ELIA
Estudios de Lingüística Inglesa Aplicada

Exploring the Use of Cognates in Lexical Availability: Differences between Heritage and Non-Heritage Learners.

Exploración del Uso de Cognados en la Disponibilidad Léxica: Diferencias entre Alumnos de Herencia y No Herencia.

Leah Geoghegan
Universidad de la Rioja
legeogw@unirioja.es

Abstract

Having previous knowledge of at least two languages may lead to advantages in the acquisition of an additional language in areas such as vocabulary, as learners may be better equipped to make cross-linguistic comparisons. However, while studies have been carried out on performance in the L1, there has been a clear lack of research on the lexical production of monolingual and bilingual EFL learners and, in particular, of bilingual heritage learners who have the same mother tongue. In addition, although cross-linguistic influence may also lead to negative effects when the influence derives from a lexical error, lexical fluency research to date has scarcely focused on the errors made by learners in lexical availability tasks. The current paper addresses these issues to determine whether heritage learners demonstrate an advantage in terms of cross-linguistic comparisons. To this effect, the study uses a semantic fluency task to analyse the English lexical availability of 134 tenth-grade L2 monolinguals and

L3 bilinguals to determine the positive and negative influence of cognates on their lexical productions. Results indicate that quantitative and qualitative differences depend on the lexical domain at hand, though very little negative lexical transfer was observed in either group. These results offer illuminating insights into the differences between heritage and non-heritage learners' lexical availability and should encourage stakeholders in our increasingly multilingual classrooms to foster the language acquisition of these bilingual learners.

Keywords: heritage learners; lexical availability; cognates; vocabulary production.

Resumen

Tener conocimientos previos de al menos dos lenguas puede suponer ventajas en la adquisición de una lengua adicional en áreas como el vocabulario, ya que los alumnos pueden hacer más fácilmente comparaciones interlingüísticas. Sin embargo, mientras que se han realizado estudios sobre el rendimiento en la L1, hay una clara falta de investigación sobre la producción léxica de los estudiantes monolingües y bilingües de inglés como lengua extranjera, y, en particular, de los bilingües de herencia que tienen la misma lengua materna. Además, aunque la influencia interlingüística también puede producir efectos negativos cuando la influencia se deriva de un error léxico, la investigación sobre la fluidez léxica hasta la fecha apenas se ha centrado en los errores cometidos por los alumnos en tareas de disponibilidad léxica (DL). El presente trabajo aborda esta cuestión para determinar si los alumnos de herencia demuestran una ventaja en comparaciones interlingüísticas. Para ello, utiliza una tarea de fluidez semántica para analizar la DL de 134 monolingües L2 y bilingües L3 de 4º ESO para determinar la influencia positiva y negativa de los cognados en sus producciones léxicas. Los resultados indican que las diferencias cuantitativas y cualitativas dependen del dominio léxico, aunque se observó muy poca transferencia léxica negativa en ambos grupos.

Estos resultados ofrecen una visión esclarecedora de las diferencias entre esos alumnos y deberían animarnos a fomentar la adquisición lingüística de estos alumnos bilingües en nuestras aulas cada vez más multilingües.

Palabras clave: alumnos de herencia, disponibilidad léxica; cognados; producción de vocabulario

1. Introduction

Over the past two decades, there has been a steady increase in the number of foreign students enrolled in the Spanish education system. In the 2023-2024 academic year, for example, enrolments increased a further 7.1%, exceeding one million for the first time (Ministerio de Educación, Formación Profesional y Deportes, 2025). By country, the largest group comes from Morocco (18.9%), followed at a distance by Romania (9.3%), Colombia (8.8%), Venezuela (5.9%), Peru (4.3%), China (4.1%) and the Ukraine (3.8%). In some cases, these individuals arrived in Spain relatively recently and have varying levels of Spanish proficiency. In other cases, the students' families moved to Spain before they were born, so these students have grown up speaking Spanish their entire lives. In addition to the official language of the country in which they live, in this case Spanish, these students may speak another language at home with their family.

