Productive vocabulary development and the frequency factor

Let's now turn to the frequency of the words produced. The three groups tend to produce more high frequency responses to stimuli, that is to say words from the first and second frequency bands, than low frequency words, as shown in Tables 6.3 and 6.4.

	L1	**L2**	**L3**	**L4**	**L5+**
UG1	50.68%	18.93%	10.07%	6.25%	14.04%
UG2	48.50%	20.46%	10.56%	6.36%	14.09%
UG3	46.28%	19.88%	10.51%	6.62%	16.68%

Table 6.3 Frequency of responses for Flex15 Nov in percentage

	L1	**L2**	**L3**	**L4**	**L5+**
UG1	49.13%	26.20%	9.82%	5.00%	9.82%
UG2	48.26%	26.23%	8.84%	5.21%	11.44%
UG3	50.56%	22.15%	9.88%	7.25%	10.13%

Table 6.4 Frequency of responses for Flex15 May in percentage

In the November test, more than two-thirds of words produced by UG1 and UG2 are high frequency words and in the May test, two-thirds of all three groups' answers are high frequency words. This supports the idea that "a large proportion of the high-frequency vocabulary is known both receptively and productively" (Nation, 2010: 371). There is also a clear frequency factor though as the percentage of responses tend to decrease as the frequency bands increase, hence showing that as for receptive vocabulary learning, productive vocabulary learning is influenced by frequency. This validates the model that underpins the Flex15 test, "a model of vocabulary

acquisition whereby the order in which learners acquire words aligns with the frequency of those words in language use" (Fitzpatrick and Clenton, 2017: 850).

For the three groups combined, there is a statistically significant increase of high frequency: L1/L2 answers from November to May and a statistically significant decrease of low: L3/L4/L5 + answers from November to May. The increase in high frequency words is driven by UG2 and the decrease in low frequency words on the other hand is statistically significant for UG3 only. It is unlikely that knowledge of low frequency words has decreased from November to May and the different proportions of high and low frequency words in both tests might be a test effect, due to the selection of lexical cues. As Nation notes too, "increases in vocabulary size as measured by direct measures of vocabulary are not necessarily reflected in an increase in vocabulary in use" (Nation, 2010: 371). That is to say, a word that is not elicited by Flex15 could still be known and used in speech or in writing. Another explanation could be that lexical improvement from November to May did occur but the nature of this improvement cannot be captured by the Flex15 test. For instance, Horst and Collins, reporting on data from another productive test, the Lexical Frequency Profile, noted that the proportion of words produced in the 1,000 band did not decrease with time but participants' lexical production on the other hand became more "register appropriate, diverse, and morphologically complex" (2006: 102). Fitzpatrick and Clenton support this view and note that "lexical improvement can also be represented in extended uses of (frequent) words, morphological sophistication, collocational awareness and so on" (2017: 857).

By the end of the academic year, L1/L2 responses represent 75.33% of the overall UG1 responses, 74.49% of the UG2 ones and 72.71% of the UG3 ones. Even though there is an institutional effect in the sense that a larger percentage of rarer words is produced by more advanced participants, the proportion of high frequency answers is still greater for all three groups. So whilst the frequency factor can be observed in the development of both receptive and productive vocabulary, the former can develop faster in the lower bands than the latter. The reason may be that the development of receptive vocabulary knowledge depends on varied input from all discourse types such conversation and written texts including novels and newspaper articles for instance. Both high and low frequency words are encountered on different occasions and in different contexts, hence leading to strengthening of these words in the mental lexicon. As we have seen earlier, receptive knowledge testing is also more sensitive, allowing for partial knowledge.

On the other hand, productive vocabulary development depends on written and spoken output. As knowledge of the most frequent 3,000 word families is said to be sufficient to understand a

conversation (Nation, 2006; Van Zeeland and Schmitt, 2013), logically the productive knowledge of these same 3,000 word families is enough to converse both in an L1 and an L2. So the need for a lower productive vocabulary size is also one of the reasons why rarer words are lagging behind in this vocabulary knowledge dimension. But as the data suggest, rarer words are being learned, probably driven by more varied and more frequent output but also driven by new topics participants need to learn about, discuss and write about. Therefore, a productive test in which cue words would relate to new topics for instance, might be helpful in tapping into rarer words even faster. The Brainstorm Frequency Profile test (BFP), which requires participants to write as many words as possible about a given topic could be useful in that way. However, when comparing the results of 80 L1 Japanese learners of English who took the BFP and the Lex30 tests, Fitzpatrick and Clenton (2017: 853-854) found that the BFP elicited "a smaller average percentage of infrequent words and a greater range of scores". So it seems that as far as productive knowledge is concerned, we have to accept that low frequency words are lagging behind, but it may also be true that productive tests do not full justice to word knowledge development.

The medium sized correlation between passive and active vocabulary that was observed in November (r = .435), is unusual. Indeed many studies have reported strong positive correlations between receptive and productive vocabulary size, such as Laufer (1998) with a correlations of .72 for ESL and .89 for EFL students; and Meara and Fitzpatrick (2000) with a 0.84 correlation between a receptive Yes/No vocabulary test of the 10, 000 most frequent words in English and Lex30. But as Nation (2010: 371) notes, the ratio of receptive vocabulary to productive is not constant. On the one hand, we know that learners' receptive vocabulary size is greater than their productive vocabulary size. But both kinds of vocabulary develop differently and we do not know how systematically new receptive vocabulary knowledge could be transferred to productive vocabulary knowledge. In addition, Nation and Webb (2011) note that productive vocabulary tests are unlikely to relate easily to vocabulary size because these tests do not give credit for partial knowledge (in Fitzpatrick and Clenton, 2017: 847), as also mentioned above.

Some high frequency stimuli such as *jardin* (garden), *espoir* (hope) and *mémoire* (memory) can trigger 100% of L1/L2 first responses from UG1 and UG2 but none of these words trigger 100% high frequency answers from UG3. We could see in this the sign that UG3 productive vocabulary is more varied and reflects the higher quantity of rarer words UG3 participants have learned compared to UG1 and UG2. An increase in the use of low frequency words by learners is often associated with progress and language sophistication (Laufer and Nation, 1995); that is to say "the

better a learner is, the more likely they are to use more infrequent vocabulary in production" (Milton 2009: 131).

Low frequency stimuli (beyond the 2K band threshold) elicit more low frequency answers, as shown in Table 5. Out of the 8 low frequency stimuli, some elicit more low frequency words with an average of 55.79% L1/L2 answers for UG1 (compared to 70.23% overall), 57.56% for UG2 (as opposed to 70.20%) and 50.51% for UG3 (as opposed to 65.68%). The low frequency word that elicits the least high frequency responses is *ordinateur* (computer) (4428) with an average of L1/L2 responses of 16.66% (UG1: 16.66%; UG2: 33.32%; UG3: 0%), followed by *assiette* (plate) (2818) with an average of 38.88% of L1/L2 answers (UG1: 30%; UG2: 40%; UG3: 46.66%). On the other hand, the cognate, *stupide* (3651) elicits 47.11% of L1/L2 answers. Another cognate *crayon* (4939) elicits 54.05% of L1/L2 answers (UG1= 75%; UG2= 53.84%; UG3 = 33.33%), lundi (2430), 56.42% (UG1= 55%; UG2 = 50%; UG3 = 64.28%), *tristesse* (sadness) (3371) 63.09% (UG1= 66.66.%; UG2= 58.33%; UG3= 64.28%).

Low frequency stimulus	**L1/L2 percentage responses overall**	**UG1 L1/L2 responses**	**UG2 L1/L2 responses**	**UG3 L1/L2 responses**
Lundi (2430)	56.42%	55.00%	50.00%	64.28%
Assiette (2818)	38.88%	30.00%	40.00%	46.66%
Tristesse (3371)	63.09%	66.66%	58.33%	64.28%
Stupide (3651)	47.11%	36.36%	58.33%	46.66%
Football (3600)	82.53%	100.00%	83.33%	64.28%
Ordinateur (4428)	16.66%	16.66%	33.32%	0.00%
Crayon (4939)	54.05%	75.00%	53.84%	33.33%

Table 6.5 Low frequency stimuli and their high frequency responses per year group in percentage

Not all low frequency stimuli elicit low frequency answers though; the cognate *loyal* (6135) elicits 78.2% of high frequency responses (UG1 = 66.66%; UG2= 83.33%; UG3 = 84.61%), and *football* (3600), 82.53% of high frequency answers (UG1= 100%; UG2= 83.33%; UG3 = 64.28%).

Fitzpatrick and Clenton (2017) report three studies in which the Lex30 test was used and note that the proportion of infrequent words produced in these three studies (words beyond the 1K band)

were 44%, 38% and 33%. These were estimated as successful rates and higher than other productive tests had reported. With proportions of infrequent words produced in the present study in the Flex15 November test of 49.32% (UG1), 51.5 (UG2) and 53.72% (UG3) as well as in the Flex15 May test of 50.87% (UG1), 51.74% (UG2) and 49.44 (UG3), the results show that not only is Flex15 an adequate tool to elicit low frequency words but also that the participants' lexicon has developed into these bands. So even if we have seen earlier that a way of tapping into more detailed productive knowledge would be useful, the productive vocabulary in this present study is developing well.

Nouns are the most common first answer to noun stimuli but also to verbs like *demander* (to ask). Some adjective stimuli elicit a vast majority of adjective stimuli such as *stupide* (91.5% of first answers are adjectives for UG1, 92.30% for UG2 and 86.66% for UG3). But other adjectives' first answer are less adjective driven, like the stimulus *loyal* which elicits 55.55% nouns, 33.33.% adjectives and 11.11% verb for UG1; 66.66% nouns and 44.44% adjectives for UG2; 53.84% adjectives and 46.15% nouns for UG3. The adjective *rapide* (fast) elicits the adverb *vite* (quickly) (44.44% for UG1; 66.66% for UG2; 57.14%) which is often mistaken for an adjective by students. The vast majority of the three groups' first answers are either paradigmatic or syntactic associations with the stimuli. The fact that there are hardly any phonological associations show that participants are all advanced learners, as this latter association is generally associated with beginners who have less knowledge of their vocabulary's grammatical functions.

Going back to the research question related to productive vocabulary, we conclude that the frequency factor does play a part in the development of this type of knowledge but that the rate of high frequency words is still much higher than that of lower frequency words. On the other hand, the test might not capture all aspects of productive lexical development. We did not find a significant correlation between receptive and productive vocabulary.

6.3 Proficiency

The next research question is: to what extent does overall L2 proficiency explain the pattern of L2 vocabulary development over time?

The results from the Elicited Imitation test show that the mean score from the 40 L2 French participants is 69.7 (out of a possible 120 maximum score) with a standard deviation of 23.37. This

average is close to Tracy-Ventura et al's mean of 62.9 when testing 29 UG2 L2 French language university students (2014: 152). The mean scores per year group increase from 56.0 for UG1, to 64.7 for UG2 and finally to 85.0 for UG3, suggesting, as one would expect, that proficiency improves with time and that second year students are more proficient than first year students but less proficient in their L2 than final year students, on the whole. It also shows that as reported by Wood Bowden (2016), EIT can discriminate between learners as UG3, the most advanced group, is statistically significantly different from UG1 and UG2. The standard deviations for UG1 and UG2 are very similar at 23.06 and 23.64 respectively but much lower for UG3 at 13.77. In other words, the greatest individual differences are found for UG1 and UG2.

How can we explain these individual differences in EIT? As the test relies on bottom-up information , that is to say learners hear the sentence, build a representation of it, hold it in their memory and then repeat it, performance on the task can be successful only if "listeners have sufficient lexical and grammatical knowledge about the target language" (Gaillard and Tremblay, 2016: 427). The great individual variation and the very low scores of some participants, in particular in UG1 show that some of them do not have sufficient L2 French lexical and grammatical knowledge to process the stimuli. This is a real issue at this level, where processing the language for incidental learning for instance is crucial. The weaker students therefore appear to be at disadvantage compared to others as their progress might be impeded. It has to be noted though that the standard deviation for UG2 reported by Tracy-Ventura et al (2014: 152) is 17.97, much lower than those reported here. More studies reporting EI scores and standard deviations are needed to determine if, as observed here, with time, there is more homogeneity in proficiency levels and also to determine to what extent different institutional groups tend to vary.

As for the range of scores, for UG1 the minimum score is very low at 16.0, compared with 29.0 for UG2 and a much greater minimum score of 66.0 for UG3. In terms of maximum scores, UG1 is 87.0 and UG2 105 and UG3 107.0. However, some UG2 participants score almost as highly on the EI test as some UG3 participants. These results seem to confirm that the proficiency level of the final year students who have spent the previous academic year in a French-speaking country is much greater and significantly different to first and second year students' proficiency level.

There is a strong positive correlation between EI scores and end of year exam grades *r =.* 64 and between EI scores and X-Lex *r* = .69; both of these correlations are statistically significant. A strong correlation between EI scores and end of year exam grades has already been observed before, for instance Tracy-Ventura et al (2014) reported a correlation of r= .78 between these two variables. However there is a weaker correlation between EI scores and Flex15 scores *r*= .30, which just fails

to achieve statistical significance. Similar results were reported by Tracy-Ventura et al with *r* =. 32. Having said that, it is possible that with a bigger group, statistical significance might have been reached. While the correlation between X-Lex and EI scores in this study is much stronger than in Tracy-Ventura et al (r= .12), in general these findings support Tracy-Ventura et al's conclusion of not using "this lexical test measure as an indication of individual proficiency development, unless that development is investigated across major time and concomitant learning intervals" (2014:156).

Next, let's consider individual differences in proficiency. For instance, out of the 13 participants reaching a proficiency score of 80 or more out of 120, 7 are UG3, 4 are UG2 and 2 are UG1, as shown in Table 6.6. This means that a couple of first year students' proficiency level is as high as that of some of the final year students. To achieve this level of proficiency, we can speculate that these UG1 and UG2 students must have engaged in a fair amount of speaking and writing, possibly of reading and listening too. If this is the case, either they received exceptional amounts of exposure to French before arrival at university, or else motivation must be driving them. We will come back to this point when analysing the motivation questionnaire. But many UG1 and UG2 participants who scored highly on the EI test also have a large receptive vocabulary hence supporting De Bot's idea that, lower proficiency levels are due to "a lower level of activation due to lower frequencies and less interaction as the network of this language subset will be relatively small" (De Bot, Lowie and Verspoor, 2005:46). On the other hand, it also seems that receptive vocabulary is not the only predictor of proficiency as some students with only average receptive vocabulary also can score highly in the EI test, which is the case for UG205 for instance.

Participant	**Proficiency score (out of 120)**	**X-Lex May score (out of 7000)**	**Flex15 May L3/L4/L5+ answers**	**End of year exam grade (in percentage)**
UG207	82	3700	9	77
UG101	83	2900	16	73
UG204	85	5100	19	79
UG105	87	5700	19	78
UG205	91	3100	12	73
UG202	105	6500	25	87

Table 6.6 UG1 and UG2 proficiency scores of 80 and over

As for the EI test's item difficulty, sentence length is very often an important predictor but other predictors of difficulty can include morphosyntactic complexity, or interaction between phonology and syntax, prosody and register as noted by Tracy-Ventura et al and confirmed in the present study. For instance, in both studies sentence 29 *l'examen n'était pas aussi difficile que ce que vous m'aviez dit* (The exam was not as difficult as you told me it would be) with its two clauses and 14 morphemes was more difficult to repeat than sentence 27 *le nombre de fumeurs de cigare augmente chaque année* (the number of cigar smokers goes up each year) which contains a similar number of syllables but only one clause and 9 morphemes. Out of the 10 most difficult sentences, 7 are common to all three groups, including sentences 29 and 30. This means that difficulties related to complexity in terms of clauses and morphemes persist over the years of instruction at university.

We also saw that lexical frequency could be a difficulty predictor as well. Item 12: *le restaurant est censé servir de très bons plats* (the restaurant is supposed to serve very good food) has the same number of syllables as item 13: *je veux une maison où mes animaux peuvent habiter* (I want a home where my pets can live) and yet the mean score for item 12 is 1.55 whereas the mean score for item 13 is 3.02. But in item 12, participants had more difficulty repeating the two low frequency words (*censé: 9832; plat: 4817)* whereas most content words in item 13 fall within the 2K frequency band (*maison, animal, habiter*). Other example of possible lexical complexity that could explain the low repetition rate of short stimuli include item 7: *Après le repas, j'ai fait une paisible sieste* (After dinner, I had a peaceful nap). This sentence was correctly repeated by only 5 participants out of 40, and this was mostly due to them being unable to repeat correctly the last two low frequency words: *paisible* (peaceful) and *sieste* (nap), found in 3347th and 7770th place respectively in Baudot's frequency list.

Lexical frequency could also explain why most participants can repeat some sentences with a fair amount of syllables quite easily, like items 27 (17 syllables) *Le nombre de fumeurs de cigares augmente chaque année* (The number of cigar smokers goes up every year), and item 25 *Le voleur que la police a arrêté était très grand et mince* (The thief that the police arrested was very tall and thin) (16 syllables), with means of 3.15 and 2.55 respectively. In sentence 27, all the content words are from the 1K frequency band apart from *cigare* (cigar) which is in the 6K band but is also a cognate. In item 25 *Le voleur que la police a arrêté était très grand et mince* (The thief that the police arrested was very tall and thin), again all content words are from the first 2K frequency band apart from *voleur* (thief) from the 4K band, which, if the lexeme was not fully understood might be helped with the associative power of "police". Designing proficiency tests such as EITs

including only very high frequency words would have the advantage to eliminate lexical difficulty in order for researchers to focus on other sources of difficulties such as syntactic ones. As we take the view that grammatical knowledge is embedded in lexical items, it would be interesting to investigate whether some morpho-syntactic difficulties are lessened with the use of high frequency vocabulary and if so, which ones. But going back to the research question, we can say that the overall L2 proficiency is related to receptive vocabulary hence showing that the larger one's vocabulary is, the more proficient he or she is.

6.4 Individual differences and exposure to L2 French

6.4.1 Motivation

The other question investigated in this study is: to what extent can the individual difference factor of motivation explain learners' variation in L2 development? Referring to Dörnyei (2005)'s L2 Motivational Self System, the aim was to find out whether motivational factors could influence vocabulary development and proficiency development. Indeed, the three components of this system, namely *Ideal L2 Self, Ought-to Self* and *Learning Experience* have a strong influence on learners and on their L2 development. The relationship between proficiency and motivation might have been researched more often than vocabulary and motivation, but we believe the link between these last two has some great potential to explain variation in university settings (Tseng and Schmitt, 2008). In addition to investigating the relationship between vocabulary learning/motivation and proficiency/motivation among L2 French learners in a UK university, we also consider critically the relevance of the L2 Self model for motivation for learning languages other than English (LOTEs) as recommended by Ushioda and Dörnyei (2017).

The results of the motivational questionnaire for all 35 participants showed that the three top motivational factors were attitudes towards the L2 community, interest in the French language and attitudes towards learning French, and that the least important factor was the Ought-to Self**.** Interestingly, when the results are analysed per year group, the top and bottom factors for all three year groups are the same as the whole group. The top motivational factors seem to suggest that these advanced learners are motivated by their own experience of L2 learning and L2 use, hence supporting Dörnyei and Ushioda's point that advanced learners "may feel motivated to pursue language study because they perceive that this is what they are good at or what they

enjoy most, and where therefore their future potential must lie" (2011: 61). But as we have seen from the language learning questionnaire, which we will discuss a little bit later, the Ideal Self too seems to play an important role as becoming fluent is the most commonly given answer by participants to the question why they decided to study French at university.

