

**THE INFLUENCE OF AGE ON VOCABULARY ACQUISITION
IN ENGLISH AS A FOREIGN LANGUAGE**

Tesi doctoral presentada per

Immaculada Miralpeix Pujol

com a requeriment per a l'obtenció del títol de

Doctora en Filologia Anglesa

Programa de Doctorat: *Lingüística Aplicada*
(Bienni 2000-2002)
Departament de Filologia Anglesa i Alemanya

Directors: **Dra. Carme Muñoz Lahoz i Dr. Paul M. Meara**

Universitat de Barcelona

2008

CHAPTER 1

INTRODUCTION

1.1. Age and Second Language Acquisition

During the 1990s, a growing number of governments in European countries brought down the age at which a Foreign Language (FL) was first introduced at schools. In Spain, the age at which children have started learning English was first moved from 11 (6th grade in primary education) to 8 (3rd grade)¹. This gave rise to new generations of students who, by the end of secondary education, had been learning a FL for at least ten years, three more than the previous generations who started at 11.

Several studies have shown the benefits of starting to learn a new language as early as possible. This has mostly been shown in naturalistic situations, that is, with people who learn the language in the country where the language is spoken. It normally happens in these cases that although children have a slower rate of development in the target language, they quite often surpass older learners in the long run achieving a

¹ Later on, an educational law passed in May 2006 made the study of English compulsory in primary education from Grade 1 onwards, therefore the official starting age was brought down to 6. The law has started being implemented in the present school year (2007-08). However, as reported by the European Commission (2005), since 2002 certain autonomous communities in Spain have been able to be teaching English from the age of 3 onwards.

superior ultimate attainment. In formal settings, that is, in non-immersion contexts, results are not so clear-cut: research demonstrates that ‘the sooner the better’ assumption cannot be applied in instructional settings in a straightforward way. As in naturalistic settings, older learners usually exhibit a quicker rate of development. Nevertheless, regarding ultimate attainment, differences in favour of early starters are rarely found.

Most of the studies concerned with age and FL learning have focussed on phonology and syntax and very few have dealt with vocabulary. One of the reasons to analyse these areas and not vocabulary may be that it is considered that words can be learned throughout the whole life and therefore lexis is not usually seen as having any age constraint as syntax or phonology may have. However, age may be argued to have an effect on the acquisition of vocabulary as well. First of all, some aspects of vocabulary are believed to be learned implicitly, and it has been argued that those aspects of language that are learned implicitly are maturationally constrained. Secondly, although neurological differences between syntax and vocabulary have been found, especially as regards brain location, these differences do not necessarily entail that if the acquisition of grammar is affected by age, that of vocabulary may not be influenced by learners’ age.

Therefore, we believe that there is a gap in research as regards the effects of age on the acquisition of vocabulary that the present study wants to bridge. In particular, *the first concern in our study is to explore if students who started learning English earlier will have better vocabularies than those who started later.*

1.2. Vocabulary acquisition

1.2.1. The importance of words

There are also other reasons why the present work focusses on lexis. They all have to do with the centrality of the lexicon in language learning, which can be appreciated in different aspects.

Firstly, the fact that vocabulary learning is essential in the learning of any language has been sometimes illustrated with the case that when we go abroad we take a dictionary, not a grammar book. This would be an example of what Wilkins means when asserting that “without grammar very little can be conveyed; without words, nothing can be conveyed” (1972:111). Similarly, Hatch (1983:74) acknowledged that “when our first goal is communication, when we have little of the new language at our command, it is the lexicon that is crucial [...] the words [...] will make basic communication possible”. Indeed, “lexical knowledge is known to be an absolutely crucial factor across the whole spectrum of L2 activities” (Singleton, 1999:3). This same idea is also found in Henriksen, Albrechtsen and Haastrup (2004), Kelly (1991), Linnarud (1986) and in Meara (1996:35), who points out that lexical competence is “at the core of communicative competence”. Vocabulary has also repeatedly been shown to be related to proficiency: the bigger your vocabulary is, the more proficient in a language you are. Therefore, vocabulary has been proved to be one of the best indicators showing how well one writes, reads, understands and speaks any language.

Secondly, after a period of neglect following the success of the Audiolingual Method and Communicative Language Teaching, there have been teaching programmes

assigning more and more importance to vocabulary, stemming from views that see language as ‘grammaticalised lexis’, not ‘lexicalized grammar’ (Lewis, 1993). Skehan (1998b) also reviews linguistic analyses that claim a more extended role for lexis in language. Thus, we find, for instance, the work by John Sinclair, highlighting the use of large lexical collocations, or methodologies by Michael Lewis, putting forward the “lexical approach” to language teaching, i.e. developing syllabuses completely based on vocabulary instead of grammar (see Harwood 2002 for an up-to-date review of key factors in the lexical approach). As Cook notices, one of the major challenges of learning and using a language may not be being competent at syntax but mastering all the aspects related to the lexicon (Cook, 1991).

Thirdly, learners themselves think that vocabulary is one of the most difficult components to master in a FL. Even after they have more or less mastered syntax, they still feel that their vocabulary is not good enough (Laufer, 1986). Vocabulary is an area which is perceived as being difficult to tackle and one which learners quite often think that they are unsure about. Furthermore, as stated above, learning words is a process that can take place during the whole life expanding until adulthood. We should not forget either that errors at the lexical level can make communication extremely complicated or even impossible.

1.2.2. The importance of measures

In spite of being such an extremely important area of language acquisition, there is not a wide range of reliable measures to control vocabulary development and test lexical proficiency at different levels.

