



ELSEVIER

Contents lists available at ScienceDirect

System

journal homepage: www.elsevier.com/locate/system

Combining corpora and experimental data to investigate language learning during residence abroad: A study of lexical sophistication

Nicole Tracy-Ventura

University of South Florida, USA

ARTICLE INFO

Article history:

Received 16 September 2017

Accepted 17 September 2017

Available online xxx

Keywords:

Study abroad

Lexical sophistication

Spanish

Lexical frequency profiling

Learner corpus research

ABSTRACT

This study investigates the development of lexical sophistication as a result of residence/study abroad. Both knowledge and use of low-frequency vocabulary were analyzed before and at the end of a 9-month stay abroad. Data were collected from 27 English-speaking university learners of Spanish who spent an academic year in Spain or Mexico. Vocabulary knowledge was assessed twice using the Spanish version of the Swansea Levels 'X-lex' test (Meara & Milton, 2003) which tests recognition of words from different frequency bands. Vocabulary use was assessed based on oral and written data from a learner corpus collected at the same time points. A profiler for lexical sophistication was created based on Spanish frequency data from the *Corpus del Español* which calculates the ratio of low-frequency words to total words produced. Results demonstrate significant increases over time in both knowledge and use of low-frequency vocabulary. Findings are discussed in light of previous research on study abroad and the analysis of lexical sophistication.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Within the field of second language acquisition, the effect of residence/study abroad on language development has received considerable attention. However, most of this research has examined changes in overall proficiency and oral performance (e.g., fluency and accuracy) more so than other linguistic areas (Collentine, 2009; Kinginger, 2011; Llanes, 2011). In particular, vocabulary development has received little attention. It is also rare to find research comparing oral and written language abilities.

Of the few studies which have examined second language vocabulary acquisition during residence abroad (RA), most either assessed vocabulary knowledge using controlled elicitation techniques (Dewey, 2008; Fitzpatrick, 2012; Pizziconi, 2017) or vocabulary use via measures of lexical complexity (Foster, 2009; Serrano, Tragant, & Llanes, 2012). To date, none of these studies have investigated lexical sophistication, "a person's command of less common words" (Jarvis, 2013, p. 16) in oral or written production, yet it is an area in which increases would be expected. The ability to use low-frequency vocabulary is considered a sign of advanced proficiency (Lindqvist, Bardel, & Gudmundson, 2011), and much previous RA research has demonstrated that general proficiency is one area in which improvements are found (e.g., Guntermann, 1995; Hernández, 2010; Segalowitz & Freed, 2004). However, due to the limited amount of RA research focusing on vocabulary development, little empirical evidence exists. Studies that have investigated lexical sophistication more generally have primarily

E-mail address: nkt@usf.edu.

<https://doi.org/10.1016/j.system.2017.09.022>

0346-251X/© 2017 Elsevier Ltd. All rights reserved.

focused on L2 English, with a few studies on L2 French and Italian (e.g., Bardel, Gudmundson, & Lindqvist, 2012). No such studies exist on L2 Spanish, a popular language investigated in the study abroad literature (see Yang, 2016).

By analysing two different sources of data, the current study investigates how both knowledge and use of low-frequency vocabulary develops as a result of a 9-month stay abroad. Longitudinal data were collected from 27 English-speaking university learners of Spanish who spent an academic year in Spain or Mexico as part of their degree requirements. Vocabulary knowledge was assessed twice, before and at the end of their stay, using the Spanish version of the Swansea Levels 'X-lex' test (Meara & Milton, 2003) which tests recognition of words from different frequency bands based on L1 corpora. In addition, vocabulary use was assessed using learner corpus data that includes three different communicative tasks (oral interview, oral picture-based narrative, and written argumentative essay), collected at the same time points as the X-lex test. A profiler for lexical sophistication was created specifically for this study which provides a total of lemmas produced in the first five frequency bands (1–1000, 1001–2000, 2001–3000, 3001–4000, 4001–5000) established from the *Corpus del Español* (Davies, 2006).