Instrumental-prevention also plays an important role, in fourth place. Participants are high achievers because to study languages at university, they had to gain high grades; this motivation factor is still with them, they do not want to be considered as weak students. As Lanvers stated, these students are "keen to comply with external demands and prevent negative outcomes" (2017: 525). Instrumental-promotion is in fifth place, that is to say students consider studying French useful to get a job, or to study abroad for a long period of time, but it is not the most important factor. And yet, we cannot conclude that the international status of English has diminished the extent of instrumental motivation for learning French at university firstly, because it still ranks in 5th place out of 12, secondly, because many other recent studies have shown it is a motivational factor for language learning in UK universities (Canning, 2011; Busse, 2013; Oakes, 2013). Thirdly, instrumental-promotion is frequently cited in the motivation questionnaire, as 43% of the participants explicitly say that they study French because they want to work abroad or because speaking French will be beneficial for their careers. The relatively low position of instrumental-promotion might be due to the questions included in the questionnaire. It is also possible that some participants study French for some future career, but because they do not have a specific career in mind yet, they might score lower on these types of questions. After all, as Lanvers states "language degrees are considered to have less obvious follow-on career paths compared to some other degrees" (2017: 520).

Interestingly, ideal self and culture interest are in the middle range though they are often cited as reasons why participants decided to study French at university. Integrativeness comes 8th showing that even if students have positive attitudes toward the L2 community, that is to say that they like to go to French-speaking countries, to meet people from French-speaking countries and to learn more about people from these places, they are not particularly willing to integrate. This attitude on the part of UK language students not wanting to integrate has also been observed by Oakes (2013) and Busse and Williams (2010). Mitchell et al (2017) also found that UK students of languages on their year abroad are not strongly motivated to integrate either. Rather, they see themselves as global citizens who can go around different countries using languages that they consider as useful skills. As we have seen, if "international posture" that is to say having an interest in the target language, being willing to go overseas to stay and work and to interact with

the L2 community without wanting to be more like them is a concept found among users of English as a lingua franca like in Japan (Yashima, 2002) or Indonesia (Lamb 2004), research increasingly show this is also the case for learners of LOTEs, at least language learners from the UK.

The bottom two motivational factors are French anxiety and Ought-to self. The French anxiety factor is quite low for these high achieving students, presumably because they are strong language learners. In addition to a good academic track record throughout their academic life in French, a lot of them have enjoyed their language experience throughout school and college as shown in the answers given to why they have chosen to study French at university.

As for the Ought-to Self, being the lowest ranked motivation factor, this means that learners of LOTEs in the UK do not feel pressure from parents, teachers or peers to learn languages. In that sense, the data supports Dörnyei and Al-Hoorie's (2017) idea that the Ought-to Self might not even be that useful for investigating the motivation of learning LOTEs. In the answers provided by participants, one participant declared overtly than he/she does not want to speak only English and by extension to be seen as a monoglot British speaker. Apart from this example, there is no evidence of the anti Ought-to self as named by Thompson and Vásquez (2015), or of the rebellious profile identified by Lanvers (2016). That is to say that unlike Oakes (2013) who found that university learners of LOTEs in the UK learn an L2 to defy the belief that British are mostly monoglot, we did not find this.

However, in the current model, there is no component to capture this potential motivational element of doing something because one does not have to. So it could be, as Dörnyei suggested, that it is a matter of rethinking the relevance of the Ought-to Self factor or maybe of developing a new component. Also, in the UK, to study in a "good" university, students need high grades. In a sense, is this not connected to the Ought-to Self? And despite lack of evidence in this research project, there is considerable anecdotal evidence that, once at university, students continue to aim at high grades to get a high degree classification, often associated with better career prospects after university. Again, this type of motivation could not be equated exactly with instrumentation-promotion and could possibly be part of the Ought-to factor. Therefore, it might be necessary when investigating LOTEs to expand the question beyond "the individual-psychological dimension to a macro-sociological perspective by considering factors in the socio-political and educational environment" (Ushioda and Dörnyei, 2017: 452).

6.4.2 Motivation and lexical development

Analysing the motivational factors of the students who scored the lowest on the receptive vocabulary test in May, instrument-promotion was slightly higher than for the whole group and so was French language anxiety. Other motivational factors were not very differently ranked than for the whole group, so that the weakest group does not differ hugely from the whole group apart from these two motivational factors. The fact that French language anxiety in particular is higher is interesting because this could have some negative impact on learning. It is well known that "a tense classroom climate is one of the most potent factors that undermine learning effectiveness and L2 motivation" (Dörnyei and Ushioda, 2011: 110). Language anxiety can also have some negative effects on autonomous learning as well, as discussed later.

The group with the largest vocabulary size has however a rather different motivational profile from that of the whole group. First of all, the first motivational factor for this group is instrumental-prevention, compared to being fourth for the whole group, meaning that participants who have a large receptive vocabulary towards the end of the academic year, are motivated to work in order to avoid obtaining low grades. This supports the view that an active and independent attitude to learning is beneficial (Barfield and Brown, 2007; Benson, 2007, 2010; Little, 1991; Pemberton, Toogood and Barfield, 2009).

Culture interest and linguistic self-confidence are also more important for this group, the former jumping from 7th place for the whole group to 3rd and from 9th to 6th for the latter. This means that in addition to spending time on learning French, they do so while accessing resources about the culture that will enhance their receptive vocabulary, but not only this. As they are also more confident in their use of French, they might access resources that less confident users would not use, or engage in activities that others might not, like befriending Erasmus students or joining French speaking clubs to practise their speaking. So the data seem to support that the three factors that are most important to develop a large receptive vocabulary are: spending time learning, engaging with authentic resources about the culture, and being confident in one's abilities.

	Ranking for lowest achieving participants	Ranking for highest achieving participants
attitudes towards L2 community	1	5
instrumentality-promotion	3	8
instrumentality-prevention	6	1
linguistic self-confidence	10	6

Table 6.7 Major motivational factors differences between lowest and highest scores in vocabulary test

When comparing the lowest and the highest achieving participants' results, they differ on four motivational factors, as we can see in Table 6.7. It is clear that prevention and linguistic self-confidence are the most important factors as far as vocabulary development is concerned. On the one hand, students who have a smaller receptive vocabulary size are more motivated by the prospect of travelling to French-speaking countries, of learning about French speaking countries and of living and studying abroad. On the other, participants with higher receptive vocabulary size are motivated by studying hard to avoid being seen as weaker students and they believe they are good at French. Good grades are not everyone's first main goal when studying languages and should not be seen as the goal to have in mind to develop large receptive vocabulary, but rather the way to get good grades is key, that is to say, investing time studying beyond the classroom, and expanding the learning experience to self-study time.

The second motivational factor that differs between the low and the high groups is linguistic self-confidence which is higher for high achievers. As we know, "if students have basic doubts about themselves, they will be unable to bloom as learners" (Dörnyei and Ushioda, 2011: 120) and as such it is important to find strategies to maintain and improve linguistic self-confidence. Many strategies have been designed to help build learners' self confidence in the context of the language classroom and to help them "to dispel misconceptions and unrealistic fears" (Dörnyei and Ushioda, 2011: 78). But given that language teaching contact time is quite small for university students, , it is possible that the low achievers' negative attitudes about their linguistic confidence affects their independent learning, that is to say, they are not confident about their ability to learn French as autonomous learners. One way to improve their motivation when the teacher is

not present or when they are not working with other learners is to remind them of relevant self-regulatory strategies, leading to self-motivation as defined by Ushioda (1997: 41):

> This capacity entails taking personal control of the affective conditions and experiences that shape one's subjective involvement in learning. It entails minimising the damage when these experiences are negative, and maximising the subjective rewards when these experiences are positive, and so fostering optimum motivational conditions for continued engagement in language learning.

In addition to self-regulatory strategies, to improve students' low linguistic self-confidence, teachers can encourage these students to evaluate their achievements in a positive light. Attributions of test results and other assessment outcomes to effort rather than ability will also make them realise that they can change, and spend more time to improve next time which is a very important factor in vocabulary and proficiency development. Also, L2 vocabulary learning for advanced university students becomes less deliberate and more incidental. There is plenty of anecdotal evidence that some students can find this shift difficult because they feel they are faced with learning thousands of new words in order to become "fluent". One way of improving their confidence in vocabulary is to share useful information about vocabulary learning such as the coverage the 2,000 most frequent words offer. A receptive test could be made available to them to reassure them that they already have the knowledge required to build upon to become more autonomous learners too. Other comprehensive information about the necessity to revisit and to use new words could be shared too as well as strategies and examples for how to do it.

6.4.3 Motivation and proficiency

We also considered the relationship between motivation and proficiency and found that instrument-promotion and French anxiety were the two motivational factors that characterised the low proficiency group. Interestingly, they were also the factors we noticed for the low vocabulary group. In mainstream motivational research, self-efficacy plays "a role in whether a possible self seems attainable and therefore triggers motivated behaviour" (Busse, 2013: 392). Indeed if a learner is able to see that a goal is within his or her reach, he or she will invest more effort. The present data support this finding by explicitly showing the link between negative self efficacy and low attainment in both vocabulary and proficiency development. In other words, students who have higher anxiety towards French and have lower linguistic self-confidence most

probably spend less time on French because they are anxious about their ability and consequently fall behind.

In addition, as MacIntyre, Noel and Clément (1997: 278) found out, anxious learners tend to underestimate their competence compared to less anxious learners who tend to overestimate their performance. So, not only might more anxious students spend less effort learning, but when evaluating what they can do they underestimate themselves. No wonder in these circumstances that they might find "confronting their perceived limitations as painful and demotivating" (MacIntyre and Gardner, 1989). But however pessimistic their view on their progress might be, the present data does underline that it concretely hinders proficiency and therefore self-efficacy needs to be maintained in order for advanced learners to develop their mastery of the language.

Many tasks in L2 learning can be sources of anxiety. Speaking and writing that involve communicating with others for instance can trigger anxiety, and even comprehension tasks such as listening when one might feel anxious to ask for clarification if failing to understand part of the message (MacIntyre et al, 1997: 279). But the two statements related to French anxiety in the motivation questionnaire were directly referring to speaking: "I get nervous when I speak in my French class" and "I worry about making mistakes when speaking French", therefore the link between French anxiety and proficiency is clearly present. We take MacIntyre et al (1997) and Young (1991)'s view that this link is reciprocal. In other words, if anxious learners are reluctant to speak because they are afraid of making mistakes for instance, they do not get a chance to re-assess their actual proficiency level and that in turn increases their anxiety level. In addition to that, these students do not progress as quickly as others who speak since, as we have seen, they do not have as many chances as less anxious students to practise their L2 competence, strengthening their lexical knowledge and improving automatic retrieval of words. Anxiety increases with less speaking and less speaking clearly impedes proficiency. This corresponds with Busse's findings too, that self-efficacy "clearly plays a role for students' engagement with language learning" (Busse, 2013: 393). In the L2 motivational model, this would also suggest that anxiety also curbs the Ideal L2 Self-development as the gap between the existing Self and the future Self, a Self that will be fluent, is not narrowing.

For the highest proficiency group, the two motivational factors that diverge the most from the whole group are culture interest, placed in third place for this group compared to 8th for the whole and linguistic self-confidence, in 4th place for the high achievers in proficiency and 9th for the whole group. What is very interesting is that culture interest and linguistic self-confidence were also the two motivational factors that differ the most for the high achievers in vocabulary. In

other words, motivational factors can indeed to some extent explain students' variation, but also the motivational factors that seem connected to the development of both vocabulary and proficiency are the same: culture interests and linguistic self-confidence. That is to say students who believe they are good at French and are interested in the culture and learn about it by explicitly saying they are accessing a range of authentic materials are developing their vocabulary and their fluency to a greater extent. This support the notion that incidental learning is key to vocabulary development at this advanced level, but also that a positive "can do" attitude is very important in language learning.

Table 6.8 below reminds us that the top three motivational factors are the same for low vocabulary and low proficiency: attitudes towards L2 community, interest in the French language and instrumental-promotion. A vision of an L2 self speaking the language and using it to travel is motivating but if it is not followed by a motivational factor that explicitly refers to spending time and effort accessing authentic materials either for gaining grades or for finding out about the culture, the result may be a low level of development in both vocabulary and proficiency.

Spending time and effort learning outside the classroom is of course what teachers recommend students to do and the fact that this advice is not always followed is a well-known problem for language teachers. However, one possible way to improve motivation to dedicate more time and effort can be to demonstrate and regularly emphasise how vocabulary gain means more ease in learning and therefore more ease in accessing a range of authentic resources. This could be a way of improving low linguistic self-confidence whilst supporting Ushioda and Dörnyei (2017: 453)'s argument that "a less instrumentalist and more holistic view of language learning as a process of expanding people's meaning-making repertoires" should be favoured. In addition to reinforcing strategies and the message that incidental learning outside the classroom is key, Oakes (2013: 189) also suggests students should have more contact time, that is to say more language classes as well as the opportunity to have content modules (i.e. history, literature modules)delivered in French rather than in English. As incidental learning with the help of a teacher leads to more vocabulary learning, this suggestion has the possibility to improve all students' vocabulary but particularly that of low achievers, which in turn might give them more confidence and therefore more autonomy in their learning outside the classroom. On the other hand, because direct vocabulary teaching in class can be very time consuming and not always successful, the goal of more contact time should be to help learners become more independent so as to develop their own vocabulary.

	Vocabulary	Proficiency
Low group	1. Attitudes towards L2 community 2. Interest in the French language 3. Instrumentality-promotion	1. Attitudes towards L2 community 2. Instrumentality-promotion 3. Interest in the French language
High group	1. Instrumentality-prevention 2. Interest in the French language 3. Culture interest	1. Interest in the French language 2. Attitudes towards L2 community 3. Culture interest

Table 6.8 Major motivation factors associated with vocabulary and proficiency development

Going back to our research question, we can conclude that motivation can to some extent explain individual differences in performance. High achievers in both vocabulary and proficiency tend to be linguistically confident and interested in the L2 culture. They declare accessing authentic materials to learn more about the L2 culture. On the other hand, low achievers in vocabulary and proficiency tend to be more anxious and instrumental-promotion is a feature of this group too. But while they have a clear vision of why learning L2 French could be useful for the future, the data suggest that they are not narrowing the gap between this Ideal Self and their current Self as their L2 French development is lagging behind.

6.4.4 Exposure to L2 French

The last research question of this project is: what is the relationship between the extent of exposure to French and vocabulary learning? To investigate participants' exposure to L2 French outside the classroom, the activities they listed were categorised according to Nation's four strands (Nation, 2010, Webb and Nation, 2017): meaning focused input, that is to say learning through reading and listening; meaning focused output, learning through speaking and writing; language-focused learning, deliberate attention is given to language features; fluency

development, use of familiar materials as quickly as possible in listening, speaking, reading and writing. The rationale for the four strands proposed by Nation (2007) states that "for vocabulary learning programme to be well balanced, the time spent both inside and outside the classroom should be divided equally" between the four strands (Webb and Nation, 2017: 179). Classifying activities in this way enables us to investigate the activity patterns of advanced learners in a university programme during independent study, and whether there is a relation between these patterns and vocabulary development.

Before focusing on vocabulary development tasks, let's summarise the data found about the tasks undertaken to develop all aspect of L2 outside the classroom. 49.5% of these refer to activities categorised as meaning-focused input, while at 33.7%, language-focused activities are the second most common type reported, and include keeping lists of vocabulary and idioms. Finally, 16.8% of language development activities can be classified as meaning-focused output, which include speaking with Erasmus students or during Café Parler, organised by the French society at university. There are no activities that could be classified as fluency development, but that is rather due to the definition of the strand than a total lack of fluency activity among the participants as, for instance, speaking in French to Erasmus students can usually be seen as fluency development. The other issue with the data about L2 French exposure is that the frequency of occurrence for these activities is not always provided and when it is, it can be reported vaguely as "once or twice a week" (UG209), "frequently" and "very infrequently" (UG203) to "once every few months" (UG202).

Bearing in mind therefore that frequency information should be considered cautiously, what the results show is that despite Webb and Nation's recommendation, there is a real imbalance within the strands as almost half of French language activities outside the classroom are meaning-focused input. One of the reasons is probably that reading and listening materials are much easier to access. Materials or links towards materials are made available to students in all language modules but also, they have access to the university library, not to mention the internet, another great source of French meaning focused input. Unlike the advocates of meaning-focused input like Krashen (1981, 1985) who believes activities like reading are the most important ones in L2 development, Nation advocates a variety of meaning-focused input but also a more balanced distributions of the activities across the four strands as we have seen above. The result that language-focused activities represent the second type of activities is on the other hand very positive and shows that students are leading their own deliberate vocabulary learning using flashcards for instance. This type of "individualised deliberate vocabulary study has been

thoroughly researched and shown to be highly effective" (Webb and Nation, 2017: 186), that is to say leading to increased levels of both receptive and productive knowledge.

However, as we have seen in chapter 3, the form-meaning knowledge that can be reinforced using flashcards needs to be expanded to include both receptive and productive knowledge. The fact that only 16.8% of activities are meaning-focused output not only means that the productive knowledge we have just mentioned is hampered but this might also have a negative effect on students' motivation. After all, when asked what the main reason was for them to study French at university, as we have seen in the previous section, becoming fluent was the most common answer. Of course, meaning-focused output is somehow more difficult to put in place, particularly as part of independent language study, because it either requires someone else to speak too or the idea of someone having to mark written work. But there might be ways within the course to develop independent activities that could lead to more speaking and more writing without taxing other people or requiring too many extra resources as we will suggest in the conclusion.

Let's now discuss the type of exposure for developing vocabulary in particular. Out of the 190 language learning activities listed, 72, that is to say 37.9% were directly concerning the development of French vocabulary outside the classroom. The majority of vocabulary related activities are language-focused for all three groups (UG1, 80%; UG2, 71%; UG3, 74.1%) and include, as we have mentioned, activities such as making electronic or paper flashcards. Vocabulary related activities that are meaning-focused output is low: 15% for UG1, 0% for UG2 and 3.7% for UG3. It is not clear why the UG1 percentage is higher as they do not seem to have more opportunities to speak French than the other two groups, in particular UG3 who might have kept contact with L1 French native speakers who they met during their year abroad.

6.4.5 Exposure and lexical development

Turning to the link between exposure and vocabulary learning, we considered again the group who scored the lowest in the May X-Lex test and the group that scored the highest and looked at the type of exposure they had. The number and the type of activities that both groups listed did not differ much, apart from the meaning-focused input of reading which is more present in the responses of the high X-Lex group. The differences between the two though were more linked to the frequency of the activities, with the high group being more specific and showing that activities were taking place on more regular basis than for the low group. In addition, there seems to be some disengagement among the low group which lists activities such as listening to the radio in

the background. Similar findings from the low proficiency group lead us to conclude, from the data available, that lower level of vocabulary and lower level of proficiency are a result of a few factors including: an exposure to French that is too limited, participants' attention to the language from the input activities that is too low and inadequate access to resources appropriate for their level.

For successful incidental learning, defined as per Huckin and Coady (1999: 185) as "the learning of new words as a by-product of a meaning-focused communicative activity, such as reading or listening, and interaction", students need to be actively engaged for learning to take place. This of course refers as well to Schmitt's noticing hypothesis we have discussed in chapter 3 according to which attention from the learner "appears to play a crucial role in both implicit and explicit language learning" (2000:9). So activities undertaken without sufficient willingness to notice the language are unlikely to impact on language learning. But the level of noticing also depends on the level of proficiency of the learners, that is to say, a more proficient learner will have a larger vocabulary size that will enable him or her to notice and focus more on new words. And as this lower level involvement during incidental learning seems to be a characteristic of students whose vocabulary and proficiency is lagging behind, it might be that the involvement issue is a by-product of their L2 capacities.

Interestingly, what we can also notice from the exposure data is that participants are aware that they should be practising the four skills of reading, writing, listening, speaking, autonomously, even though as we have seen there is in fact an imbalance between the types of activities. They also make the most of resources they have found useful in their language learning history such as flashcards. All this means that, as expected of advanced and motivated students like these ones, they are willing to take on board strategies and advice that would enhance their learning. One of the strategies that we can reinforce and that Webb and Nation (2017) advocate is to equip learners with strategies to learn how to become a more efficient autonomous learner and developing vocabulary beyond the 3K band in order to be able to guess successfully the meaning of new words from authentic materials.