Berthele & Udry (2022) have pointed out that these heritage learners may be at an advantage when learning an additional language. For instance, a bilingual heritage speaker who has grown up in Spain speaking Spanish in school and Arabic or Romanian at home may avail themselves of their previous linguistic knowledge to facilitate their acquisition of a foreign language such as English. As highlighted by Agustín-Llach (2019, p. 53), numerous studies, such as Gabrys-Barker (2006), Kroll, Bogulski, and McClain (2012) and Otwinowska (2016), have corroborated this advantage for bilinguals, finding that bilinguals learning another foreign language have "a higher repertoire of learning strategies, an increased metalinguistic

awareness, better memory performance, and are more verbally creative."

Numerous studies have also emerged over the past ten years regarding bilingual heritage speakers specifically, supporting this multilingual advantage for such learners when learning English as a Foreign Language (EFL). However, Lorenz et al. (2020) point out the controversy regarding this advantage, as it may depend on different factors such as the type of bilingual, the type of language skill under analysis, cognitive skills, language background, age, language status and influence, the context, socioeconomic status, and literacy. In terms of the type of language skill, previous research has found that general proficiencies, such as oral and writing skills, may be more positively affected than skills related to grammatical development (Cenoz, 2003; Siemund & Lechner, 2015; Siemund & Mueller, 2020). Regarding vocabulary, heritage learners may have an advantage in this area when learning a foreign language. Klein (1995), for example, indicates that as bilinguals have two prior lexicons, their lexical awareness is greater than that of monolinguals learning an additional language. As a result, bilinguals may be better prepared to make cross-linguistic comparisons. Therefore, they may be able to take advantage of the presence of cognates to help them learn the foreign language vocabulary more easily (Molnar, 2008; Otwinowska, 2016). This benefit, referred to as the cognate facilitation effect, suggests that bilinguals find cognate words easier and faster to recognise than non-cognate words (Costa, Caramazza & Sebastian-Gallés, 2000; Muñoz, 2020). Thus, when learning a foreign language, heritage learners may be able to draw on their existing languages and use the presence of cognates to facilitate the acquisition of the foreign language.

However, most previous research on heritage learners has been carried out on performance in the L1, while the lexical production of second language (L2) and third language (L3) EFL learners has largely been overlooked. Furthermore, the few studies dealing with L3 EFL learners have generally included a very small number of bilingual heritage learners, generally with varying L1s,

which is evidently problematic if the purpose of the research is to compare the influence of the learners' vocabulary in one language on another. In addition, while proficiency in multiple languages may provide potential advantages, the influence of other previously known languages may also lead to negative effects when the influence derives from a lexical error (Agustín-Llach, 2019). While this is an important issue, lexical fluency research to date has scarcely focused on the errors made by learners in lexical availability tasks. The present study addresses these gaps in the literature by focusing on heritage and non-heritage learners' lexical availability in EFL. In addition, the study examines not only the quantitative differences between these groups of learners, but also the qualitative differences in positive and negative cross-linguistic influence (CLI). To this effect, the following section will survey the key concepts and previous research in these areas, namely, lexical availability, CLI, and previous research on L2 and L3 EFL learners.

2. Literature Review

2.1 Lexical Availability

Lexical availability refers to "the words that come to mind in response to topics related to daily situations" (Jiménez Catalán & Fernández Fontecha, 2019, p. 77). For example, in a category such as 'Animals', it is very likely that the first words which come to mind are those such as *cat* or *dog*. The origins of lexical availability date back to the 1950s in France, where word frequency lists were typically used as a means of deciding the vocabulary which was to be taught to learners of French (López Morales, 2014). This practice was thought to be logical as it was expected that words on these lists would be the most useful and the most used. However, common everyday words such as *fork*, *subway* or *tooth* did not feature on these lists, which led researchers to question the usefulness of word frequency as a criterion for selection (Sánchez-Saus Laserna, 2024). As an alternative, Michéa (1953) proposed using lexical availability,

a formula which "proved more effective at determining the rate of word usage than the frequency-based methodology, as it took into account the distribution of frequencies in the text and avoided potential oversights" (Sánchez-Saus Laserna, 2024, p. 3). This distinction between frequency and availability appeared to be crucial, and led to the first lexical availability studies in the 1960s and 1970s, such as Gougenheim et al. (1964) in France, Dimitrijévic (1969) in Scotland, Mackey (1971) in Canada, and López Morales (1973) in Puerto Rico.