2. Literature review

2.1. Lexical sophistication and lexical frequency profiling

Lexical sophistication is an aspect of lexical competence that distinguishes between the use of basic and advanced vocabulary. The difference has traditionally been based on how frequently words occur in language as measured in large reference corpora such as the *British National Corpus* (<http://www.natcorp.ox.ac.uk/>). Basic vocabulary items are considered those which are the most frequently occurring. Therefore, use of low frequency words is considered a sign of more sophisticated and advanced vocabulary knowledge. Kyle and Crossley (2015) provide an overview of the various ways that lexical sophistication has been measured in previous research. In addition to corpus-derived frequency counts, they discuss the number of different texts a word occurs in (*range*), the frequency of multi-word units (*n-grams*), the frequency of words and multi-word units more typical of academic language, as well as psycholinguistic properties of words. So far, the most well attested method has been corpus-derived frequency counts, described in more detail next.

One of the primary methods used to assess frequency is lexical frequency profiling (LFP: Laufer & Nation, 1995). LFP is typically based on word families and involves comparing each word produced by L2 learners to the sum of all the word family members (e.g., *writes*, *writing*, *written*) in a target language corpus. Results are provided in the form of frequency bands: the 1000 most frequent words, the 2000 most frequent words, etc. The popular VocabProfile program, available on the web via Tom Cobb's Compleat Lextutor website (<http://www.lexutor.ca/vp/eng/>), provides output based on the first 2000 words in English, in addition to the percentage of words that come from the Academic Word List (Coxhead, 2000). Any other words occurring in the text are included in the percentage of "off-list" words. Studies of L2 English have demonstrated that the LFP can be used in longitudinal designs to capture changes over time (e.g., Laufer, 1994; Tonkyn, 2012). Far fewer studies have used the LFP method in other languages, but this number is on the rise. For example, a French version of the LFP, developed by Cobb and Horst (2004), is now available on the Lextutor website and has been used in other studies on L2 French (e.g., Ovtcharov, Cobb, & Halter, 2006).

Lexical sophistication has also been a variable of interest in research using cross-sectional designs to examine whether it can be used as a reliable predictor of proficiency levels (e.g., Crossley & McNamara, 2012; Kyle & Crossley, 2015). Lindqvist et al. (2011) investigated L2 French, as well as L2 Italian, to examine whether LFP could be used to distinguish between proficiency levels in oral production. In comparison to most previous studies using LFP, the L1 corpus data that formed the basis of their results was based on spoken corpora. Their findings were in support of LFP analysis as a reliable method for distinguishing between proficiency levels, and the results also demonstrated that their advanced learner group achieved L1-like lexical profiles. Other work by Bardel et al. (2012) brought attention to how low-frequency cognates and thematic words found in language textbooks should be integrated into LFP programs. Their study demonstrated the superiority of the new method for distinguishing between proficiency levels. Although LFP appears to distinguish between proficiency levels in other Romance languages, this method has not yet been applied in L2 Spanish research. Fairclough and Belpoliti (2016) used LFP to investigate Spanish heritage language learners' use of low-frequency vocabulary in writing samples taken from a Spanish placement test, but their focus was not on development. Therefore, it is unknown whether LFP could be used to capture development in longitudinal designs and if low-frequency cognates would also be used frequently by English L1 learners of Spanish.

Despite the popularity of LFP, which is considered a band-based frequency measure, some researchers (e.g., Crossley, Cobb, & McNamara, 2013; Meara, 2005) have questioned whether it is sensitive enough to measure development. To examine this question, Crossley et al. (2013) investigated whether a count-based measure would be more appropriate for distinguishing between proficiency levels. A count-based measure "involves averaging individual word frequencies to one-integer frequency ratings for texts" (p. 967). In other words, each text would have one score which would be the average word frequency of all the words in the text based on their frequencies in a reference corpus (e.g., the British National Corpus). Based on results of their study comparing band-based and count-based measures, Crossley et al., concluded that both methods have their advantages. Because no studies comparing language development before and after study abroad have used LFP, the current study adopts this method as a first step in examining if learners increase their use of low frequency vocabulary after a 9-month stay abroad.

Download English Version:

<https://daneshyari.com/en/article/6849477>

Download Persian Version:

<https://daneshyari.com/article/6849477>

[Daneshyari.com](https://daneshyari.com)