Going back to our research question, we can conclude that not only there is a link between exposure and L2 French vocabulary learning but we also notice that the level of engagement plays an important role in incidental learning of these participants. Reading seems to be the activity that really drives the vocabulary acquisition of students though. We also noticed that there is little speaking and writing practice outside the classroom. This could slow the development of

productive vocabulary knowledge whilst potentially also demotivating some students whose main aim is to become fluent in the language.

6.5 Reflections on Levelt and Nation's frameworks

The results discussed in this chapter have also some pedagogical implications for the teaching and the learning of L2 French at advanced level that will be discussed in the next chapter. For now though, in the light of these findings, we will reflect on the theoretical frameworks adopted in this research project, Levelt's language processing model, related usage-based theories of L2 development, and Nation's word learning in instructional contexts. With Levelt, we have adopted the view that the mental lexicon is central to both speech and comprehension, and is accessed through the different modules that mirror each other for the speaking and listening processes. For instance, in speaking, the preverbal message that originates in the conceptualiser passes through the formulator where lemmas are retrieved, activated and selected; they consequently "activate syntactic procedures that correspond to their syntactic specifications" (Levelt, 1993: 6). In other words, syntactic form originates from meaning, via lexical selection; this is the lexical hypothesis. Furthermore, we have adopted the usage based perspective on L2 acquisition supported among others by Verspoor, Lowie and De Bot (2008), according to which language processing can be equated to language learning, so that for example each time a new L2 lemma is activated, it becomes better developed and more richly connected to other elements in the lexicon.

In this section, we summarise in turn, how the theoretical framework has helped us interpret particular findings of the study, as follows:

The place of words in L2 learning

Word class preferences

Role of input/ input frequency in vocabulary acquisition

Role of intentional learning and noticing

Language education

6.5.1 The place of words in L2 learning

First of all, the study shows a strong correlation between X-Lex May scores and EIT scores hence supporting Levelt's model according to which word knowledge is central to proficiency; that is to say, the more L2 words learners know, and the better they know them, the more grammatical structures they can activate, and consequently the more proficient they are overall. The fact that there is a noticeable frequency factor in both X-Lex Nov and X-Lex May tests also confirms the view of usage-based theorists such as De Bot et al. that input processing is central to L2 learning, i.e. learners process and therefore learn frequent words and their syntactic features first. Consequently and as demonstrated by the Flex15 tests data, frequent words are also retrieved better. In a recall or naming task, this would have also influenced the speed of access as high frequency words are generally accessed faster than low frequency ones. This was a written exercise so there is no information about the speed of access.

In this study, we have also adopted the view that learning of L2 words is incremental. Learning is first the creation of a L2 form-meaning link and then gradually syntactic information about the words is learned, as instances of its use in different L2 contexts are processed. However, the tests that we have used in this study do not allow us to investigate the suggested gradual nature of L2 learning that is to say from fuzzy form-meaning links to elaborated lemmas.

What we have noticed though is that some closed word categories such as interjections and conjunctions are more difficult to process for all learners. For instance within the 5K most frequent words in the X-Lex Nov test, closed words that were tested include the prepositions *sur* (on) and *environ* (around), both in the 1K band and the pronoun *sien* (his/hers/ones), the interjection *hélas* (alas) and the conjunction *sinon* (otherwise) found in the 2K band. This pronoun, this interjection and this conjunction, despite their high frequency, were not easily recognised by participants, to the extent that, as we have seen in Chapter 5, all three year groups were 2K deficient, in large part due to these words. So despite their high frequency in language use and regardless of learners' proficiency, these words have not been learned. In De Bot's interpretation, the semantic information from the conceptual system has not been copied onto the lemma. In his model, Levelt states that word frequency and semantic markedness each have their effects at the level of lexical access proper, but these examples show that semantic markedness can have more effects than word frequency on accessing some word categories such as conjunctions for instance.

6.5.2 Word class preferences

The main aim of the Flex15 tests was to yield some information about the development of productive vocabulary and its change over the years in terms of the frequency factor as has been discussed in Chapter 5. However, we also took this opportunity to look into the types of relationships between items in the lexicon when analysing the first response to each stimulus word. We found that nouns are the most common first answer to noun stimuli but also to verbs like *demander* (to ask). The vast majority of the three groups' first answers are either paradigmatic or syntactic associations with the stimuli. The fact that there are hardly any phonological associations show that participants are all advanced learners, as this latter association is generally associated with beginners. But the preference for a particular word class still needs some explanation. As we have seen above, our EI data suggests that lemmas of content words, particularly of nouns (their meaning and their syntactic information) are easier to acquire than lemmas of function words. This in turn could mean that in the productive test, nouns are most accessible for processing, and that therefore nouns are activated first. In Levelt's model, in speaking for instance, it is suggested that "there is a preference hierarchy for grammatical functions, with the subject function being the most highly valued one" (Levelt, 1989: 233).

However, as we have seen, most of the first answers are also high frequency words, that is to say from the first 2K bands from Baudot's French frequency list. So it seems that whatever the type of word category, frequent words will be accessed first. What we could speculate is that nouns are the word category that tend to be semantized the fastest, and when this is coupled with their frequency, they are the most common category produced.

6.5.3 Role of input/ input frequency in vocabulary acquisition

As we have seen, acquiring a word in L2 is first knowing its meaning (supported by connections with L1 words and their meanings), and then gradually building syntactic information about it. Even though it is not known how much of this information is learned when a learner encounters a new word, it has been recognised that multiple encounters in different contexts of the same words are needed to fully develop this knowledge. Nation supports this idea and believes that in L2 instruction, many encounters with the same words are needed for their complexity to be grasped by the learners, so that input should be both varied and frequent.

In the context of the study, the L2 input accessed by participants outside the classroom was investigated and most of the activities linked to this self-study input can be qualified as incidental learning, so that learning new words is a by-product of those tasks. If incidental learning is the

way L1 words are generally learned, we have comparatively little evidence about its effect on L2 vocabulary. Yet it is a widely accepted idea (Webb and Nation, 2017: 54) that an incidental approach to L2 learning is beneficial. What this study shows is that this approach has indeed a positive effect on L2 vocabulary development among advanced learners, because when participants regularly undertake meaning focused input activities such as reading and listening, this leads to a bigger receptive vocabulary and a better proficiency level. The model that underpins this study also suggests that high proficiency learners should be more efficient incidental learners, because they have deeper and richer knowledge of the surrounding words to which the new word is connected, they can make more use of the lexical context than lower level learners. Participants whose receptive vocabulary is small tend to undertake meaning focused input activities too but less regularly, and the way they report these activities seem to suggest that they approach them in a less engaged manner.

The results validate the usefulness of an incidental approach to vocabulary learning for advanced learners. In a university setting with a variety of tasks and the amount of time dedicated to self-study when this incidental learning can take place, this is good news. However, the study also confirms that not all learners benefit the same way from this approach and from the same input. Indeed, we have noticed that there can be great variety of lexical knowledge and proficiency within year groups, even though this is to some extent overcome by the rich and prolonged L2 input offered by the year abroad .

6.5.4 Role of intentional learning and noticing

Why do some learners seem to benefit more from incidental vocabulary learning activities than others do? To answer this question, we turned back to the psycholinguistic processing and acquisition point of view according to which the learner “is not a passive container that receives some input, but a very active problem solver who anticipates, predicts, and thereby interacts with the input continuously in a constructivist manner” (Verspoor, Lowie, De Bot, 2008: 19). In Nation’s framework, the same idea is shared, that is to say for learning to take place, the learner should be able to notice and Nation notes that “motivation and interest are important enabling conditions for noticing” (Nation, 2010: 63).

The present study does confirm the positive effect of intentional learning and noticing on L2 vocabulary development and on L2 proficiency. This is shown in the finding that instrumental-prevention type of motivation did correlate with high receptive vocabulary scores and higher levels of proficiency.

Additionally, in Nation's framework, learners should use the vocabulary they learn to automatize the use of these words; according to Nation, the four strands should be as much as possible well balanced. However, in this study, as far as the incidental vocabulary activities undertaken by participants are concerned, we found evidence of an imbalance between meaning focused input (learning through reading and listening), meaning-focused output (learning through speaking and writing), language-focused learning (deliberate attention to language features) and fluency development (use of familiar materials as quickly as possible in listening, speaking, reading and writing). These findings point clearly to a need for pedagogical intervention that we will develop further in the next chapter.

For both Levelt and Nation, input processing depends to some extent on the learners' engagement and ability to notice. However, the present research study has already shown that not all input can be processed in the same way. In other words, some input is more resistant to assimilation. For instance, the EIT results show that out of the 10 most difficult sentences to repeat, 7 are common to all year groups. This also seem to confirm De Bot's point according to which "within a lemma, meaning and syntactic information may not be inextricably linked" (De Bot, 1992: 21). This could be the case in those 7 sentences and it would be interesting to investigate in more detail in the underlying semantic and/or syntactic reasons why those cases are more difficult to process than others.

6.5.5 Language education

The psycholinguistic processing and acquisition framework and the instructional framework we have selected have implications for education policy. It is important when investigating the vocabulary learning of the participants to consider the context in which they learn the language, in this case, L2 French at university in the UK. The language learning of our participants has been shaped by the language policy of the country that has influenced not only the way students have been taught but also what they have been taught. We showed in Chapter 2 that L2 achievement levels in much of UK language education are often problematic. Understanding the reasons why learners know what they know is important because like Levelt from a psycholinguistic point of view and like Nation from an instructional point of view, we believe that ongoing L2 learning is influenced by what is already known. There is in this view a real link between language policy, language teaching, language learning and language use.

Measuring language and in particular measuring vocabulary in this research project, as we have seen, is not at all easy. We have adopted the frequency model to track acquisition of vocabulary, and despite some limitations in our tests, we have been able to contribute new knowledge on the

extent of vocabulary known by specialist languages undergraduate students. These results have clear pedagogical implications, which will be detailed in the conclusion.

6.6 Conclusion

This chapter has summarized the present study's findings and discussed them with reference to each of the research questions and in relation to relevant previous studies. The study confirms that the frequency factor is observable in the development of the receptive and productive vocabulary knowledge of advanced learners and that "frequency is a key determiner in the learning of vocabulary" (Brown, 2012: 20). However, there are also some individual differences and whereas the frequency factor is observable across the first 5K bands by the end of the academic year, many participants do not display this regular profile. The study also confirms previous studies' findings that the uptake of receptive vocabulary within an academic year is small. This suggests that once students have reached the 2,000 word threshold, the quality of their receptive vocabulary knowledge rather than solely its quantity seem to improve. The study also demonstrates that more advanced learners know rarer words and are able to access them more easily than less advanced ones. However, no systematic correlations between receptive and productive vocabulary was found at the end of the academic year.

Whilst the three year groups were not significantly different in May as far as their receptive and productive vocabulary were concerned, the study on the other hand did point to a significant difference in proficiency between UG3 and the other two groups. The data also suggest that individual differences which are quite large in EIT for UG1 and UG2 tend to reduce for UG3. Correlations between the Elicited imitation test and end of year exam grades as well as receptive vocabulary were found. It was also suggested that the EIT does display some lexical difficulty that can be attributed to the frequency factor.

The motivation questionnaire revealed that high achievers in both vocabulary and proficiency tend to be linguistically confident and interested in the L2 culture whilst low achievers in vocabulary and proficiency tend to be more anxious and to favour instrumental-promotion. Not only does this have some pedagogical implications as we have seen, but it also shows that there is a strong link between receptive vocabulary and proficiency development. The study also confirms that some of the motivational constructs, such as the Ought-to Self, that were typically used for

studying learners of L2 English are not entirely suited for the study of motivation for learning LOTEs.

Finally, the study confirms that exposure has an impact on L2 lexical and proficiency development. Language learning at this advanced stage is largely done through incidental learning and it was identified that the commitment of time, the relevance of the resources and the engagement of the learners are essential for vocabulary uptake. The study also pointed out an imbalance in the activities undertaken to develop L2 French outside the classroom, with few opportunities for speaking, writing and fluency development. This an important area that need to be addressed pedagogically as productive language use is essential to strengthen vocabulary knowledge and to develop proficiency.

Chapter 7 Conclusion

7.1 Introduction

In this chapter, a summary of the key findings are given as well as considerations for pedagogical implications for the teaching and learning of L2 French at university. Limitations of the current study as well as implications for further studies will be discussed.

7.2 Key findings

The main aim of the present study was to investigate the lexical development of advanced university students of L2 French in the UK with a unique triangulation approach to see how this lexical development would be associated with overall proficiency and motivation. The multi tools approach enabled us to gather data about vocabulary development of participants from three different year groups as well as longitudinal data over one academic year for each year group. The productive vocabulary data from the X-Lex tests showed that UG2 was the only group that made statistically significant improvement from November to May. However, out of all three groups, UG1 was the one with the highest receptive vocabulary gain but due to the group size it failed to reach statistical significance. The other finding from the X-Lex was that despite a statistically significant difference between UG1 and UG3 in November, by the end of the academic year, all three groups had relatively similar mastery of the first 5K words. The receptive vocabulary results also confirm prior studies' findings such as David's (2008) that show that L2 French advanced learners tend to display quite modest uptake in receptive vocabulary from one year group to the next at undergraduate level. This would seem to demonstrate that, at advanced level at least, vocabulary does not only develop in size but most probably in quality. The frequency factor is also observable from the first to the fifth thousand bands with some individual differences. For instance, despite groups displaying a frequency effect from 1k to 5k bands at the end of the academic year, almost half of the participants do not.

The Flex15 test also yielded some interesting findings about the participants' productive vocabulary. For instance, as expected, more experienced learners know rarer words and are

producing them more easily than less advanced learners. But on the whole, students' productive vocabulary tends to be majorly made up of high frequency words, that is to say, of words from the first 2,000 most frequent words. Rarer words are acquired and there is an institutional effect to this acquisition in the sense that UG3 have access to more rare words than UG2 and in turn, UG2 tend to know more rare words than UG1. However, the development of rarer words seems slower than the development of high frequency words. No systematic correlations were found between the receptive and the productive vocabulary of the participants. It is difficult to say whether this is due to the fact that both receptive and productive vocabulary knowledge are simply different or if the tests used are possibly responsible for this lack of correlation.

The Elicited Imitation test clearly showed that the proficiency of UG3 was clearly different from the proficiency of UG1 and UG2. That is to say with instructional level, participants might not differ a lot on receptive and productive vocabulary measurements but their proficiency is much higher. The results also seem to show that advanced learners could be aiming for an average of 65 out of 120. More tests would be needed of course, but as with receptive vocabulary, mounting evidence is helpful to distinguish the levels students should aim at. It also seems that with instructional years, there is more homogeneity, that is to say, while UG1 and UG2 can have very different proficiency scores, the UG3 range tends to be smaller. However, other studies using EITs across institutional year groups would need to be carried out to refute or confirm this. There was a strong correlation between EIT scores and X-Lex scores showing that the two are associated. Again, this does not confirm previous studies such as Tracy-Ventura et al (2014) and further studies might be needed. On the other hand, the present study confirms the correlation between EI scores and end-of-year exam grades as per other studies (Tracy-Ventura, 2014).

The motivation questionnaire data showed that participants are very motivated by learning more about the language and by the prospect of becoming fluent. Instrumental-promotion is slightly lower than one would have anticipated though and the present study does confirm recent findings that seem to suggest that L2 learners of LOTEs do not particularly want to integrate in the L2 community. It is also clear that the Ought-to Self is not an important factor as it stands for L2 advanced French learners in the UK and in Chapter 6, we have suggested some ways of rethinking the Ought-to Self component, at least in the context of L2 learning in the UK.

Our goal to investigate the relationship between motivation and the development of vocabulary and proficiency yielded some interesting findings. First, the motivational factors of instrumentality-prevention and culture interest in particular seem to be linked with the development of larger vocabulary and higher proficiency whereas the factor of anxiety and

instrument-promotion were associated with smaller receptive vocabulary size and lower proficiency level. Therefore, not only can motivation explain to some extent vocabulary and proficiency development differences but also, vocabulary development and overall proficiency development depend both on similar motivational components, at least from the data collected in the present study.

L2 exposure is important for L2 vocabulary and proficiency development but the way it is reported by participants needs to be considered with caution. In some cases, independent study activities reported would most probably have little impact on participants' lexical or proficiency development because of the lack of active engagement. As far as L2 exposure is concerned, there are three main differences between low and high proficiency groups: frequency of exposure, participants' engagement with materials and relevance of materials. In addition to these three differences, it seems that, according to the data provided by these particular participants, reading is the activity that differentiates the most between low and high vocabulary and proficiency groups. That is to say, participants whose receptive vocabulary knowledge and proficiency level are higher, declare reading on a more regular basis compared to students whose lexical knowledge and proficiency level are lower.

This study on real language learners acquiring vocabulary over a long time-scale, does show that there is an imbalance in incidental activities undertaken by participants outside the classroom. The data show that meaning focused output (writing and speaking) is infrequent. As a consequence, the productive knowledge of lexical items might be impeded as there are few opportunities to expand the knowledge beyond making form-meaning connections. Having said that, let's not forget that all participants have written and spoken assessments as part as their French language module and as such might actually develop the syntactic and semantic meaning of lexical items more than it first appears from the L2 French exposure. What the data suggest though, is that independent planning and independent undertaking of meaning focused output activities is lagging behind meaning focused input (reading and listening) and so is fluency development. Practical issues, as well as Webb and Nation's definition of fluency development, can to some extent explain why these two strands of activities are less prioritised. We will suggest in the next sections ways of addressing these two strands.

7.3 Pedagogical implications

Based on the data collected and taking into account the profile of students and their learning environment, that is to say advanced students of L2 French in a British university, pedagogical implications in two major areas, namely incidental vocabulary and fluency development can be drawn. Before looking at these pedagogical implications, we need to note that the study also reveals that the L2 French pedagogical approach at this university is successful as students improve from year to year. All the data collected, from receptive vocabulary, productive vocabulary and proficiency tests, confirm this point. The Modern Languages and Linguistics department at this university has a solid reputation for the teaching of languages and regularly seeks and acts upon students' feedback. Teaching materials and activities also reflect current good practice and take into account students' needs and interests. Assessments do differentiate between learners and correlate well with other measures of proficiency such as EITs. Expanding vocabulary, grammar and proficiency knowledge as well as practising the four skills regularly during independent study time is strongly recommended and strategies, activities and feedback are provided to help achieve these goals. But as often in language learning though, individual differences lead to different learning outcomes.

Receptive vocabulary test results in November show that all groups have a 2K band deficit, that is to say, that their knowledge of the most frequent 2,000 words is not as strong as one would expect. The threshold of 2,000 words in language learning and teaching is recognised as very important because, the knowledge of these words facilitates all language learning activities, i.e. reading, writing, watching etc. Therefore it is in students' interest to develop and consolidate this knowledge. This could be achieved in different ways. For instance, a receptive vocabulary test could be made available to students, in particular to first years so that they can evaluate their lexical vocabulary. In addition, other resources that reinforced the knowledge of very frequent words should be specifically shown to students. Resources could also be created to practise both receptively and productively these words to develop syntactic and semantic knowledge that will facilitate incidental learning through authentic materials. With stronger knowledge of the 2,000 most frequent words, students are able to comprehend more contexts in which additional new words occur and therefore in turn can lower the learning burden of new words as well as consolidating the syntactic and semantic information of known words.