In these initial studies and in the research since then, lexical availability has been examined via the data-collection instrument known as the Lexical Availability Task (LAT), leading to clear benefits in methodological homogeneity. Generally carried out by means of a paper-and-pencil questionnaire, participants are first presented with prompts (with each prompt appearing on a different page with numbered lines), and then asked to write down all words that come to mind in response for two minutes (Samper Hernández & Jiménez Catalán, 2014). To date, most studies have also focused on the 16 prompts first proposed by Gougenheim et al. (1964) in the first lexical availability studies, allowing cross-comparison across a wide range of studies. The systematicity in the assessment of lexical availability has led researchers to conclude that the LAT is a reliable task with consistent results when testing the L1 as well as EFL (Canga Alonso, 2017). In addition, the task allows researchers not only to determine the lexical resources available to learners, but also to collect this rich set of data in a remarkably efficient and economical way (Jiménez Catalán & Fitzpatrick, 2014).

2.2 Cross-linguistic Influence and Lexical Availability

Cross-linguistic Influence (CLI) is "the influence of a (previously known) language on another language (being learned)" (Agustín-Llach, 2019, p. 50). When referring to the lexical level, CLI may also be referred to as "lexical transfer." In other words, a word from a

language which the speaker already knows may be transferred to the target language. This influence may be positive, such as when words in two languages are cognates, that is, words which share formal and semantic specifications in the native or source language and in the L2. For example, a Spanish speaker learning English may instantly recognise the word *broccoli*, even when they have never seen it before, simply due to its similarity to the word in their L1: *brócoli*. In such cases, when an L1 and a TL contain a high number of cognates, learners are evidently at an advantage, as linguistic resources are maximized: with one single lexical entry, learners have knowledge of two languages. Conversely, negative CLI may also be observed when two words which appear to be very similar actually have different meanings, such as in the case of so-called *false friends*. For example, an English speaker learning Spanish may mistakenly use the word *embarazada*, believing it to mean *embarrassed*, rather than its real meaning, *pregnant*.

In terms of lexical availability, researchers such as Jarvis (2009) have highlighted the assumed simultaneous activation of the L1 and the TL. As a result, previous linguistic knowledge may play a determining role in what words learners produce because if both languages are activated at the same time, it is very likely that there will be CLI in the retrieved words. Regarding monolingual and bilingual learners of a new foreign language, previous research has been inconsistent. On the one hand, Bogulski and Kroll (as cited in Kroll, Bogulski & McClain, 2012) have found little difference between these groups, which they attribute to the bilingual learners' inability to apply the inhibitory strategy for the new L2. On the other hand, Kaushanskaya and Marian (2007) have found that bilingual learners have a superior inhibitory control mechanism when using the L2. In addition to these different results, as Agustín-Llach (2019) points out, there has also been a clear scarcity of research examining CLI in lexical availability research. This research novelly addresses this issue, also comparing L2 and L3 learners, the results of which are presented in the following section.

2.3. Lexical Availability Research on L2 and L3 EFL learners

As previously noted, most lexical availability research with heritage and non-heritage learners to date has focused on the L1, largely neglecting the lexical availability of EFL learners. Some recent studies have, however, aimed to address this gap, focusing on learners for whom English is the L3, and who speak either Spanish at home (L1) and Basque at school (L2), or another language at home (L1) and Spanish at school (L2).

Firstly, Agustín Llach (2023) investigated the lexical profiles of 12th-grade L2 and L3 EFL learners via a lexical fluency task. A total of 42 participants were divided into three groups depending on their monolingual or bilingual status: one group had 14 Spanish monolingual EFL learners (monolingual EFL learners); the second group had 14 Spanish-Basque bilinguals also learning EFL (educational bilingual EFL learners); and the third group had 14 bilingual L3 EFL learners (environmental bilingual EFL learners). The last group of heritage learners included learners with a range of different L1s in the home context, including Arabic (n = 3), Romanian (n = 5), Portuguese (n = 2), Armenian (n = 2), Georgian (n = 1), and Russian (n = 1). Participants completed the Oxford Placement Test (OPT) and a LAT with four different prompts: 'Food and Drink', 'Hobbies', 'Animals', 'Town' and 'Countryside'. Results indicated no statistically significant differences between the number of words produced by the different groups of learners, suggesting little difference in the learners' lexical availability. However, analysis by means of graph theory metrics revealed that educational bilinguals exhibited higher levels of lexical organization as well as stronger connections among the nodes. These learners can thus be deemed to have a mental lexicon which is, as the researcher notes, better organised, more compact, and more stable.