Vocabulary knowledge can be of different types, one of the most common distinctions is that between receptive and productive vocabulary knowledge. There is some discrepancy in the fact that receptive and productive aspects of vocabulary knowledge can be seen as a continuum or can be considered as two different states in which a word could be at different moments of acquiring a language. In spite of this lack of agreement, it is generally assumed that reception precedes production and that receptive vocabulary is much larger than productive vocabulary. According to Nattinger (1988) and Meara (1990), receptive vocabulary is recognized when it is seen or heard but an external support is needed (visual, audio etc.). Productive vocabulary is retrieved from the speaker's or writer's memory without any external support.

There are different ways of measuring lexical reception and production. However, trying to quantify both receptive and productive vocabularies is not an easy matter, and as Melka (1997) points out, we do not have enough tools to do it successfully. Revising studies dealing with vocabulary measurement, one realises that new ways of coping with lexical diversity and dealing with vocabulary size are very much needed. This would pave the way for language acquisition and teaching research; for instance, we could study in more detail under which conditions lexicons develop more quickly or how far students have gone in a particular language course.

In our context, learners are Catalan/Spanish bilinguals who study English as the third language. According to the curricular requirements at the end of secondary education, students should be capable of reading and understanding basic texts on different topics, recognising about 5,000 words and communicating effectively in English, both in oral and written language (Dept. Ensenyament, 1990), which implies having a good productive vocabulary. Actually, after ten years of instruction, these would not seem difficult demands. Laufer (1992), for instance, has indicated that the minimal number of words constituting the lexical threshold for text comprehension is 3,000, i.e. a word knowledge inferior to 3,000 words will make reading comprehension very difficult. However, Waring and Nation (2004:12) have alerted that to gain adequate comprehension of a text a learner should have a coverage rate of 80%, that is, s/he should know more than the 80% of words in the text. Now, how do we measure then if our learners are accomplishing these goals?

There are multiple choice or translation tests and checklists for receptive vocabulary, but how shall we evaluate the language they produce? Some of the most typical measures of lexical richness have been recently called into question (Malvern et al., 2004), and if estimates of receptive vocabulary are being difficult to achieve, ways of estimating productive vocabulary are even more scarce: it is still very difficult to have a general approximation to how many words can be retrieved from memory in different situations. Therefore, the second aim of this dissertation is *to survey the widely used measures to compute lexical richness that exist nowadays and make new proposals to measure productive vocabulary size if the existing measures are found not to be fully satisfactory.*

1.3. The present study

In order to accomplish the purposes presented in the previous sections, this study has been structured and carried out in the following way:

Chapters 2 and 3 are devoted to giving an account of the literature of the research fields that the present work aims at making a contribution to: age studies and vocabulary measures studies. Chapter 2 reviews the literature on age and Second Language Acquisition (SLA), focussing especially on the reasons why vocabulary has not had an important role in research on age (section 2.5.1) and on the main findings of research dealing with age and Second Language (SL) vocabulary acquisition, particularly in formal contexts, as this is the case under study. Chapter 3 gives an overview of the measures that have been used to assess vocabulary knowledge. The studies reviewed are presented as attempts to give answers to frequently asked questions like ‘what do we mean by saying that a writer has a rich vocabulary?’ or ‘how can richness be assessed?’. As one of the objectives of the dissertation is that of obtaining productive vocabulary size estimates for different tasks, a comprehensive revision of studies that have tried to quantify the amount of words known by learners in the SL is presented. The chapter also tackles the question of the difficulties of estimating vocabulary size and wonders about which sort of ‘theoretical vocabulary’ should be used to evaluate the ‘observed vocabulary’, which is the vocabulary produced by the students.

Once the background has been provided, chapter 4 presents the methodology and procedure used for the analysis of the data. The instruments for the data collection are described and the participants that took part in the present research are presented. These

participants form part of the *Barcelona Age Factor* (BAF) Project, which analyses the effects of age on the acquisition of English as a FL in instruction settings. It is a long-term study which started in 1995 and which still continues nowadays in the English Department at the University of Barcelona, although the data analysed in this dissertation was collected between 1995 and 2004. The data analysed here comes from three oral and one written task as well as an English cloze. In addition to information on the method, the chapter also gives account of decisions that were taken previous to the data analysis and which are based on a pilot study presented in section 4.3.4.1. Some aspects related to the tools used and the statistical analyses performed are also considered.

Chapters 5, 6 and 7 consist in the analysis of the data, which comes from groups of students that share (or differ in) the age they started learning the FL (AO: Age of Onset), the age at which they were tested (AT) and the amount of exposure (AE) that they have received. The control of these three different variables allows for different studies that want to throw light on the research questions proposed. In the first place, chapter 5 uses intrinsic measures for the assessment of the students' production. Some traditional measures are adopted, as well as a new one: D, whose behaviour will be carefully examined in relation to the other indices (sections 5.2.2 and 5.3.4). Besides the analysis of cross-sectional data, the chapter also presents crucial information on the lexical development of longitudinal participants. In chapter 6, we make use of different extrinsic measures such as *Lexical Frequency Profiles* or *P_Lex* to examine the learners' productive vocabulary. The results obtained aim at complementing those displayed in the previous chapter and together with those resulting from chapter 7, offer a realistic

and complete picture of the productive vocabulary of the learners analysed.

Chapter 7 proposes a method to estimate the vocabulary size of the groups of learners in the study. A theoretical framework is set up for the inferential procedure presented and also the computation, which has been programmed into a new computer tool called *V_Size*, is exhaustively examined. In addition to the experimental work with the program (explained in section 7.3), estimates of productive vocabulary are calculated for the learners in the different groups.

Although each of the three chapters on data analysis contains its own discussion of the results, a general discussion that takes into account all the results from the work carried out is offered in chapter 8, where conclusions are also provided and issues for further research identified.

To close this dissertation, a list of the bibliographical references mentioned in the present study and seven appendices are included at the end.