Detecting lower lexical vocabulary knowledge, planning and creating resources that front load these high frequency words can be pedagogically implemented and can have at least three

advantages. First, if weaker students make the most of this, it has the potential to improve their vocabulary learning. Secondly, these pedagogical approach also reinforces the message that vocabulary learning requires varied input and therefore that a lot of language learning at advanced level has to be undertaken outside the classroom. By explicitly explaining why decontextualized learning is not enough and by explicitly explaining why, to start with, a good command of the most frequent words is necessary (to make the most of the context), students are given further understanding and therefore control over their learning. This is in line with current efforts made to help students become efficient independent learners. Thirdly, this approach has also the potential to motivate some students further. Indeed, as we have seen, learning about the language itself is a strong motivation factor for them and a declarative knowledge of how to maximise learning can have a positive effect.

Another way to improve students' motivation when the teacher is not present or when they are not working with other learners is to remind them of relevant self-regulatory strategies, leading to self-motivation as defined by Ushioda (1997: 41):

> This capacity entails taking personal control of the affective conditions and experiences that shape one's subjective involvement in learning. It entails minimising the damage when these experiences are negative, and maximising the subjective rewards when these experiences are positive, and so fostering optimum motivational conditions for continued engagement in language learning.

In addition to self-regulatory strategies, to improve students' low linguistic self-confidence, teachers can encourage these students to evaluate their achievements in a positive light. Attributions of test results and other assessment outcomes to effort rather than ability will also make them realise that they can change, and spend more time to improve next time which is a very important factor in vocabulary and proficiency development. Also, L2 vocabulary learning for advanced university students becomes less deliberate and more incidental. There is plenty of anecdotal evidence that some students can find this shift difficult because they feel they are faced with learning thousands of new words in order to become "fluent". Comprehensive information about the necessity to revisit and to use new words could be shared too as well as strategies and examples for how to do it.

Spending time and effort learning outside the classroom is of course what teachers recommend students to do and the fact that this advice is not always followed is a well-known problem for language teachers. However, one possible way to improve motivation to dedicate more time and

effort can be to demonstrate and regularly emphasise how vocabulary gain means more ease in learning and therefore more ease in accessing a range of authentic resources. This could be a way of improving low linguistic self-confidence whilst supporting Ushioda and Dörnyei's (2017: 453) argument that "a less instrumentalist and more holistic view of language learning as a process of expanding people's meaning-making repertoires" should be favoured. In addition to reinforcing strategies and the message that incidental learning outside the classroom is key, Oakes (2013: 189) also suggests students should have more contact time, that is to say more language classes as well as the opportunity to have content modules (i.e. history, literature modules) delivered in French rather than in English. As incidental learning with the help of a teacher leads to more vocabulary learning, this suggestion has the possibility to improve all students' vocabulary but particularly that of low achievers, which in turn might give them more confidence and therefore more autonomy in their learning outside the classroom. On the other hand, because direct vocabulary teaching in class can be very time consuming and not always successful, the ultimate aim of more contact time should be to help learners become more independent.

Another pedagogical implication that can be drawn from this study is that meaning focused output activities and fluency development could be more present in the learning experience of these advanced students. One way of doing is so would be to use the tools available from a Virtual Learning Environment (VLE) such as Backboard to create such activities. As far as meaning focused output is concerned, writing and speaking tasks (such as Vlogs for the latter) could be set on weekly basis. Each students could complete them and record their writing and their speaking under their Blackboard profile. Language tutors could view them every few weeks and comment or provide short feedback, without marking or assessing these productions. These independent writing and speaking activities have potential for both learners and teachers. Learners of course maintain regular writing and speaking practice while teachers have a way of assessing students' commitment outside the classroom.

As we know many language development activities at this advanced level are independent activities but as there is no need to submit or to record work, it is not always easy for teachers to know how much work is actually done outside the classroom. Therefore, these activities will have the advantage to inform the teacher whilst still being in line with the role of the teacher as a facilitator. Finally, these types of independent meaning focused output activities have the potential to improve the linguistic self-confidence and to lower the anxiety of low achieving students. The data show that students whose vocabulary and proficiency lag behind tend to be more anxious and we took the view that this leads to less practice, therefore less feedback and

therefore ongoing low confidence about their abilities. Being able to practice in safe surroundings and to get feedback on the work produced might in that sense be helpful for these students.

Time restrictions could also be added to these online writing and speaking activities. For instance, instructions could specify that students should complete the tasks in 3 minutes .This way learners will be able to develop their fluency in the way described by Webb and Nation (2017). Other ways of implementing more meaning focused output and fluency development activities could be to implement extra-curricular activities or modules. An obvious one would be for instance a creative writing class that would provide weekly writing practice and would have the potential to motivate students who are already keen writers. But implementing such modules might not always be possible because of other considerations such as financial and strategic visions of universities.

7.4 Limitations of the present study and implications for further studies

The most obvious limitation of this study is the small sample size that does not allow to generalize findings and, at times, probably prevented results from reaching statistical significance. Another limitation is that not all 40 participants participated in all the tests. The duration of the study also somewhat limits the findings because, even though it spans over three years, it does not follow the same cohort over three years. Finally, instrument design could have been improved; for example, rather too many cognates were included in the receptive and productive vocabulary tests.

But despite its limitations, the present study demonstrates the existence of links between vocabulary learning, proficiency development, motivation and exposure to L2 French over an extended period of time. Further studies investigating the proficiency of L2 French advanced learners from different curricular groups will be needed to investigate whether the homogeneity in proficiency that was observed with institutional level is confirmed. The motivation of L2 French learners in British university also raises some interesting points such as the construction and the evolution of the instrument-promotion factor. It is one of the components that is often advertised for language learning but from the current findings, it seems to be associated with lower vocabulary and proficiency development. The present study also demonstrates that the internet plays an important role today in the development of vocabulary (online flashcards), access to L2

resources (watching, reading). When investigating the L2 exposure of advanced learners, further studies might focus on the types of internet activities undertaken and whether different technological approaches to language learning can also help explaining individual differences in vocabulary and proficiency.

In exploring the lexical development of advanced learners over a three year period, this study has addressed Webb and Nation (2017: 57)'s call for investigating the contribution of incidental learning over a long period of time. In using the French EIT version developed by Tracy-Ventura et al (2014), the study also contributes to the gathering of currently limited data on global L2 French oral proficiency from different curricular levels. Finally, in exploring to what extent the individual factor of motivation could explain learners' variation, the study also has also considered the question of the L2 Motivational Self system relevance for the study of languages other than English (LOTEs), raised by Ushioda and Dörnyei (2017).

Suggestions for further research include replication studies of the present study with larger numbers of students and with other languages (Spanish, German, Chinese etc). Running a similar study which would be longitudinal rather than cross sectional would also be a useful way to confirm the present study findings. The actual instruments used in the study could also be improved, with, for instance, a different approach to item selection for Flex15 in order to create a productive test that would be closer to the original Lex30 test.

A possible agenda for future research studies also includes the development of the Flex15 test as a productive test for L2 French vocabulary that would not only enable the tester to determine the frequency of responses and investigate word associations, but also to gather more information about the quality of the learner's word knowledge. One possible way of doing so would be to create a two-step Flex test. The first step would be very similar to the current test with stimulus words for which the learner needs to produce word associations. The answers could then be categorised per frequency band and following some criteria that would need to be defined to answer specific research questions, some of the learner's answers could be used in a second test dedicated to measuring the quality of the productive vocabulary. To do so, the four level scale mapping the Capture Zones of Productive Vocabulary developed by Fitzpatrick and Clenton (2017) could be used.

Another possible future study would be to investigate the learning outcomes of the least reported strand in vocabulary learning, that is to say the fluency development strand, described by Nation as the use of familiar materials as quickly as possible in listening, speaking, reading and writing.

Activities could be created following this definition and could be included in a vocabulary development program. Timed-activities and their results could be accessed through a Virtual Learning Environment and could also be implemented in the classroom. The potential impact of fluency activities in speaking and listening on productive vocabulary development and proficiency level would be interesting to investigate.

Another future study that would be interesting and potentially impactful in psycholinguistic and instructional frameworks would be to investigate the syntactic information that learners acquire when encountering a word for the first time. To determine words learners do not know, they could be included first within a receptive vocabulary test. All words that are not known could then be included in a reading task and be the focus of exercises manipulating these words to determine how much syntactic information has been processed from the overall input. This would be particularly interesting for researchers in the field of fundamental psycholinguistics.

To summarise, words are central to proficiency, as demonstrated with the correlation between receptive vocabulary and proficiency tests, and input drives learning, which leads receptive and productive vocabulary knowledge to be influenced by the frequency factor. However, as far as the development of rarer words and the quality of known words are concerned, this study confirms the view that intentional learning is needed to complement exposure to input. The present study has not only shed some light on the L2 lexical development of advanced learners in relation to proficiency and motivation but it also provides useful information about L2 vocabulary learning in the UK. Given that foreign language education in the UK system is in a weak state, and levels of achieved proficiency are low, including levels of vocabulary knowledge, it is important to understand the vocabulary knowledge and development amongst university students because they are among the most successful L2 learners in the system. Their vocabulary knowledge and what might speed or impede its development is of special interest as an indicator of quality in the wider educational system, i.e. it shows the most that even specialist students are likely to achieve under current UK conditions, and also provides some pointers to how this achievement might be improved.

Appendix A

French X-Lex Vocabulary test

Please look at these words. Some of these words are real French words and some are invented but are made to look like real words. Please tick the words that you know you can use. Here is an example.

Chien √

Conseil	joueur	deviner	Honteux	Lifrer
Peu	eltrisse	copain	Porc	Rigoler
Abjecter	blesser	éteindre	Fiable	Dépouille
Fort	dépenser	signard	Aviser	Malin
Près	Volet	souhait	Précont	Exode
Conseil	morceau	cibler	Sel	Expected
Partout	Pluie	dour	crâne	Orgueil
Viser	soigner	humeur	Jerette	Cage
Naître	défaulter	vendeur	Alourdir	Vendu
Blanc	Huile	odeur	Boursier	Oubli
Espoir	épargner	conjoint	Entrance	Traite
Peur	sonner	quai	Nouer	vulgaire
Salle	lancement	écourt	Fromage	crête
Crétale	couler	clef	Fleche	nadoir
Terrain	formirique	habituellement	Rayer	voisinage
Lien	Rater	tueur	équivaloir	écrasant
Mer	jaune	gestide	Néfaste	inonder
Attente	mensonge	deuxièmement	Fillette	forestier
arguable	Onde	char	Périr	viable
Niveau	disabilité	tante	Navette	ancrer
Clair	Aîné	tardif	Diroir	performant
Euplain	pente	adieu	Foudre	bosniaque
Parmi	Aile	impuissant	Détente	tirôt
Oeil	Lèvre	haïr	Cochon	allouer

(Adapted from Milton, J 2009)

Appendix B

French learner questionnaire

This survey is conducted to better understand the belief and thoughts of learners of French. This questionnaire is made up of two sections. Please read the instructions and write your answers. The results of this survey will be used for research purpose only so please give your answers sincerely. Thank you for your help!

Part I

In this part, we would like you to tell us how much you agree or disagree with the following statements by simply circling a number from 1 to 6. Please do not leave out any items.

Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agree
1	2	3	4	5	6

Ex If you strongly agree with the following statement, write this:

I like skiing	1	2	3	4	5	**6**

1 When my teachers give me an optional assignment, I certainly do it	1	2	3	4	5	6
2 I spend a lot of time and effort studying French	1	2	3	4	5	6
3 I think I am doing my best to learn French	1	2	3	4	5	6
4 I can imagine myself living abroad and having a discussion in French	1	2	3	4	5	6
5 Whenever I think of my future career, I imagine myself using French	1	2	3	4	5	6
6 I can imagine myself speaking French as if I were a native speaker	1	2	3	4	5	6
7 Studying French is important to me in order to gain the approval of my peers	1	2	3	4	5	6
8 Studying French is important to me in order to gain the approval of my parents	1	2	3	4	5	6
9 It will have a negative impact on my life if I don't learn French	1	2	3	4	5	6
10 Studying French is important to me in order to gain the approval of my tutors	1	2	3	4	5	6
11 My parents speak French well	1	2	3	4	5	6
12 Studying French is important to me because I think it will someday be useful in getting a job	1	2	3	4	5	6
13 Studying French is important to me because I would like to live abroad for a longer period of time	1	2	3	4	5	6

14 Studying French is important to me because it offers a new challenge in my life	1	2	3	4	5	6
15 I study hard for French otherwise I will not gain good grades	1	2	3	4	5	6
16 Studying French is important to me because I don't like to be considered a weaker student	1	2	3	4	5	6
17 I like the atmosphere of my French classes	1	2	3	4	5	6
18 I find learning French very interesting	1	2	3	4	5	6
19 I would like to have more French classes	1	2	3	4	5	6
20 I feel excited when I hear French spoken	1	2	3	4	5	6
21 I am interested in the way French is used in conversation	1	2	3	4	5	6
22 I find the difference between English and French vocabulary interesting	1	2	3	4	5	6
23 I like the rhythm of French	1	2	3	4	5	6
24 I get nervous when I speak in my French class	1	2	3	4	5	6
25 I worry about making mistakes when speaking French	1	2	3	4	5	6
26 Learning French is important in order to learn the culture of its speakers	1	2	3	4	5	6
27 I would like to become more similar to people who speak French	1	2	3	4	5	6
28 I consider myself to be good at French	1	2	3	4	5	6
29 I consider myself to be good at languages in general	1	2	3	4	5	6
30 French is my favourite language	1	2	3	4	5	6
31 I like accessing a variety of French authentic materials	1	2	3	4	5	6
32 Internet is my main source of French authentic materials outside classroom-related materials	1	2	3	4	5	6
33 I like studying content modules (e.g. about French culture, literature, history or linguistics)	1	2	3	4	5	6
34 I would like my French content modules to be taught in French	1	2	3	4	5	6
35 I like to travel to French-speaking countries	1	2	3	4	5	6
36 I like meeting people from French-speaking countries	1	2	3	4	5	6
37 I like to learn more about people from French-speaking countries	1	2	3	4	5	6

Part II

*Please write your responses in the spaces and delete when necessary **

Personal details

1. Date of birth: 2. Nationality:

3. Degree title and year of study:

Languages background

4. Are languages other than English spoken in your family home? YES/ NO*

If yes, which language(s):

If no, skip to question 9

5. How would you describe your spoken level in this/these language(s): Beginner/advanced/fluent*

6. Do you read and write in this/these language(s)? YES/NO*

If yes, what kind of materials do you read (eg newspapers) and how often?

7. Do you write in this/these language(s)? YES/NO*

If yes, describe your written level in this/these language(s): Beginner/advanced/fluent*

8. Have you taken a formal examination in this/these language(s) for eg GCSE? YES/NO*

If yes, what grade did you get?

French language learning background

9. How long have you been learning French as a foreign language?

10. What grade did you get at GCSE?

11. What grade did you get at A-level?

12. Have you ever visited France or a French speaking country? YES/NO*

If yes, give details about when you went and how long for:

13. Do you enjoy reading in English? YES/NO*

If yes, what kind of materials do you like to read (e.g. fiction) and how often do you read on average?

14. Rate your language skills in French from 1 to 4 (1 being weak and 4 being excellent)

Reading:	Listening:
Writing:	Speaking:
Grammar:	Vocabulary:

15. What activities do you enjoy in French language classes?

16. What activities do you not enjoy in French language classes?

17. According to you what makes a good language learner?

18. Why did you choose to study French at university?

19. Outside French language classes, and related coursework, what activities do you do to practise, consolidate or further your language knowledge and skills? List them all eg: studying grammar, studying vocabulary, tandem exchange with a native speaker, watching DVDs, chatting online to a native speaker, reading etc, detailing why you favour such or such activities and when/how often you practise French this way

20. What strategies do you make use of to build your French vocabulary?

21. Which other languages do you study at the university and what stage do you study them at?

This is the end of the questionnaire. Thank you for your cooperation!

Appendix C

Participant Information Sheet (Face to Face)

Study Title: Lexical development in advanced learners of French

Researcher: Virginie Pignot-Shahov **Ethics number**:

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

The aim of this research is to gather data concerning the vocabulary development of undergraduate students of French over a four-year degree period. In the last few years, vocabulary has regained an important place in the field of research and yet the uptake of vocabulary in advanced learners of French is not widely studied. With your help, this research aims at documenting this vocabulary development and the various factors that might affect it. As a Teaching Fellow of French, this study is part of my MPhil/PhD and is funded by the University of Southampton.

Why have I been chosen?

You have been chosen because your profile as an advanced learner of French is very relevant to the study.

What will happen to me if I take part?

You will complete two vocabulary tasks in November 2014 and in May 2015. The first task is a vocabulary recognition test. In other words, you will be presented with a number of French words and you will need to state whether you know these French words by ticking them. This task will take about 10 minutes. The second task will consist of writing four words that come to your mind when reading some trigger words. These two tasks will be repeated in May 2015 (note that the words in the tests will be different) and you will also be asked to complete a short questionnaire about your language learning which should take no more than 20 minutes to complete. No preparation is required for any of the tasks. In addition, there will be a proficiency test in November 2015 and a memory test in May 2015.

Are there any benefits in my taking part?

Yes there are. First of all, you get the opportunity to learn more about vocabulary as a research area which will also enable you to self-reflect on your vocabulary learning. You will also get the opportunity to practise your written skills and to obtain feedback on your vocabulary richness. Please note that the tests for this research are completely separate from your degree assessments and that your test results will not be discussed or shared with other teachers.

Are there any risks involved?

None beyond the normal risks of everyday life on campus.

Will my participation be confidential?

I will comply with the Data Protection Act/University policy and the information will be stored and remain confidential. The data will be anonymised before I analyse it, and will be kept on a password protected computer.

What happens if I change my mind?

You have the right to withdraw from this project at any time. No other aspect of your programme of study will be affected if you choose to withdraw from the research.

What happens if something goes wrong?

In the unlikely case of concern or complaint, please contact the Chair of the Faculty Ethics Committee Prof Chris Janaway (023 80593424, c.janaway@soton.ac.uk) or Research Governance (02380 595058, mad4@soton.ac.uk)

Where can I get more information?

If you have any questions, please do not hesitate to contact me directly (02380 593942, V.Pignot-Shahov@soton.ac.uk)

Appendix D

CONSENT FORM *(FACE TO FACE: Insert Version number)*

Study title:

Researcher name: Virginie Pignot-Shahov

Staff/Student number: 2220946

ERGO reference number: 7759

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (insert date /version no. of participant information sheet) and have had the opportunity to ask questions about the study.	☐
I agree to take part in this research project and agree for my data to be used for the purpose of this study	☐
I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected	☐

Data Protection

I understand that information collected about me during my participation in this study will be stored on a password protected computer and that this information will only be used for the purpose of this study. All files containing any personal data will be made anonymous.

Name of participant (print name)……

Signature of participant……..…..

Date……

Tips for Designing a Consent Form

(Please delete this section before submitting your documents to your Ethics Committee/RGO)

The main features of a good consent form are:

Date and Version number

It is important that the consent form is version numbered and dated so it is possible to track changes if and when they occur.

Use of Ethics reference

This is evidence of ethical approval and will reassure participants – enter the Submission ID generated when you create a submission in ERGO

Use of itemised statement to allow each component of the research to be agreed to

Information will be commensurate with the study. For example, in an interview study you may want consent to (i) interview and (ii) tape the interview.

Use of Initial boxes

Better than tick boxes to minimise fraud

Confirmation of the right to withdraw

You may wish to include a separate statement on confidentiality/anonymity but this is often best explained in the participant information sheet

Space for printed names, signatures and dates

A space for the name and signature of the person taking consent is also desirable if different from the named researcher

For studies involving minors/vulnerable adults

For studies involving minors/vulnerable adults, consent must normally be obtained from the parent/guardian/carer. It is also desirable for the participant to sign an assent form to indicate their willingness to take part.

There are situations where it may be appropriate to use 'opt-out consent' (informing parents/carers of the study and that if they do not respond to inform the researcher that

*they **do not** want their child/dependent to take part then it will be assumed that their consent is given). If you are in any doubt about the method of consent required you should seek advice from your local Ethics Committee or the Research Governance Office.*

When conducting research outside the UK, there may be different consent practices in use. In these cases you should take local advice and make sure that your consent requests reach all relevant stakeholders.