Secondly, Fernández-Fontecha, Jiménez Catalán and Ryan (2024) also investigated the lexical availability of L2 and L3 EFL, again taking not only a quantitative approach, but also a qualitative one whereby they analysed the participants' lexical organisation and

production strategies as well as global and local semantic relatedness. The participants included 32 L1 Spanish EFL learners (monolingual L2 EFL learners) from La Rioja and 28 L1 Spanish, L2 Basque EFL learners (bilingual L3 EFL learners). Both groups of students were in 12th grade and they were homogeneous in terms of their L1, Spanish, which was used by all participants in the home context, and their English proficiency level. The key difference between the groups was that the L3 group attended a Basque-medium school where all content subjects were delivered in this language. Learners completed a linguistic background questionnaire, the Oxford Placement Test, and a LAT addressing the category 'Animals'. In addition to comparing the number of responses by each group, lexical organisation and production were also analysed using word2vec, a distributional semantic analysis model, by means of the VFClust tool. Results again revealed a lack of quantitative differences between the two groups, as L2 and L3 learners produced a similar number of words. However, the L3 group produced more diverse word classes and grammatical categories, more cognates, and more idiosyncratic responses related to their family life and episodic memories. The L2 group, on the other hand, were found to be more homogeneous in terms of how they explored the semantic field locally, excelling in most semantic relatedness variables. The authors suggest that these differences may be attributed to an activation of their extra language, Basque, by the L3 learners.

Finally, Agustín-Llach (2019) also investigated the issue with 12th grade EFL learners: 86 Spanish L1 learners and 9 learners with another home language (5 Romanian, 1 Arabic, 2 Portuguese and 1 Basque). Participants completed a pen-and-paper LAT with a total of 15 prompts: 'Parts of the Body', 'Clothes', 'The House', 'Black and White', 'Food and Drink', 'Make', 'Animals', 'Sad', 'School', 'The Town', 'The Countryside', 'Love', 'Professions', 'Hobbies', and 'Hate'. While the cohort was very similar to the above studies, this study is particularly novel given its focus on examining lexical access and lexical CLI through the LAT. CLI instances were classified in terms

of phonetic spelling, lexical creation, literal translation, semantic extension, and cognate use. Results showed relatively little presence of negative CLIL, while cognate use was the most frequent type of CLI. Comparison between the two groups revealed no statistically significant differences in either the mean number of tokens produced or any of five categories of lexical CLI.

The above results are thus somewhat inconclusive regarding the advantage for L3 heritage learners in terms of lexical availability and CLI. The first two studies indicate that, although little quantitative difference has been observed between bilingual and monolingual learners' lexical availability in EFL, there may be some qualitative benefits in terms of CLI. The third study, however, suggests no difference in the use of cognates between these two groups of learners. However, in the above research, it is important to distinguish between what Agustín Llach (2023) refers to as educational bilingual EFL learners (i.e., learners enrolled in a school where all content classes are delivered in a language other than the L1) and environmental bilingual EFL learners (i.e., learners who speak a different language at home than that spoken in the country in which they are living). The latter group is akin to what we refer to here as L3 heritage bilinguals. The numbers of environmental bilingual EFL, or heritage, learners in the above studies are, however, notably very small: 14 in Agustín Llach (2023) and 9 in Agustín-Llach (2019). This is understandable, as the cohort of L3 heritage bilinguals in these studies will naturally represent the percentage of these learners in the typical classroom. However, an undesirable consequence of this small corpus is the linguistic heterogeneity among the heritage learners: in one study, 14 learners had six different L1s, while in the other, 9 learners had four. This is evidently problematic if the goal of the study is to investigate lexical transfer, as this factor will necessarily differ depending on the various languages involved. In other words, there is little reason to expect similar levels of lexical transfer between an EFL learner who speaks Spanish and Arabic and one who speaks Spanish and Russian. Thus, it is necessary to carry out further research on the lexical transfer of

a larger group of L3 heritage learners who speak the same L1. Furthermore, as suggested by Agustín-Llach (2019), CLI may also lead to negative effects when the influence derives from a lexical error. While this is an important issue, lexical fluency research to date has scarcely focused on the errors made by learners in lexical availability tasks.