For studies conducted via the internet

For studies in which participants will take part online, (e.g. web surveys and questionnaires), the Participant Information Sheet and Consent Form are typically combined on one webpage. A clear Statement of Consent should be provided with which participants must agree in order to continue.

In most cases participants will need to be over 16 to participate, so they must also confirm this. Note that if the study is to be advertised on the web then you need to indicate on the IRGA Form on ERGO that portions of the research may take place outside of the UK.

Remind participants in internet based research that they may withdraw from the study at any time by closing the webpage. If the study may lead to any distress (e.g., reporting anxiety, recalling unpleasant events), provide one or more suitable "after-care" contact details on the information page in case a participant withdraws before reaching the debriefing page.

Be clear about whether or not internet data will be anonymous. If you are collecting any identifying information (e.g., email address), please state whether this information can be linked to data.

In summary, a combined PIS and Consent Form for internet based research should include the following paragraphs:

1. Opening statement
2. Description of research
3. *[Statement of risks + aftercare details] (only if necessary)*
4. Statement of Consent.

Example opening statement for internet based research:

Please read this information carefully before deciding whether to take part in this research. You will need to indicate that you have understood this information before you can continue. You must also be aged over 16 to participate. By ticking the box at the

bottom of this page and clicking 'Continue', you are indicating that you are aged over 16, and you are consenting to participate in this survey.

Example statement of risks with aftercare details, to be provided at the end of an online study:

We have tried to ensure that the questions in this study do not cause any distress. However, it is not uncommon to experience some anxieties or concerns when completing questionnaires about [*topic*], and support is available. If participating in this study raises any issues for you, we recommend that you contact one of the following resources: [*provide details*]

Example statement of consent (before participant clicks to continue):

Statement of Consent

I have read and understood the information about this study. In consenting, I understand that my legal rights are not affected. I also understand that data collected as part of this research will be kept confidential and that published results will maintain that confidentiality. I finally understand that if I have any questions about my rights as a participant in this research, or if I feel that I have been placed at risk, I may contact the Administrator of the Ethics Committee, Humanities, University of Southampton, SO17 1BJ, UK. Phone: +44 (0)23 8059 4663, Email: C.Cooke@soton.ac.uk.

I certify that I am 16 years or older. I have read the above consent form and I give consent to participate in the above described research.

Use of user-friendly language

Please use clear and simple English while preparing your consent form. Other languages may be used where appropriate.

Appendix E

French X-Lex Vocabulary test **Student number**

Please look at these words. Some of them are real French words and some are invented but are made to look like real words. Please tick the words that you know and can use.

sur	quotidien	haine	colline	expecter
dire	crier	signard	féliciter	tiède
ainsi	prémetance	ruche	marin	vaguement
pays	hiver	conseiller	dissoudre	éblouir
manchir	sien	quasi	aléatoire	ananas
connaître	abattir	souffle	précont	ras
beau	sinon	Fuir	vider	farouche
tête	témoin	plaindre	Las	nadoir
coeur	sable	Dour	chêne	d'emblée
crétale	pouce	soumis	jerette	ours
ajouter	paroi	écourt	entamer	sauvegarder
regard	serrer	coller	hâter	tirôt
environ	vernique	quatrième	restreint	souriant
lumière	échec	rayonnement	entrance	chagrin
lancer	veille	échantillon	brisé	entasser
abjecter	dépriver	gestide	couteau	lenteur
foi	pire	interdire	guérir	rangé
volonté	respirer	préalable	trottoir	rivage
faible	dent	lâcher	pourcent	lifrer
goût	hélas	poing	diroir	saut
métracte	appui	pêche	Uni	vers
hier	recueillir	écraser	colis	cendrier
Scolaire	envers	Sel	couché	cuisson
Maintenir	inscrire	doucement	gratuit	frôler

Appendix F

French X-Lex Vocabulary test **Student number**

Please look at these words. Some of them are real French words and some are invented but are made to look like real words. Please tick the words that you know and can use.

Ce	sang	Voile	bondir	serre
Falloir	oser	Lent	clabrer	bricoleur
Non	né	rideau	laine	octe
Puis	feuille	ultimation	rocher	grenier
Empâtrer	plaire	décor	souple	nouer
Comprendre	assis	goutte	bijou	apiculteur
Obtenir	tentre	vusier	plaie	rubir
Milieu	soulever	pente	culon	aveu
jeu	office	jadis	enfer	broder
Aige	taille	roue	barre	os
Bras	prier	plafond	éveiller	conte
Soleil	abeille	fleuve	pinceau	rente
Voisin	sureux	proine	gant	touceul
Mener	genou	aube	grappe	char
Auquel	nu	gras	fusil	gamin
Joyance	revue	brossin	giste	remise
Courant	néanmoins	abri	sueur	talon
Champ	plassard	miel	sirop	déferler
Cadre	aile	chevalier	boucle	rojette
Pentée	nolois	barbe	essor	écorce
Diriger	véhicule	muet	tirage	file
Âme	fier	taureau	matrer	induit
Janvier	grève	vitre	devise	gueule
Fer	tort	col	ivre	hasarder

masser	égarer
plaider	soupaire
voûte	farine
avribeur	logis
canton	pointu
coulisse	bail
épave	écourt
gâter	déprimer
boînase	sitôt
patin	allumette
ancre	tirôt
carrefour	arrondir
brigeable	brûlure
horloge	trifler
vibrer	civique
défier	corvée
fourmi	emprise
mât	entouré
seau	faiblir
évirelle	freiner
tiroir	hâtif
toux	inculpé
tracteur	indéfinement
araignée	gruyère

Appendix G

Flex15 vocabulary test **Student number**

You will need about 10 minutes to do this test. The test consists of 15 items.
For each item, write four words in French which you think are related to it.

E.g: **Animal** tigre zoo lion nature

1. **photo**……
2. **fête**……
3. **jardin**……
4. **espoir**……
5. **carte**……
6. **ciel**……
7. **crayon**……
8. **stupide**……
9. **ordinateur**……
10. **demander**……
11. **mémoire**……
12. **voiture**……
13. **vacances**……
14. **assiette**……
15. **futur**……

Appendix H

Flex15 vocabulary test **Student number**

You will need about 10 minutes to do this test. The test consists of 15 items.
For each item, write four words in French which you think are related to it.

E.g: **Animal** tigre zoo lion nature

1. **journal**……
2. **lundi**……
3. **mer**……
4. **savoir**……
5. **pluie**……
6. **oreille**……
7. **argent**……
8. **rapide**……
9. **loyal**……
10. **maison**……
11. **radio**……
12. **tristesse**……
13. **football**……
14. **pouvoir**……
15. **tradition**……

Appendix I

Elicited Imitation – Newly developed French stimuli, literal English translation, and parallel English stimuli from Ortega *et al.* (1999)

Syllable numbers for the English version in parentheses

1. *Je dois aller au coiffeur*. ('I have to go to the hairdresser's')
I have to get a haircut (7 syllables)
2. *Le livre rouge est sur la table.* ('The red book is on the table')
The red book is on the table (8 syllables)
3. *Les rues sont larges dans cette ville*. ('The streets are wide in this city')
The streets in this city are wide (8)
4. *Il prend une longue douche tous les matins*. ('He has a long shower every morning')
He takes a shower every morning (9)
5. *Qu'as-tu dit que tu vas faire aujourd'hui*? ('What did you say that you are going to do today?')
What did you say you were doing today? (10)
6. *Je ne sais pas s'il sait très bien conduire*. ('I don't know if he can drive all that well')
I doubt that he knows how to drive that well (10)
7. *Après le repas, j'ai fait une paisible sieste.* ('After dinner I had a peaceful nap')
After dinner I had a long, peaceful nap (11)
8. *Il est possible qu'il se mette à pleuvoir demain.* ('It is possible that it might rain tomorrow')
It is possible that it will rain tomorrow (12)
9. *J'adore les films, surtout ceux qui finissent bien*. ('I love films, especially those that have happy endings')
I enjoy movies which have a happy ending (12)
10. *Les maisons sont très jolies mais inaccessibles*. ('The houses are very nice, but inaccessible')
The houses are very nice but too expensive (12)
11. *Le petit garçon, dont le chaton est mort, est triste*. ('The little boy whose kitten died is sad')
The little boy whose kitten died yesterday is sad (13)
12. *Le restaurant est censé servir de très bons plats*. ('The restaurant is supposed to serve very good food')

That restaurant is supposed to have very good food (13)

13. *Je veux une maison où mes animaux peuvent habiter*. ('I want a home where my pets can live')

I want a nice, big house in which my animals can live (14)

14. *Tu aimes bien écouter de la musique country, n'est-ce pas*? ('You like listening to country music, don't you?')

You really enjoy listening to country music, don't you (14)

15. *Elle a fini de peindre les murs de son appartement.* ('She finished painting the walls in her flat')

She just finished painting the inside of her apartment (14)

16. *Traverse la rue au feu rouge et puis continue juste tout droit*. ('Cross the street at the red light and then just continue straight on')

Cross the street at the light and then just continue straight ahead (15)

17. *La personne que je vois a un sens de l'humour fabuleux*. ('The person I'm dating has a great sense of humour')

The person I'm dating has a wonderful sense of humour (15)

18. *Elle ne commande que de la viande et ne mange aucun légume*. ('She only orders meat and never eats vegetables')

She only orders meat dishes and never eats vegetables (15/16)

19. *J'aimerais que le prix des maisons de ville soit accessible.* ('I would like town houses to be more accessible')

I wish the price of town houses would become affordable (15)

20. *J'espère que le printemps arrivera plus tôt que l'an dernier*. ('I hope spring arrives sooner than last year')

I hope it will get warmer sooner this year than it did last year (16)

21. *Un de mes meilleurs amis s'occupe des enfants de mon voisin*. ('One of my best friends looks after my neighbour's children')

A good friend of mine always takes care of my neighbour's three children (16)

22. *Le chat noir que tu as nourri a été chassé par le chien*. ('The black cat that you fed was chased by the dog')

The black cat that you fed yesterday was the one chased by the dog (16)

23. *Avant de pouvoir sortir, il doit finir de ranger sa chambre*. ('Before going out, he has to finish tidying his room')

Before he can go outside, he has to finish cleaning his room (16)

24. *Je me suis bien amusé lors de notre sortie à l'opéra* ('I had a great time when we went to the opera')

The most fun I've ever had was when we went to the opera (16)

25. *Le voleur que la police a arrêté était très grand et mince*. ('The thief that the police arrested was very tall and thin')

The terrible thief whom the police caught was very tall and thin (17)

26. *Pourriez-vous s'il vous plaît me passer le livre qui est sur la table*? ('Would you please pass me the book that is on the table?')

Would you be so kind as to hand me the book which is on the table? (17)

27. *Le nombre de fumeurs de cigares augmente chaque année*. ('The number of cigar smokers goes up each year')

The number of people who smoke cigars is increasing every year (17/18)

28. *Je ne sais pas si le train de 11h30 a déjà quitté la gare*. ('I don't know if the 11.30 train has already left the station')

I don't know if the 11:30 train has left the station yet (18)

29. *L'examen n'était pas aussi difficile que ce que vous m'aviez dit.* ('The exam was not as difficult as you had said')

The exam wasn't nearly as difficult as you told me it would be (18)

30. *Il y a énormément d'individus qui ne mangent rien du tout le matin*. ('There are a lot of people who do not eat anything in the mornings')

There are a lot of people who don't eat anything at all in the morning (19)

Tracy-Ventura, N, McManus, K, Norris, J.M. and Ortega, L (2014) "Repeat as much as you can": Elicited imitation as a measure of oral proficiency in L2. In, Leclercq, Pascale, Edmonds, Amanda and Hilton, Heather (eds.) *Measuring L2 Proficiency: Perspectives from SLA.* Bristol, GB, Multilingual Matters, 143-166. (Second Language Acquisition, 78).

Appendix J

SCORING GUIDELINES FOR ELICITED IMITATION TASK

L2 FRENCH

SCORE 0

Criteria	**Examples**
• Nothing (Silence)	
• Garbled (unintelligible, usually transcribed as XXX)	
• Minimal repetition, then item abandoned: - Only 1 word repeated - Only 1 content word plus function word(s) - Only 1 content word plus function word(s) plus extraneous words that weren't in the original stimulus - Only function word(s) repeated **NOTE: with** only, just, yet (meaningful adverbs), score 1	- le voleur bla bla bla la commerce (25) - suive la rue jusqu'au feu rouge (16) - bla bla bla dans la ville (2)

SCORE 1

Criteria	Examples
• When only about half of idea units are represented in the string but a lot of important information in the original stimulus is left out	- je ne sais pas… conduire (6) - J'adore **la** film… (9) - quitter la gare (28)
• When barely half of lexical words get repeated and meaningful content results that is unrelated (or opposed) to stimulus, frequently with hesitation markers	- **XXX** dans ce ville - bla bla bla tu vas faire aujourd'hui (5) -le petit garçon dans le château est très triste (13)
• Or when string doesn't in itself constitute a self-standing sentence with some (targetlike or nontargetlike) meaning (This may happen more often with shorter items, where if only 2 of 3 content words are repeated and no grammatical relation between them is attempted, then score 1)	Je sais passe s'il a un très bien conduire
• Also when half of a long stimulus is left out, and the sentence produced is incomplete	

SCORE 2

Criteria	Examples
• When content of string preserves at least more than half of the idea units in the original stimulus; string in meaningful, and the meaning is close or related to original, but it departs from it in some slight changes in content, which makes content inexact, incomplete, or ambiguous	J'adore les films, surtout **les** qui finissent bien J'aimerais que **les** prix **de** maison **est plus** accessible **Les voleurs** étaient très grand et mince (25) Je ne sais pas si le train de **11h00** a déjà quitté la gare. **La maison** est très jolie mais **accessible** (12)
• Cases of extra marking or more marked morphology should be considered as meaning change. For example, a present tense repeated as past or as future should be scored as meaning change (score 2).	
• Similarly, singular/plural differences between stimulus and repeated string change the meaning, not only the grammar (score 2).	Je veux une maison ou **les** animaux **peut** habiter
• Changes of person (il for elle) change the meaning, so score 2	**La rue** sont larges dans **le** ville Il prend **les longues douches** tous les matins Avant de pouvoir sortir, **elle** doit finir de ranger sa chambre.

SCORE 3

Criteria	Examples
• Original, complete meaning is preserved as in the stimulus. Strings which are quite ungrammatical can get a 3 score, as long as exact meaning is preserved. Synonymous substitutions are acceptable. • Changes in grammar that don't affect meaning should be scored as 3. For instance, failure to supply past tense (had>have) and missing articles should be considered grammar change only (score 3). • Ambiguous changes in grammar that COULD be interpreted as meaning changes from a NS perspective should be scored as 2. That is, as a general principle in case of doubt about whether meaning has changed or not, score 2.	Pourriez-vous me passer le livre qui est sur la table? Je dois aller **à** coiffeur J'espère que le printemps arrivera plus tôt que **l'année dernière** J'espère que le printemps **arrive** plus tôt que l'an dernier

SCORE 4

Criteria	Examples
• Exact repetition: String matches stimulus exactly. Both form and meaning are correct without exception or doubt. • Missing liaison acceptable	Je **dois aller** au coiffeur (1)

Appendix K

French learner questionnaire

This survey is conducted to better understand the belief and thoughts of learners of French. This questionnaire is made up of two sections. Please read the instructions and write your answers. The results of this survey will be used for research purpose only so please give your answers sincerely. Thank you for your help!

Part I

In this part, we would like you to tell us how much you agree or disagree with the following statements by simply circling a number from 1 to 6. Please do not leave out any items.

Strongly disagree	Disagree	Slightly disagree	Slightly agree	Agree	Strongly agree
1	2	3	4	5	6

Ex If you strongly agree with the following statement, write this:

I like skiing	1	2	3	4	5	**6**

1 When my teachers give me an optional assignment, I certainly do it	1	2	3	4	5	6
2 I spend a lot of time and effort studying French	1	2	3	4	5	6
3 I think I am doing my best to learn French	1	2	3	4	5	6
4 I can imagine myself living abroad and having a discussion in French	1	2	3	4	5	6
5 Whenever I think of my future career, I imagine myself using French	1	2	3	4	5	6
6 I can imagine myself speaking French as if I were a native speaker	1	2	3	4	5	6
7 Studying French is important to me in order to gain the approval of my peers	1	2	3	4	5	6
8 Studying French is important to me in order to gain the approval of my parents	1	2	3	4	5	6
9 It will have a negative impact on my life if I don't learn French	1	2	3	4	5	6
10 Studying French is important to me in order to gain the approval of my tutors	1	2	3	4	5	6
11 My parents speak French well	1	2	3	4	5	6
12 Studying French is important to me because I think it will someday be useful in getting a job	1	2	3	4	5	6
13 Studying French is important to me because I would like to live abroad for a longer period of time	1	2	3	4	5	6
14 Studying French is important to me because it offers a new challenge in my life	1	2	3	4	5	6
15 I study hard for French otherwise I will not gain good grades	1	2	3	4	5	6
16 Studying French is important to me because I don't like to be considered a weaker student	1	2	3	4	5	6
17 I like the atmosphere of my French classes	1	2	3	4	5	6

18 I find learning French very interesting	**1**	**2**	**3**	**4**	**5**	**6**
19 I would like to have more French classes	**1**	**2**	**3**	**4**	**5**	**6**
20 I feel excited when I hear French spoken	**1**	**2**	**3**	**4**	**5**	**6**
21 I am interested in the way French is used in conversation	**1**	**2**	**3**	**4**	**5**	**6**
22 I find the difference between English and French vocabulary interesting	**1**	**2**	**3**	**4**	**5**	**6**
23 I like the rhythm of French	**1**	**2**	**3**	**4**	**5**	**6**
24 I get nervous when I speak in my French class	**1**	**2**	**3**	**4**	**5**	**6**
25 I worry about making mistakes when speaking French	**1**	**2**	**3**	**4**	**5**	**6**
26 Learning French is important in order to learn the culture of its speakers	**1**	**2**	**3**	**4**	**5**	**6**
27 I would like to become more similar to people who speak French	**1**	**2**	**3**	**4**	**5**	**6**
28 I consider myself to be good at French	**1**	**2**	**3**	**4**	**5**	**6**
29 I consider myself to be good at languages in general	**1**	**2**	**3**	**4**	**5**	**6**
30 French is my favourite language	**1**	**2**	**3**	**4**	**5**	**6**
31 I like accessing a variety of French authentic materials	**1**	**2**	**3**	**4**	**5**	**6**
32 Internet is my main source of French authentic materials outside classroom-related materials	**1**	**2**	**3**	**4**	**5**	**6**
33 I like studying content modules (e.g. about French culture, literature, history or linguistics)	**1**	**2**	**3**	**4**	**5**	**6**
34 I would like my French content modules to be taught in French	**1**	**2**	**3**	**4**	**5**	**6**
35 I like to travel to French-speaking countries	**1**	**2**	**3**	**4**	**5**	**6**
36 I like meeting people from French-speaking countries	**1**	**2**	**3**	**4**	**5**	**6**
37 I like to learn more about people from French-speaking countries	**1**	**2**	**3**	**4**	**5**	**6**

Appendix L

Language learning background questionnaire

*Please write your responses in the spaces and delete when necessary **

Personal details

1. Date of birth: 2. Nationality:

3. Degree title and year of study:

Languages background

4. Are languages other than English spoken in your family home? YES/ NO*

If yes, which language(s):

If no, skip to question 9

5. How would you describe your spoken level in this/these language(s): Beginner/advanced/fluent*

6. Do you read and write in this/these language(s)? YES/NO*

If yes, what kind of materials do you read (eg newspapers) and how often?

7. Do you write in this/these language(s)? YES/NO*

If yes, describe your written level in this/these language(s): Beginner/advanced/fluent*

8. Have you taken a formal examination in this/these language(s) for eg GCSE? YES/NO*

If yes, what grade did you get?

French language learning background

9. How long have you been learning French as a foreign language?

10. What grade did you get at GCSE?

11. What grade did you get at A-level?

12. Have you ever visited France or a French speaking country? YES/NO*

If yes, give details about when you went and how long for:

13. Do you enjoy reading in English? YES/NO*

If yes, what kind of materials do you like to read (e.g. fiction) and how often do you read on average?

14. Rate your language skills in French from 1 to 4 (1 being weak and 4 being excellent)

Reading:	Listening:
Writing:	Speaking:
Grammar:	Vocabulary:

15. What activities do you enjoy in French language classes?

16. What activities do you not enjoy in French language classes?

17. According to you what makes a good language learner?

18. Why did you choose to study French at university?

19. Outside French language classes, and related coursework, what activities do you do to practise, consolidate or further your language knowledge and skills? List them all eg: studying grammar, studying vocabulary, tandem exchange with a native speaker, watching DVDs, chatting online to a native speaker, reading etc, detailing why you favour such or such activities and when/how often you practise French this way

20. What strategies do you make use of to build your French vocabulary?

21. Which other languages do you study at the university and what stage do you study them at?

This is the end of the questionnaire. Thank you for your cooperation!

Appendix M

UG302

Sentence	Answer	Score-criteria
1		4
2	Le vero rouge est sur la table.	2 String preserves at least more than half of the idea units.
3	Les roues sont larges dans cette ville.	2 String preserves at least more than half of the idea units.
4	Il prend **un** longue douche tous les matins.	3 Ungrammatical but the original, complete meaning is preserved.
5		4
6	Je sais pas s'il sait très bien conduire.	3 Original, complete meaning is preserved.
7	Après le repas, j'ai fait une petite bleummmm sieste.	2 String preserves at least more than half of the idea units.
8		4
9	J'adore les films surtout tous ce qui finissent bien.	2 String preserves at least more than half of the idea units.
10		4
11	Le petit garçon dont le château est mort est triste.	2 String preserves at least more than half of the idea units.
12	Le restaurant sans arrêt un très bons plats.	1 String doesn't in itself constitute a self-standing sentence.
13	Je veux une maison où les animaux peuvent habiter.	2 String preserves at least more than half of the idea units.
14	Tu aimes bien écouter la musique country, n'est-ce pas ?	2 String preserves at least more than half of the idea units.
15	Elle a fini de peindre les murs dans son appartement.	2 String preserves at least more than half of the idea units.
16	Traverse la roue au feu rouge et puis continue tout droit.	2 String preserves at least more than half of the idea units.

17	La personne que je vois a un sens d'amour fabuleux.	2 String preserves at least more than half of the idea units.
18	Elle ne commande que de la viande et ne mange que des légumes.	2 String preserves at least more than half of the idea units.
19	J'aimerais que les prix des maisons en ville soit accessibles.	2 String preserves at least more than half of the idea units.
20	J'espère que le printemps arrivera plus tôt que le dernier.	2 String preserves at least more than half of the idea units.
21		4
22		4
23	Avant de pouvoir sortir, il doit finir de sortir sa chambre.	2 String preserves at least more than half of the idea units.
24	Je me suis bien amusé lors de la sortie à l'opéra.	2 String preserves at least more than half of the idea units.
25	Le voleur que la police a arrêté a été très grand et mince.	2 Grammatical change that affect meaning.
26	Pourriez svp passer le livre qui est sur le table ?	2 String preserves at least more than half of the idea units.
27		4
28	Je sais pas si le train de onze heures et demie est déjà quitté la gare.	2 String preserves at least more than half of the idea units.
29	L'examen est pas si difficile à ce que tu m'avais dit.	1 Only about half of idea units are represented.
30	Il y a énormément des individus qui ne mangent rien le matin.	2 String preserves at least more than half of the idea units.

Appendix N

L2 exposure data from L2 language learning questionnaire

	Meaning focused input (learning through reading and listening)	**Meaning-focused output (learning through speaking and writing)**	**Language-focused learning (deliberate attention to language features)**
UG101	1. Subscribing to email newsletters Le Monde « à la une » and about.com (2h-3h/week).	2. Chatting online with native speakers (1h/week). 1. Integration of vocabulary into sentences.	3. Studying grammar (1h-2h/week). 4. Proof-reading GCSE/A-level essays for friends. 2. Quizlet.com, online flashcards 3. Writing my own vocabulary lists on paper.
UG102	1. watching films (sometimes). 2. Watching news online on France 24 (sometimes). 3. Reading books (sometimes).		1. Extracting vocabulary from articles and books.
UG103	1. Reading newspapers and magazine from the LLC (once a week). 2. Listening to French radio when pottering around.	1. Writing summary sheets on topics studied in class to include key sentences and vocabulary.	2. Recording new vocabulary in a book. 3. Highlighting and annotating new vocabulary in articles I read.
UG104	1. Reading magazines (every 2 weeks). 2. Listening to French songs, have them in the background (every day)	3. Speaking with Erasmus students (every week)	1. Making a list of new words, memorize them and test myself 2. Quizlet, online flashcards
UG105	1. Watching DVDs (once a week)		1. Quizlet, online flashcards

UG106	1. Watching vlogs	2. Emailing French friends (regularly)	3. Going over grammar points and tenses studied in class (regularly) 1. Quizlet, online flashcards (once a week) 2. Make a list of new words
UG107	1. Reading French books (1/year) 2. Reading French news (once or twice/week) 3. Watching films (once/week)	4. Talking to French friends (once every two weeks)	1. Make a list of new words
UG108	1.Watching DVDs (one a term) 2.Reading the news (once or twice/week)	3.Chatting online to native speakers (once every ¾ weeks)	4.Writing grammar notes (once every 2 weeks) 1. Make lists of new words from classes 2. Using synonym dictionary to find alternatives to common words
UG110	1.Reading books 2.Watching films	3.Speaking French at "café parler" 1.Writing	2.Noting key phrases 3.Repetition
UG111	1.Watching films (a few times/month) 1.Reading	2.Speaking to my Erasmus friend (once/week)	2.Looking up unknown words
UG201	1.Reading parallel books and news articles 2.Watching DVDs 1.Reading books to see vocabulary in context		2. Anki flashcard app 3. Making my own flashcards and lists
UG202	1.Listening to French/Francophone music (weekly) 2.Reading blogs, newspaper articles online (a few times/week)	5.Chatting with native speakers (weekly/fortnightly)	6.Studying grammar 2. Looking up words in dictionary and record them if I feel they are relevant/interesting

	3.Reading French books (once every few months) 4.Watching French films (1-2/month) 1. Reading		
- UG203	1.Reading (frequently). 2.Listening to radio (in the background) when I work 3.Listening to (French) music (all the time). 4.Watching DVDs (very infrequently). 1.Reading.	5.Speaking at "Café parler" (once/week). 6.Messaging French native speakers (infrequently).	2.Repetition/recital and memory testing.
UG205	1.Watching films. 2.Listening to music.		1.Looking up unknown words in the dictionary.
UG206	1.Watching YouTube videos and French vloggers. 2.Reading French newspapers online and books. 1.Reading.	3.Talking online or in person to native speakers.	2. Looking up unknown words in the dictionary and make lists.
UG207	1.Reading France 24 news with app	2.Conversing (in informal French) with French Erasmus buddy. 3.Speaking French at "café parler" for one hour (every week)	4.Looking up grammar points that interest me. 1.Recording new words and trying to learn the list. 2.Creating puns out of words to remember them better: "amener is for people because it has the word men in it, not apporter".
UG208	1.Listening to French songs because it is a		4.Testing myself on verbs.

	good way to learn vocabulary. 1.Reading online articles (sometimes). 2.Watching Youtube videos (sometimes). 3.Watching French films with English subtitles (occasionally). 2.Reading articles to build my vocabulary.		3.Making lists of words I am unsure of in class.
UG209	1.Watching films (1-2/month) 2.Listening to music while completing work (most days) 3.Reading books (1/2 months) 4.Reading online newspapers (every other day) 5.Listening to the radio (every other day)		6.Studying grammar (once or twice a week) 1.Making vocabulary lists, check them, memorise them. 2.Making flashcards and learn them. 3.Revising frequent words with Duolingo.
UG210	1.Listening to French radio (2-3 times a week) 2.Listening to Michel Thomas teach yourself French (3 or 4 times a week) 3.Watching French films (once/month) 4.Looking at French websites (once a week or so)	6.Emailing my former A-level teacher who is a native French.	7.Studying grammar (once a week) 2.Writing words I don't know from set work.

	5.Reading one novel for my content module. 1.Reading an article or something online once a week and write down any vocabulary		
UG211	1.Watching Youtube videos 2.Skimming through the news	3.Speaking to French friends	1. Quizlet, online flashcards
UG212	1.Reading French 2.Watching French films (regularly) 1.Reading French newspapers online for vocabulary 2.Listening to French radio for vocabulary	3.Emailing French friends (every few weeks)	4.Looking back over grammar done in class to consolidate 3.Translating what is around me in French and thinking through in French 4.Discussing vocabulary with my dad who speaks French
UG301	1.Reading news articles 2.Watching films and TV programs (sometimes) 1.Reading		2.Making my lists of words to learn 3.Imagining my own sentences with the new vocabulary 4.Searching for examples of new words use on “webcorp.org”
UG302	1.Watching the odd French film (occasionally)	2.Speaking with native speakers on tandem exchanges	1.Learning vocabulary with AnkiDroid app
UG303	1.Reading books, newspapers and fictions 1.Reading	2.Writing an explanation in French of words I don’t know.	2.Writing down words I don’t know when I read with an explanation in French
UG304	1.Watching DVDs (twice/month)	2.Speaking with a native speaker on Skype (every week)	1.Writing down new vocabulary (when speaking and reading) and looking it up

UG306	1.Watching films, TV game shows, series (once/week) 2.Reading news (most days) 3.Listening to music (once/week)	4.Speaking to French speakers 1.Making an effort to insert new vocabulary into speech and assignments	2.Making lists of new vocabulary
UG307	1.Listening to French music 2.Reading newspapers online 1.Reading	3.Chatting with French natives (from time to time)	2.Looking up words when reading Le Monde or books for pleasure
UG308	1.Reading news online 2.Watching videos online	3.Speaking at "café parler"	1.Keeping vocabulary and idiom lists 2.Building up a vocabulary wall in my house
UG309	1.Watching news items (every day) 2.Reading magazine articles from the LLC 3.Reading a few pages of a novel (every day)	4.Chatting in French (once a week)	1.Creating flashcards with new vocabulary from class 2.Creating flashcards online and sharing them with peers 3.Writing down new vocabulary
UG310	1.Listening to music and radio 2.Watching films and series 3.Reading news headlines 4.Reading French versions of books I have read in English (e.g. Hunger games)	5.Chatting to French friends made during the year abroad and with my tandem partner	1.Making up vocabulary tests for myself and my friends 6.Going over specific grammar points that need to be worked on 2.Writing down new vocabulary when speaking and reading
UG311	1.Reading news (2-3 times/week)	2.Chatting with native speakers via Skype (once/week)	1.Creating quizlet online flashcards

UG312	1.Watching DVDs (a few times a month) 2.Reading French blogs online (weekly) 1.Reading and watching to see vocabulary in context	3.Speaking French with my tandem partner (4 times/month)	4.Studying grammar 2.Studying vocabulary 3.Making flashcards with new words learned over the semester
UG313	1.Listening to the French news (every day) 2.Watching French films/series (often) 1.Listening		2.Looking up new words when listening to the radio
UG314	1.Reading articles about politics and current affairs (every week) 2.Reading French literature (once every two weeks)	3.Speaking to native speakers via Skype (once/month)	1.Looking up new words and making lists in relevant category: verbs, nouns etc
UG315	1.Reading French novels (a book every two months) 1.Reading 2.Listening to French radio (2-3 times/week)	3.Speaking with my tandem partner (once/week)	

List of references

Adolphs, S. and Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics* 24, (4), pp. 425-438.

Aitchison, J. (2003). *Words in the mind: An introduction to the mental lexicon*. Oxford: Blackwell.

Aizawa, K. (2006). Rethinking frequency markers for English-Japanese dictionaries. In M. Murata., K. Minamide., Y. Tono, and S. Ishikawa, ed., *English Lexicography in Japan*. Tokyo: Taishukan-shoten, pp. 108-119.

Alderson, C. and Hudson, R. (2013). The metalinguistic knowledge of undergraduate students of English language or linguistics. *Language Awareness*, 22, pp. 320-337.

Al-Homoud, F. and Schmitt, N. (2009). Extensive reading in a challenging environment: A comparison of extensive and intensive reading approaches in Saudi Arabia. *Language Teaching Research*, 13, pp. 383-401.

American Council on the Teaching of Foreign Languages. (2012). Oral Proficiency Interview. Familiarization manual. NY: ACTFL professional programs

Anderson, J.R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press

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

Anderson R.C. and Freebody, P. (1983). Reading comprehension and the assessment and acquisition of word knowledge. In B. Hutson, ed., *Advances in reading/language research: A research annual*. Greenwich: CT: JAI Press, pp. 231-256.

Anderson, R.C. and Shifrin, Z. (1980). *The meaning of words in context*, in Spiro et al., pp. 330-348.

Andringa, S., de Glopper, K. and Hacquebord, H. (2011). Effect of explicit and implicit instruction on free written response task performance. *Language Learning*, 61, pp. 868-903.

Ashton, C. (2003). *Key stage 3 National Strategy. Framework for teaching modern foreign languages: years 7, 8 and 9.* Retrieved from http://webarchive.nationalarchives.gov.uk/20130401151715/http://www.education.gov.uk/publications/eOrderingDownload/0289-2003PDF-EN.pdf [Last accessed on 08/06/2016]

Bachman, L.F. and Palmer, A.S. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly*, 16, pp. 449-465.

Bachman, L.F. and Palmer, A.S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.

Barcroft, J. (2009). Effects of synonymy generation on incidental and intentional L2 vocabulary learning during reading. *TESOL Quarterly*, 43 (1), pp. 79-103.

Barfield, A. and Brown, S.H. (2007). *Reconstructing Autonomy in Language Education. Inquiry and Innovation.* Basingstoke: Palgrave Macmillan.

Baudot, J. (1992). *Fréquences d'utilisation des mots en français écrit contemporain*. Montréal: Les Presses Universitaires de Montréal.

Benson, P. (2007). Autonomy in language teaching and learning. "State of the art" article. *Language Teaching,* 40, pp. 21-40.

Benson, P. (2010). *Teaching and Researching Autonomy in Language Learning*. 2nd ed. London: Longman.

Bernaus, M. and Gardner, R. C. (2008).Teacher Motivation Strategies, Student Perceptions, Student Motivation, and English Achievement. *Modern Language Journal*, 92, pp. 387–401.

Board, K. and Tinsley, T. (2015). *Language Trends. The state of language learning in primary and secondary schools in England*. Reading: Education Development Trust.

Bowles, M. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute? *Studies in Second Language Acquisition*, 33, pp. 247-271.

Brown, D. (2012). The frequency Model of Vocabulary Learning and Japanese Learners. Vocabulary Learning and Instruction, 1 (1), pp. 20-28

Burger, S. and Chrétien, M. (2001). The development of oral production in content-based second language courses at the University of Ottawa. *The Canadian Modern Language Review*, 58, pp. 84-102.

Busse, V. (2011). Language Use and Language Learning in CLIL Classrooms. *System*, 39 (4), pp. 559–561.

Busse, V. (*2013*). An exploration of motivation and self-beliefs of first year students of German, *System* 41, pp. 379–398.

Busse, V. and Walter, C. (2013). Foreign Language Learning Motivation in Higher Education: A Longitudinal Study of Motivational Changes and Their Causes. *Modern Language Journal*, 97(2), pp. 435-456.

Busse, V. and Williams, M. (2010). Why German? Motivation of students studying German at English universities. *Language Learning Journal*, 38 (1), pp. 67-85.

Byrnes, H. (2006). The outcomes of collegiate FL programs: Specifications, assessment, evaluation. *Modern Language Journal*, 90, pp. 574-601.

Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research*, 6 (2), pp. 145-173.

Canale, M. and Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. Applied Linguistics, 1, pp. 1-47.

Canning, J. (2011). *A survey of Non-specialist Language Learners in UK Higher Education.* Southampton: University Council for Modern Languages and the Subject Centre for Languages, Linguistics and Area Studies.

Carlsen, C. (2012). Proficiency level-a fuzzy variable in computer learner corpora. *Applied Linguistics*, 33, pp. 161-183.

Carroll, J.B. (1968). The psychology of language testing. In A. Davies, ed., *Language Testing Symposium: A Psycholinguistic Approach*. London: Oxford University, pp. 46-69.

Carroll, J.B. (1981). Consciousness and automatic processes in language learning. *Canadian Modern Language Review*, 37, pp. 462-474.

Chalhoub-Deville, M. (1997). Theoretical models, assessment frameworks and test construction. *Language Testing*, 14, pp. 3-22.

Chambers, G.N. (1999). *Motivating Language Learners*. Clevedon: Multilingual Matters.

Chapelle, C.A. (1998). Construct definition and validity inquiry in SLA research. In L.F. Bachman and A. Cohen, ed., *Interfaces between Second Language Acquisition and Language Testing Research*. Cambridge: Cambridge University Press, pp. 32-70.

Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.

Cheng, H. and Dörnyei, Z. (2007). The use of motivational strategies in language instruction: The case of EFL teaching in Taiwan. *Innovation in Language Learning and Teaching*, 1, pp. 153-174.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press.

Christensen, C., Hendrickson, R. and Londsdale, D. (2010). Principled construction of elicited imitation tests. Paper presented at the *Language Resources and Evaluation Conference*, Malta.

Csizér, K. and Dörneyi, Z. (2005). The internal structure of language learning motivation and its relation with language choice and learning effort. *Modern Language Journal*, 89, pp. 19-36.

Coady, J and Huckin, T. (1997). Second Language Vocabulary Acquisition. A Rationale for Pedagogy. Cambridge: Cambridge University Press

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

Coleman, J.A. (2009). Why the British do not learn languages: Myths and motivation in the United Kingdom. *Language Learning Journal*, 37, pp. 111-127.

Council of Europe. (2014). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR).* Retrieved from http://www.coe.int/t/dg4/linguistic/cadre1_en.asp [Last accessed on 08/06/2016]

Clenton, J. (2008). *Investigating the construct of the Productive Vocabulary: comparing different measures.* BAAL Annual Conference, pp. 26-28.

Cramer, P. (1968). *Word association*. New York: Academic Press.

Crookes, G. and Schmidt, R. (1991). Motivation: Reopening the research agenda. *Language Learning*, 41, pp. 469-512.

Csizér, K. and Dörnyei, Z. (2005). Language learners' motivational profiles and their motivated learning behaviour. *Language Learning*, 55 (4), pp. 613-659.

Csizér, K. and Kormos, J. (2009). Learning experiences, selves and motivated learning behaviour: A comparative analysis of structural models for Hungarian secondary and university learners of English. In Z. Dörnyei, and E. Ushioda, ed., *Motivation, Language Identity and the L2 Self*. Bristol: Multilingual Matters, pp. 98-119.

Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire.* Clevedon: Multilingual Matters.

Daller, H., Milton, J. and Treffers-Daller, J. (2007). *Modelling and assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press.

David, A. (2008). Vocabulary breadth in French L2 learners. *Language Learning Journal*, 36 (2), pp. 167-180.

David, A., Myles, F., Rogers, V. and Rule, S. (2009). Lexical development in instructed L2 learners of French: Is there a relationship with morphosyntactic development? in H. Daller., D. Malvern., P. Meara., J. Milton., B. Richards, and J. Treffers-Daller, ed., *Vocabulary studies in first and second language acquisition: the interface between theory and application*. Basingstoke: Palgrave, pp. 147-163.

De Bot, K. (1992). A Bilingual production model: Levelt's "Speaking" Model Adapted. Applied Linguistics, 13, pp 1-25

De Bot, K. (2004). The Multilingual Lexicon: Modeling Selection and Control. International Journal of Multilinguism, 1 (1), pp. 17-32

De Bot, K., Paribakht, T., and Wesche, M. (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), pp. 309-329.

De Bot, K., Lowie, W. and Verspoor, M. (2005). *Second Language Acquisition. An advanced resource book*. London: Routledge Applied linguistics.

Deese, J. (1965). *The structure of associations in language and thought*. Baltimore, MD: The Johns Hopkins Press.

De Groot, A.M.B. and Kroll, J.F. (1997). *Tutorials in Bilingualism: Psycholinguistics Perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.

Department for Education (2013). *National curriculum in England: languages programs of study*. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-languages-progammes-of-study [Last accessed on 08/06/2016]

Department for Education and Skills (2002). *Languages for All: Languages for Life*. Retrieved from http://arts.ulster.ac.uk/nils/admin/documents/docs/DfESLanguagesStrategy.pdf [Last accessed on 08/06/2016]

DeKeyser, R. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: Cambridge University Press.

Dörnyei, Z. (1994a). Motivation and motivating in the foreign language classroom. *Modern Language Journal*, 78, pp. 273-284.

Dörnyei, Z. (1994b). Understanding L2 motivation: On with the challenge! *Modern Language Journa*l, 78, pp. 515-523.

Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. (2009). The L2 Motivational Self system. In Z. Dörnyei, and E. Ushioda, ed., *Motivation, Language Identity and the L2 Self.* Bristol: Multilingual Matters. pp. 9-42.

Dörnyei, Z. and Al-Hoorie, A.H. (2017). The motivational foundation of learning languages other than Global English: Theoretical issues and research directions. *Modern Language Journal*, 101(3), pp. 455–468.

Dörnyei, Z. and Csizér, K. (1998). Ten Commandments for motivating language learners: Results of an empirical study. *Language Teaching Research*, 2, pp. 203-229.

Dörnyei, Z. and Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research*, 4, pp. 275-300.

Dörnyei, Z. Taguchi, T. (2010). *Questionnaires in Second Language Research: Construction, Administration and Processing*. 2nd eds. London: Routledge.

Dörnyei, Z. and Ushioda, E. (2011). *Teaching and researching motivation*. 2nd ed. Harlow: Pearson.

Dunkel, C. S. and Anthis, K. S. (2001). The role of possible selves in identity formation: A short-term longitudinal study. Journal of Adolescence, 24(6), pp. 765-776.

Eisenstein, M., Bailey, N. and Madden, C. (1982). It takes two: contrasting tasks and contrasting structures. *TESOL Quarterly*, 16 (3), pp. 381-393.

Ellis, N.C. (1998). Emergentism, Connectionism and Language Learning. *Language Learning*, 48 (4), pp. 631-664.

Ellis, R. (1997). Second Language Acquisition. *Issues in Applied Linguistics*, 8 (1), pp. 147.

Ellis, R. (1999). *Learning a second language through interaction*. Amsterdam: John Benjamins.

Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning*, 51 (1), pp. 1-46.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, 54, pp. 227-275.

Ellis, N.C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. Studies in Second Language Acquisition, 27, pp. 305-352.

Ellis, R. (2008). The study of second language acquisition. Oxford: Oxford University Press.

Entwisle, D.R., Forsyth, D.F. and Muuss, R. (1964). The syntactic-paradigmatic shift in children's word associations. *Journal of Verbal Learning and Verbal Behavior*, 3, pp. 19-29.

Epstein, A.G. (1978). Sprogmangler hos børn. Copenhagen: Dansk: Videnskabs Forlag

Erlam, R. (2006). Elicited Imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics,* 27, pp. 464-491.

Erlam, R. and Loewen, S. (2010). Implicit and explicit recasts in L2 oral French interaction. *The Canadian Modern Language Review*, 66, pp. 887-916.

Erler, L. and Macaro, E. (2012). Decoding Ability in French as a Foreign Language and Language Learning Motivation. *Modern Language Journal*, 95 (4), pp. 496-518.

Ervin, S.M. (1961). Changes with age in the verbal determinants of word association. *American Journal of Psychology*, 2, pp 190-202.

European Commission. (2013). *Barcelona Objectives. The development of childcare facilities for young children in Europe with a view to sustainable and inclusive growth*. Luxemburg: Publications Office of the European Union.

Fitzpatrick, T. and Barfield, A. (2009). L*exical processing in second language learners: Papers and Perspectives in honour of Paul Meara*. Cleveland: Multilingual Matters.

Fitzpatrick, T. and Clenton, J. (2010). The challenge of validation: Assessing the performance of a test of productive vocabulary. *Language Testing*, 27 (4), pp. 537-554.

Fitzpatrick, T. and Clenton, J. (2017). Making Sense of Learner Performance on Tests of Productive Vocabulary Knowledge. *TESOL Quarterly, 51 (4), pp. 844-867.*

Gaillard, S. and Tremblay, A. (2016). Linguistic Proficiency and Assessment in Second Language Acquisition Research: The Elicited Imitation Task. *Language Learning*, 66 (2), pp. 419-447.

Gallagher-Brett, A. (2006). *Hard-going but worth it: A snapshot of attitudes to reading among languages undergraduates.* Southampton: LLAS Subject Centre.

Gallimore, R. and Tharp, R.G. (1981). The interpretation of elicited sentence imitation in a standardized context. *Language Learning*, 31 (1), pp. 369-392.

García-Amaya, L. (2009). *New Findings on Fluency Measurements across three Different Learning Contexts. Selected Proceedings of the 11th Hispanic Symposium*. MA: Cascadilla Proceeding Project, pp 68-80.

Gardner, R.C. (1985). *Social Psychology and Second Language Learning: The role of attitudes and motivation*. London: Baltimore, MD: E. Arnold.

Gardner, R.C. and Lambert, W.E. (1959). Motivational variables in second language acquisition. *Canadian Journal of Psychology*, 13, pp. 266-272.

Gardner, R.C., Masgoret, A-M., Tennant, J. and Mihic, L. (2004). Integrative motivation: Changes during a year-long intermediate-level language course. *Language Learning*, 54, pp. 1-34.

Gass, S. (1999). Discussion: Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21 (2), pp. 319-333.

Gatbonton, E. and Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 3, pp. 325-353.

Goulden, R., Nation, P. and Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11 (4), pp. 341-363.

Graham, C.R., McGhee, J. and Millard, B. (2010). The role of lexical choice in elicited imitation item difficulty. In M.T. Prior., Y. Watanabe, and Lee, S.K, ed, *Selected proceedings of the 2008 Second Language research forum: Exploring SLA perspectives, positions, and practices*. Somerville, MA: Cascadilla Proceedings Project, pp. 57-72.

Graham, S. (1997). *Effective Language Learning. Positive Strategies for Advanced Level Language Learning*. Clevedon: Multilingual Matters.

Graham, S. (2004). Giving up on modern foreign languages? Students' perceptions of learning French. *Modern Language Journal*, 88 (2), pp. 171-191.

Graham, S., Richards, B. and Malvern, D. (2008). Progress in learning French vocabulary in a one-year advanced course at school. *First Language Studies*, 18, pp. 349-364.

Green, D. W. (1986). Control, activation and resource: A framework and a model for the control of speech in bilinguals. Brain and Language, 27, pp. 210-223

Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, 1 (2), pp. 67-81.

Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Multilingual and Multilingual Development*, 6, pp. 467-477.

Gu, L. (2014). At the interface between language testing and second language acquisition: Language ability and context of learning. *Language Testing*, 31, pp. 111-133.

Gutiérrez, X. (2013). Metalinguistic knowledge, metalingual knowledge and proficiency in L2 Spanish. *Language Awareness*, 22 (2), pp 176-191.

Häcker, M. (2008). Eleven pets and twenty ways to express one's opinion: The vocabulary learners of German acquire at English secondary school. *Language Learning Journal*, 36 (2), pp. 215-226.

Hafner, L.E. (1967). Using context to determine meanings in high school and college. *Journal of Reading*, 10, pp. 491-498.

Hameyer, K. (1980). Testing oral proficiency via elicited imitation. Revue de Phonetique Appliquée, 53 (1), pp. 11-24.

Harley, B., Cummins, J., Swain, M. and Allen, P. (1990). The nature of language proficiency. In B. Harley, P. Allen, J. Cummins and M. Swain, ed., *The Development of Second Language Proficiency*. Cambridge: Cambridge University Press, pp. 7-38.

Hendrickson, R., Aitken, M., McGhee, J. and Johnson, A. (2010). What makes an item difficult? A syntactic, lexical, and morphological study of elicited imitation test items. In M.T. Prior., Y. Watanabe, and S.K. Lee, ed., *Selected proceedings of the 2008 Second Language research forum: Exploring SLA perspectives, positions, and practices*. Somerville, MA: Cascadilla Proceedings Project, pp. 48-56.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21(2), pp. 303-317.

Higgins, J. (1977). The THNEED test. *Pasaa* 7 (1), pp. 363-367.

*Hilton***,** *H***.** (2008). The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal,* 36 (2), pp. 153-166.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review/La revue canadienne des langues vivantes*, 61, pp. 355-382.

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

Horst, M. and Collins, L. (2006). From *Faible* to Strong: How Does Their Vocabulary Grow?. The Canadian Modern Language Review, 63 (1), pp. 83-106

Housen, A. and Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30, pp. 461-473.

Hudson, R. (1984). Invitation to Linguistics. London: Martin Robinson

Hulstijn, J. H. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, 18, pp. 193-223.

Hulstijn, J.H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal, 91*, pp. 662-666.

Hulstijn, J. H. (2010). Measuring second language proficiency. In E. Blom and S. Unsworth, ed., *Experimental methods in language acquisition research* (EMLAR). Amsterdam: John Benjamins. pp. 185-199.

Hulstijn, J.H. (2011). Language proficiency in language native and non-native speakers: an agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly*, 8, pp. 229-249.

Hulstijn, J.H. (2012). The construct of language proficiency in the study of bilinguism from a cognitive perspective. *Bilinguism: Language and cognition*, 15, pp. 422-433.

Hulstijn, J.H. (2013). Is the second language acquisition discipline disintegrating? *Language Teaching*, 46, pp. 511-517.

Hulstijn, J. H. (2015). *Language Proficiency in Native and Non-native Speakers. Theory and research*. Amsterdam: John Benjamins Publishing Company.

Hu, M. and Nation, P. (2000). Unknown Vocabulary Density and Reading Comprehension. *Reading in a Foreign Language*, 13 (1), pp. 403-430.

Huckin, T. and Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition: A review* 21 (1), pp. 181-193.

Hymes, D. (1972). On communicative competence. In J.B Pride and J. Holmes, ed., *Sociolinguistics*, Harmondsworth, UK: Penguin Books, pp. 269-293.

Inbar, O., Donista-Schmidt, S. and Shohamy, E. (2001). Students' motivation as a function of language learning. The teaching of Arabic in Israel. In Z. Dörnyei, and R. Schmidt, ed., *Motivation and Second Language Acquisition*. Honolulu, HI: University of Hawaii Press, pp. 297-311.

Jeon, E.H. and Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning*, 64, pp. 160-212.

Juffs, A. (2009). Second language acquisition of the lexicon in W.C. Ritchie, and T.K. Bathia, ed., *The New Handbook of Second language acquisition.* Bingley: Emerald Group Publishing Limited.

Kamimoto, T. (2005). *The effect of guessing on vocabulary test scores: A qualitative analysis.* Paper presented at the European Second Language Association (EuroSLA) Conference 15, Dubrovnik, Croatia.

Kaplan, R. and Bresnan, J. (1982). *Lexical-Functional Grammar: A Formal system for Grammatical Representation*. Stanford University. Stanford.

Kelch, K. (1985). Modified input as an aid to comprehension. *Studies in Second Language Acquisition*, 7 (1), pp. 81-90.

Kempen, G. and Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. Cognitive Science, 11(2), pp. 201-258.

Kim, Y., Tracy-Ventura, N. and Jung, Y. (2016). A measure of Proficiency or short-Term Memory? Validation of an Elicited Imitation Test for SLA Research. *Modern Language Journal*, 100 (3), pp. 655-673.

Kučera, H. and Francis, W.N. (1967). C*omputational analysis of present-day American English*. Providence: Brown University Press.

Kunnan, A.J. (1998). An introduction to structural equation modelling for language assessment research. *Language Testing,* 15, pp. 295-332.

Krashen, S. D. (1981). The "fundamental pedagogical principle" in second language teaching. S*tudia Linguistica*, 35 (1-2), pp. 50-70.

Krashen, S.D. (1985). *The input hypothesis: issues and implications*. London: Longman.

Krashen, S.D. (2003). *Explorations in language acquisition and use: The Taipei lectures.* Portsmouth, NH: Heinemann.

Kroll, J.F and Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. Journal of Memory and Language, 33, pp. 149-174

Kroll, J.F. and Sunderman, G. (2003). The development of lexical and conceptual representations. In C. Doughty and M. Long, ed., *The Handbook of Second Language Acquisition. Oxford: Blackwell*, pp. 104-129.

Kruse, H., Pankhurst, J., and Sharwood Smith, M. (1987). A multiple word association probe in second language acquisition research. *Studies in Second Language Acquisition*, 9 (2), pp 141-154.

Lado, R. (1961). *Language Testing*. New York, NY: McGraw-Hill.

Lakoff, G. (1972). Hedges: a study in meaning criteria and the logic of fuzzy concepts. Papers of the Eight Regional Meeting, Chicago Linguistic Society, pp. 183-228.

Lakoff, G. (1999). Cognitive Models and Prototype Theory. In E. Margolis, and E. Laurence, ed., *Concepts*, Cambridge Massachusetts: The MITT Press.

Lamb, M. (2004). Integrative motivation in a globalizing world. *System*, 32, pp. 3-19.

Lanvers, U. (2016). Lots of Selves, some rebellious: Developing the Self Discrepancy Model for Language Learners. *System*, 60, pp. 79-92.

Lanvers, U. (2017). Contradictory *Others* and the *Habitus* of Languages: Surveying the L2 Motivation Landscape in the United Kingdom. *Modern Language Journal*, 101 (3), pp. 517-532.

Laufer, B. (1997). What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In N.Schmitt and M. McCarthy, ed.,*Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.

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

Laufer, B. (2003). Vocabulary acquisition in a second language: do learners really acquire most vocabulary by reading? *Canadian Modern Language Review*, 59 (4), pp. 565-585.

Laufer, B. (2005). Focus on Form in second language vocabulary acquisition. In S. H. Foster-Cohen., M. P. Garcia-Mayo, and J. Cenoz*, ed., EUROSLA Yearbook 5*, pp. 223-250.

Laufer, B., Elder, C., Hill, K. & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, 21(2), pp. 202-226.

Laufer, B. and Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Journal of Learning Language*, 54(3), pp. 399-436.

Laufer, B. and Hulstijn, J. H. (2001). Incidental vocabulary acquisition and in a second language: the construct of tasked-induced involvement load. *Applied Linguistics*, 22, pp. 1-26.

Laufer, B. and Nation P. (1995). Vocabulary Size and Use: Lexical Richness in L2 written production. *Applied Linguistics,* 16, pp. 307-322.

Laufer, B. and Nation P. (2012). Vocabulary. In S.M. Gass, and A. Mackey, ed., *The Routledge Handbook of Second Language Acquisition*. London: Routledge, pp. 163-176.

Laufer, B. and Paribakht, T.S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48, pp. 365-391.

Laufer, B. and Ravenhorst-Kalovski, G.C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. Reading in a Foreign Language, 22 (1), pp. 15-30

Laufer, B. and Rozovski-Roitblat, B. (2011). Incidental vocabulary acquisition: The effects of task type, word occurrence and their combination. *Language Teaching Research*, 15, pp. 391-411.

Leclercq, P., Edmonds, A. and Hilton, H. (2014). *Measuring L2 Proficiency Perspectives from SLA*. Bristol: Multilingual Matters.

Leeman, J. and King, K.A. (2015). Heritage language education: Minority language speakers, second language instruction, and monolingual schooling. In M. Bigelow, and J. Ennse-Kananen, ed., *The Routledge Handbook of Educational Linguistics*. New York: Routledge/Taylor and Francis, pp. 210-223.

Lehrer, A and Kittay, E. F. (1992). Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization. New York: Routledge

Levelt, W. J. M. (1989). Speaking: From intention to articulation. Cambridge, MA, US: The MIT Press.

Levelt. W.J.M (1993). Lexical Access in Speech Production. Cambridge. Blackwell

Li, S. (2013). The interactions between the effects of implicit and explicit feedback and individual differences in language analytic ability and working memory. *Modern Language Journal*, 97, pp. 634-654.

Little, D. (1991). *Learner Autonomy 1: Definitions, Issues and Problems*. Dublin: Authentik.

Liu, N., and Nation, I.S.P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16 (1), pp 33-42.

Li Wei, L. (2006). The Multilingual Mental Lexicon and Lemma Transfer in Third Language Learning. International Journal of Multilinguism, 3 (2), pp.88-104

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27, pp. 361-386.

Loewen, S. (2015). *Introduction to instructed second language acquisition*. New York: Routledge.
Lonsdale, D. and Le Bras, Y. (2009). *A Frequency Dictionary of French: Core Vocabulary for Learners*. Oxon: Routledge.

Macaro, E. and Masterman, L. (2006). Does intensive explicit grammar instruction make all the difference? *Language Teaching Research*, 10, pp. 297-327.

Macaro, E. (2008). The decline in language learning in England: Getting the facts right and getting real. *Language Learning Journal*, 36, pp. 101-108.

MacIntyre, P. D. and Gardner, T. (1989). Anxiety and Second-Language Learning: Toward a Theoretical Clarification. *Language Learning,* 39 (2), pp. 251-275.

MacIntyre, P. D. and Gregersen, T. (2012). Affect: The role of language anxiety and other emotions in language learning. In S. Mercer, S. Ryan, and M. Williams. Ed., L*anguage learning psychology: Research, theory and pedagogy*. Basingstoke: Palgrave, pp. 103-118.

MacIntyre, P. D., Noel, K.A. and Clément, R. (1997). Biases in Self-Ratings of Second Language Proficiency: The Role of Anxiety. Language Learning, 47 (2), pp. 265-287.

Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21, pp. 557-587.

Markman, B., Spilka, I. and Tucker, G. (1975). The use of elicited imitation in search of an interim French grammar. *Language Learning*, 75, pp. 31-41.

Markus, H. and Nurius, P. (1986). Possible selves. *American Psychologist*, 41, pp. 954-969.

Masgoret, A.M. and Gardner, R. C. (2003). Attitudes, Motivation, and Second Language Learning: A Meta–Analysis of Studies Conducted by Gardner and Associates. *Language Learning*, 53, pp. 123–163.

Marsden, E. and David, E (2008). Vocabulary use during conversation: A cross-sectional study of development from year 9 to year 13 amongst learners of Spanish and French. L*anguage Learning Journal*, 36 (2), pp. 181-198.

McDade, H.L., Simpson, M.A. and Lamb, D.E. (1982). The use of elicited imitation as a measure of expressive grammar: a question of validity. *Journal of Speech and Hearing Disorders*, 47 (1), pp. 19-24.

Meara, P. (1996a). The dimensions of lexical competence. In G. Brown, K. Malmkjær, and J. Williams, ed., *Competence and performance in language learning*. Cambridge: Cambridge University Press, pp. 35-53.

Meara, P. (1996b). The classical research in vocabulary. In Anderman, G. and Rogers, M. (eds), Words, Words, Words. Clevedon: Multilingual Matters.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In N. Schmitt, and M. Mc Carthy, ed., *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press, pp. 109-121.

Meara, P. (2005). Lexical Frequency Profiles: A Monte Carlo Analysis. *Applied Linguistics*, 26 (1), pp. 32-47.

Meara, P. (2009). *Connected Words: word associations and second language lexical acquisition*. Amsterdam: John Benjamins.

Meara, P. (2014). *Lognostics. Tools for vocabulary research*. Retrieved from http://www.lognostics.co.uk/ [last accessed on 23/07/14]

Meara, P. and Buxton, B. (1987) An alternative to multiple choice will vocabulary testing. *Language Testing,* 4 (2), pp. 142-154.

Meara, P. and Fitzpatrick, T. (2000). Lex30: An improved method for assessing productive vocabulary in an L2. *System,* 28, pp. 19-30.

Meara, P. and Jones, G. (1988). *Vocabulary size as a placement indicator*. Eurocentres. Swansea University

Meara, P and Milton, J (2003). *X_Lex, the Swansea Levels Test*. Newbury: Express

Melka Teichroew, F.J. (1982). Receptive vs. productive vocabulary: a survey. *Interlanguage Studies Bulletin* 6 (2), pp. 5-33.

Měchura, M. B. (2017). Introducing Lexonomy: an open-source dictionary writing and publishing system in *Electronic Lexicography in the 21st Century: Lexicography from Scratch. Proceedings of the eLex 2017 conference, 19-21 September 2017, Leiden, The Netherlands.*

Milton, J. (2004). Comparing the lexical difficulty of French reading comprehension exam tests. Language Learning Journal, 30, pp. 5-11

Milton, J. (2006a). Language lite: Learning French vocabulary in School. *Journal of French Language Studies* 16 (2), pp. 187-205.

Milton, J. (2006b). French as a foreign language and the Common European Framework of Reference for Languages. In *Proceedings from the Crossing Frontiers: Languages and the International Dimension conference*, Cardiff University 6-7 July, distributed by CILT, the National Centre for Languages and the Subject Centre for Language, Linguistics and Area Studies.

Milton, J. (2008). Vocabulary uptake from informal learning tasks. *Language Learning Journal*, 36 (2), pp. 227-237.

Milton, J. (2009). *Measuring Second Language Vocabulary acquisition*. Bristol: SLA.

Milton, J. and Alexiou, T. (2009). Vocabulary Size and the Common European Framework of Reference for Languages. In B, Richards et al, ed., *Vocabulary Studies in first and second language acquisition*. Basingstoke: Palgrave, pp. 194-211.

Milton, J. and Fitzpatrick, T. (2014). Dimensions of vocabulary knowledge. Basingstoke: Palgrave Macmillan

Milton, J. and Hopkin, N. (2006). Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners. *Canadian Modern Language Review*, 63, pp. 127–147.

Milton, J. and Meara, P. (1998). Are the British really bad at learning foreign languages? *Language Learning Journal,* 18, pp. 68-76.

Ministère de l'éducation (1966) *Le Français fondamental*. Paris: Institut Pédagogique National.

Mitchell, R. (2003). Rethinking the concept of progression in the National Curriculum for Modern Foreign Languages: a research perspective. *Language Learning Journal*, 27 (1), pp. 15-23.

Mitchell, R. (2010). Policy and practice in foreign language education: case studies in three European settings. *European Journal of Language Policy*, 2, (2), pp. 151-180.

Mitchell, R. (2011). Still gardening in a gale: policy, research and practice in foreign language education in England. *Fremdsprachen Lehren und Lernen*, 40, (1), pp. 49-67.

Mitchell R, Myles, F and Marsden, E. (2013) *Second Language Learning Theories*. London: Routledge.

Munnich, E., Flynn, S. and Martohardjono, G. (1994). Elicited imitation and grammaticality judgement tasks; what they measure and how they relate to each other. In E.E Tarone, S. Gass and A.D. Cohen, ed., Research methodology in second-language acquisition. Hove: Lawrence Erlbaum, pp. 227-243.

Nagy, W.E., Herman, P. A., and Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20 (2), pp 233-253.

Naiman, N. (1974). The use of elicited imitation in second language acquisition research. *Working Papers on Bilingualism*, 2, pp. 1-37.

Namei, S (2004). Bilingual lexical development: A Persian-Swedish word association study. *International Journal of Applied Linguistics*, 14 (3), pp. 363-388.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Boston: Heinle & Heinle.

Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: University Cambridge Press.

Nation, I.S.P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63 (1), pp. 59-82.

Nation, I.S.P. (2007). The four strands. *Innovation in Language Learning and Teaching,* 1 (1), pp. 1-12

Nation, I.S.P. (2010). Learning Vocabulary in Another Language. Cambridge: University Cambridge Press.

Nation, I.S.P and Laufer, B. (2012). Vocabulary. S. Gass and A. Mackey, ed.,*The Routledge Handbook of Second Language Acquisition*. New York: Routledge, pp. 163-176.

Nation, I.S.P. and Webb, S.A. (2011). Researching and analysing vocabulary. Boston, MA: Heinle Cengage Learning

Neuman, S.B., and Koskinen, P.S. (1992). Captioned television as comprehensible input: effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, 27 (1), pp. 95-106.

New B., Pallier C., Ferrand L. and Matos R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE 3. *L'Année Psychologique,* 101, pp. 447-462. http://www.lexique.org/ [Accessed 14 April 2016]

Nissen, H.B. and Henriksen, B. (2006). Word class influence on word association test results. *International Journal of Applied Linguistics*, 16 (3), pp. 389-408.

Norton, B. (2000). *Identity and Language Learning: Gender, Ethnicity and Educational Change*. Harrow: Longman.

Norris, J.M. and Ortega, L. (2012). Assessing learner knowledge. In S.M. Gass, and A. Mackey, ed., *The Routledge Handbook of Second Language Acquisition*. New York, NY: Routledge, pp. 573-589.

Oakes, L. (2013). Foreign language learning in a "monoglot culture": Motivational variables amongst students of French and Spanish at an English university". *System*, 41, pp. 178-191.

Okura, E. and Lonsdale, D. (2012). Working memory's meagre involvement in sentence repetition tests. In N. Miyake, D. Peebles, and R.P. Cooper, ed., *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society, pp 2132-2137.

Oller, J.W., Jr. (1979). Language Tests at School: A Pragmatic Approach. London: Longman

Oller, J.W., Jr. (1983). Issues in Language Testing Research. Rowley, MA: Newbury House

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition,* 21 (1), pp. 109-148.

Ortega, L., Iwashita, N., Norris, J.M. and Rabie, S. (2002). An investigation of elicited imitation tasks in crosslinguistic SLA research. Paper presented at the Second Language Research Forum, Toronto.

Oxford, R.L and Shearin, J. (1994). Language learning motivation: expanding the theoretical framework. *Modern Language Journal*, 78, pp 12-28.

Oxford, R.L and Shearin, J. (1996). Language learning motivation in a new key. In R.L. Oxford, ed., *Language Learning Motivation: Pathways to the New Century*. Honolulu, HI: University of Hawaii Press, pp. 155-187.

Palmer, H.E. (1917). The Scientific Study and Teaching of Languages. London: Harrap

Paradis, M. (1987). The Assessment of Bilingual Aphasia. Hillsdale, NJ: Erlbaum

Pauwels, P. (2012). Vocabulary materials and study strategies at advanced level. *Language Learning Journal*, 40 (1), pp. 47-63.

Pemberton, R., Toogood, S. and Barfield, A. (2009). *Maintaining Control. Autonomy and Language Learning*. Hong Kong: Hong Kong University Press.

Pérez-Vidal, C. (2014). *Language Acquisition in Study Abroad and Formal Instruction Contexts*. University of Pompeu Fabra: John Benjamins Publishing Company.

Piaget, J. (1970). *Science of education and the psychology of the child*. New York: Viking.

Pienemann, M. (1998). *Language Processing and Second Language Development: Processability Theory*. Amsterdam: John Benjamins Publishing.

Pigada, M. and Schmitt, N. (2006). Vocabulary acquisition form extensive reading: a case study. *Reading in a Foreign Language*, 18 (1), pp. 1-28.

Prutting, C.A., Gallagher, T.M. and Mulac, A. (1975). The expressive portion of the NSST compared to a spontaneous language sample. *Journal of Speech and Hearing Disorders*, 40 (1), pp. 40-48.

Purpura, J. E. (2008). Assessing communicative language ability: Models and their components. In E. Shohamy and N.H Hornberger, ed., *Encyclopedia of language and education*.2nd ed. Language testing and assessment. Dordrecht: Kluwer, pp. 53-68.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing* 10 (3), pp. 355-371.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Rebuschat, P. and Williams, J. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, 33, pp. 829-856.

Rice, M.L., and Woodsmall, L. (1988). Lessons from television: children's word learning when viewing. *Child Development*, 59 (2), pp 420-429.

Richards, B.J. and Malvern, D.D. (2007). Validity and threats to the validity of vocabulary measurement. In H. Daller, J. Milton, and J. Treffers-Daller, ed., *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press, pp. 79-92.

Richards, B.J., Malvern, D. and Graham, S. (2008). Word frequency and trends in the development of French vocabulary in lower-intermediate students during Year 12 in English schools. *Language Learning Journal*, 36 (2), pp 199-213.

Savignon, S.J. (1982). Dictation as a measure of communicative competence in French as a second language. *Language Learning*, 32 (1), pp. 33-51.

Scheibner-Herig, G., Sauerbrey, H. and Kokoschka, D. (1991). Repetition- a means to predict foreign language oral proficiency. *International Review of Applied Linguistics*, 29 (3), pp. 229-240.

Schmid, S and Lowie, W. (2011). Modeling Bilingualism: From Structure to Chaos. In Honor of Kees de Bot. Amsterdam: John Benjamins

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, pp. 129-158.

Schmidt, R. and Watanabe, Y. (2001). Motivation, Strategy use pedagogical preferences in foreign language learning. In Z. Dörnyei and R. Schmidt, ed., *Motivation and Second Language Acquisition*. Honolulu. National Foreign Language Resource Centre, pp. 313-359.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt, and M. McCarthy, ed., *Vocabulary: Description, Acquisition, and Pedagogy*. Cambridge: Cambridge University Press.

Schmitt, N. (2000). Vocabulary in Language Teaching. Cambridge: Cambridge University Press

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual.* Basingstoke: Palgrave.

Schmitt, N., Schmitt, D. and Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18 (1), pp. 55-88.

Shönefeld, D. (2001). *Where Lexicon and Syntax Meet*. Berlin: Mouton de Gruyter.

Simon, H.A. (1974). How big is a chunk? *Science*, 183, pp. 482-488.

Singleton, D. (1999). *Exploring the Second Language Mental Lexicon*. Cambridge: Cambridge University Press.
Skehan, P. (1989). *Individual differences in Second-language Learning*. London: Arnold.

Slobin, D.I and Welsh, C.A. (1968). Elicited imitation as a research tool in developmental psycholinguistics. Working Paper 10. Reprinted in Ferguson, C.A and Slobin, C.A (1973). *Studies of child language development*. New York: Holt, Rinehart and Winston, Inc., pp. 485-497.

Söderman, T. (1993). Word associations of foreign language learners and native speakers: the phenomenon of a shift in response type and its relevance for lexical development. In H. Ringbom, ed., *Near-Native proficiency in English*, Åbo: Åbo Akademi: English Department Publications.

Sonaiya, R. (1991). Vocabulary acquisition as a process of continuous lexical disambiguation. *IRAL*, 29, pp 273-284

Sonbul, S. and Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning,* 63, pp. 121-159.

Spitze, K. and Fischer, S.D. (1981). Short-term memory as a test of language proficiency. TESL talk, *Quarterly for Teachers of English as a Second Language*, 12 (4), pp. 32-41.

Stæhr, L.S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal* 36 (2), pp. 139-152.

Stolz, W.S and Tiffany, J. (1972). The production of “child-like” word associations by adults to unfamiliar adjectives. *Journal of Verbal Learning and Learning Behavior*, 11, pp. 38-46.

Taguchi, T., Magid, M. and Papi, M. (2009). The L2 Motivational Self System among Japanese, Chinese and Iranian learners of English: a comparative study. In Z. Dörnyei, and E. Ushioda, eds, *Motivation, Language Identity and the L2 Self*. Bristol: Multilingual Matters, pp. 66-97.

Tanaka, K. and Ellis, R. (2003). Study abroad, Language proficiency, and Learner’s Belief about Language Learning. *JALT Journal*, 25 (1), pp. 63-85.

Tarone, E. (2013). Requiring a proficiency level as a requirement for U.S. K-12 teacher licensure. *Modern Language Journal*, 97, pp. 528-560.

Tian, L. and Macaro, E. (2012). Comparing the effect of teacher codeswitching with English-only explanations on the vocabulary acquisition of Chinese university students: A lexical focus-on-form study. *Language Teaching Research*, 16, pp. 367-391.

Tidball, F. and Treffers-Daller, J. (2008) Analysing lexical richness in French learner language: what frequency lists and teacher judgements can tell us about basic and advanced words. *Journal of French Language Studies*, 18-3, pp. 299-313.

Thomas, M. (1994). Assessment of L2 proficiency in second language acquisition research. *Language learning*, 44, pp. 307-336.

Thomas, M. (2006). Research synthesis and historiography: The case of assessment of second language proficiency. In J. Norris and L. Ortega, ed., *Synthesizing Research on Language Learning and Teaching*. Amsterdam: Benjamins, pp.279-298.

Thompson, A.S. and Vásquez, C. (2015). Exploring motivational profiles through language learning narratives. *Modern Language Journal*, 99, pp. 158-174.

Thorndike, E. and Lorge, I. (1944). The teacher's word book of 30,000 words. New York: Columbia University.

Tokowicz, N. (2015). *Lexical processing and second language acquisition*. New York: Routledge

Tracy-Ventura, N., McManus, K., Norris, J.M. and Ortega, L. (2014). "Repeat as much as you can": Elicited imitation as a measure of oral proficiency in L2. In, P. Leclercq, A. Edmonds, and H. Hilton, ed., Measuring L2 Proficiency: Perspectives from SLA. Bristol: Multilingual Matters, pp. 143-166.

Tremblay, A. (2011). Proficiency assessment standards in second language acquisition research. Clozing the gap. *Studies in Second Language Acquisition*, 33, pp. 339-372.

Tschichold, C. (2007). On the usefulness of word frequency information. Paper to BAAL, September 2007

Tseng, W-T., Dörnyei, Z. and Schmitt, N. (2006). A new approach to assessing strategic learning: the case of self-regulation in vocabulary acquisition. *Applied Linguistics,* 27 (1), pp. 78-102.

Tseng, W-T. and Schmitt, N. (2008). Toward a model of motivated vocabulary learning: A structural equation modelling approach. *Language Learning,* 58 (2), pp. 357-400.

University of Southampton (2016) Modern Language. Retrieved from http://www.southampton.ac.uk/ml/index.page [Last accessed 08/06/2016]

Ushioda, E. (1997). The role of motivational thinking in autonomous language learning. In Little, d. and Voss, B. (eds) Language Centres: Planning for the New Millennium. Plymouth: CERCLES, Centre for Modern Languages, University of Plymouth, pp. 39-50

Ushioda, E. (1998). Effective motivational thinking: A cognitive theoretical approach to the study of language learning motivation. In E.A. Soler. and V.C. Espurz, ed., *Current Issues in English Language Methodology*. Castelló de la Plana: Universitat Jaume I, pp. 77-89.

Ushioda, E. (2001). Language learning at university: Exploring the role of motivational thinking. In Z. Dörnyei, and R. Schmidt, ed., *Motivation and Second Language Acquisition*. Honolulu, HI: University of Hawaii Press, pp. 93-125.

Ushioda, E. and Dörnyei, Z. (2017). Beyond Global English: Motivation to Learn Languages in a Mutlicultural World: Introduction to the Special Issue. *Modern Language Journal*, 101, (3), pp. 451-454.

VanPatten, B. and Williams, J. (2007). *Theories in second language acquisition: An introduction.* Mahwah: Lawrence Erlbaum Associates.

Van Zeeland, H. and Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: the same or different from reading comprehension? *Applied Linguistics*, 34 (4), pp. 457-479.

Vermeer, A. (2004). The relationship between lexical richness and vocabulary size in Dutch L1 and L2 children. In P. Bogaards and B. Laufer (eds) *Vocabulary in a Second Language*. pp. 173-189. Amsterdam: John Benjamins

Verspoor, M.H., Lowie, W.M. and de Bot, C.L.J. (2008) "Input and Second Language Development from a Dynamic Perspective" In T.Piske and M,Young-Scholten (eds) *Input Matters in SLA.* Bristol: Multilingual Matters.

Vidal, K. (2003). Academic listening: a source of vocabulary acquisition? *Applied Linguistics,* 24 (1), pp.56-89.

Vinther, T. (2002). Elicited imitation: a brief overview. *International Journal of Applied Linguistics*, 12 (1), pp. 54–73.

Webb, S. (2007). The effects of synonymy on second language vocabulary learning. *Reading in a Foreign Language*, 19 (2), pp. 120-136.

Webb, S. (2008). Receptive and productive vocabulary size. *Studies in Second Language Acquisition*, 30, pp. 79-95.

Webb, S. and Chang, A.C.-S. (2015). How does prior word knowledge affect vocabulary learning progress in an extensive reading program? *Studies in Second Language Acquisition*, 37 (4), pp. 651-675.

Webb, S. and Nation, P. (2017*). How Vocabulary is Learned*. Oxford: Oxford University Press.

Webb, S. and Paribakht, T.S. (2015). What is the relationship between the lexical profile of test items and performance on a standardized English proficiency test? *English for Specific Purposes*, 38 (1), pp. 34-43.

Webb, S., and Rodgers, M.P.H. (2009). Vocabulary demands of television programs. *Language Learning*, 59 (2), pp. 335-366.

Weinreich, U. (1953). Languages in contact, findings and problems. New York: Linguistic Circle of New York

West, M. (1953). *A General Service List of English Words*. London: Longman, Green and Co.

Wilks, C. and Meara, P. (2002) Untangling word webs: graph theory and the notion of density in second language word association networks. *Second Language Research*, 18, pp. 303–324.

Wilks, C., Meara, P. and Wolter, B. (2005). A further note on simulating word association behaviour in an L2. *Second Language Research*, 21 (4), pp. 359-372.

Williams, M. and Burden, R.L. (1997). *Psychology for Language Teachers: A social constructivist approach*. Cambridge: Cambridge University Press.

Wolter, B. (2001). Comparing the L1 and the L2 mental lexicon; a depth of individual word knowledge model. *Studies in Second Language Acquisition,* 23, pp. 41-69.

Wood Bowden, H. (2016). Assessing second-language oral proficiency for research. *Studies in Second Language Acquisition*, 38, pp. 647-675.

Worton, M. (2009). Review of Modern Foreign Languages provision in higher education in England. HEFCE. Bristol. Retrieved from http://webarchive.nationalarchives.gov.uk/20100202100434/http://www.hefce.ac.uk/pubs/hefce/2009/09_41/ [Last accessed 08/06/2016]

Wu, S-L. and Ortega, L. (2013). Measuring Global Oral Proficiency in SLA Research: A New Elicited Imitation Test of L2 Chinese. *Foreign Language Annals*, 26 (4), pp. 680-704.

Yan, X., Maeda, Y., Jing, Lv. and Ginther, A. (2016). Elicited imitation as a measure of second language proficiency: A narrative review and meta-analysis. *Language Testing*, 33 (4), pp. 497-528.

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *Modern Language Journal*, 86, pp. 54-66.

Young, D. (1991). The relationship between anxiety and foreign language oral proficiency ratings. In E.K. Horwitz and D.J Young ed., *Language anxiety*. Englewood cliffs, NJ: Prentice-Hall, pp. 57-63

Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to native speakers' lexical organisation? *Second Language Research*, 23 (2), pp. 123-153.