The present study addresses these issues. Specifically, the study seeks to compare the lexical availability of non-heritage and heritage EFL learners, including a larger cohort of heritage learners with the same L1. In addition, the study aims to compare the learners in terms of the number of words they can produce, and the instances of positive CLI, as measured by cognate use, and negative CLI, as measured by the presence of Spanish words. To this effect, the study addresses the following research questions:

> RQ1: Are there quantitative differences in the lexical availability and language proficiency of L2 monolinguals (non-heritage learners) and L3 bilinguals (heritage learners) in a Spanish context?
>
> RQ2: Are there qualitative differences in terms of cross-linguistic influence?

3. Methodology

3.1. Research Approach and Design

The current study has been carried out as part of a national project (HERPRO) which explores productive vocabulary in English as a L2 and L3 in secondary school contexts, and analyses cognitive, psycholinguistic and sociocultural dimensions. It is a coordinated project involving researchers from the University of La Rioja and the University of Extremadura, and is divided into two subprojects. The first subproject (HERLEX), from which the study at hand is derived, focuses on the psycholinguistic and sociocultural dimensions (Project PID 2022-137337NB-C22), while the second

(HERCOG) focuses on the cognitive dimension (Project PID 2022-137337NB-C22).

The project includes a total of 434 students in the 10th grade of compulsory secondary education (4th-year ESO students). Of these, 314 are native Spanish speakers learning English in the school context (henceforth, L2 non-heritage learners), while 120 are learners who, in addition to speaking Spanish and learning EFL, are heritage learners who speak another language at home (henceforth, L3 heritage learners). Languages spoken at home were predominantly Arabic (57.5%), Romanian (15.8%) and Urdu (6.7%), with Georgian, Hindi, Italian, Polish, Portuguese, Fula, Bulgarian, Ukranian, and Uzbek being spoken by the other 20% of learners. These participants came from ten secondary schools, with a mixture of both urban and rural schools: five in the autonomous community of La Rioja and five in the autonomous community of Extremadura. Data were collected over two sessions given the number of tests involved.

3.2. Instruments

As noted above, the national project seeks to address a number of different factors including productive vocabulary, creativity, and metaphorical competence. To this effect, instruments included the following:

1. Biodata and language background questionnaire
2. Oxford Placement Test (OPT)
3. Lexical Availability Task (LAT)
4. Productive Vocabulary Task
5. Creativity Task (PIC-J Prueba de Imaginación creativa para jóvenes)
6. Written composition

7. Metaphorical competence test

The current study focuses on the learners' language proficiency and their lexical availability. To this end, it firstly uses the results of the OPT (Version 2, UCLES, 2002), which contains two parts: Part 1 (Questions 1-40), which contained questions on topics like notices, word choice, and short passages; and Part 2 (Questions 41-60), which focuses on vocabulary in longer passages. As outlined in the test instructions, Part 2 should be attempted only if Part 1 is finished without issues. The students, who in general were not expected to have more than a B1 proficiency level, were given 20 minutes to complete the test to allow enough time to complete Part 1.

Secondly, learners' lexical availability was assessed by means of the LAT, a time-controlled task which activates vocabulary in response to a particular stimulus, with the aim of measuring the number of words which learners' retrieve (Fernandez-Fontecha et al., 2024). The LAT in the project included the following 13 prompts: 9 in English and 4 of those in Spanish:

1. Animals
2. Town
3. Countryside
4. Fruits and Vegetables
5. Love
6. Food and Drink
7. Festivities and Celebrations
8. Daily Activities
9. Physical Activity
10. Frutas y Verduras
11. Fiestas y Celebraciones
12. Amar / Amor
13. El campo

The current study focuses on two prompts carried out in English: one natural, or taxonomic, category ('Food and Drink') and one more open, or slot-filler, category ('Love'). As outlined above, while LATs have typically collected data by means of a paper-and-pencil questionnaire, a novelty of the current project is the collection of data through a specifically designed online application (HerPro App). The application functioned in the same way as a typical paper-and-pencil LAT: one prompt appeared at a time and the participants had to type as many responses as came to mind as a reaction to that prompt for two-minutes. At the end of this time, responses were saved automatically, and the next prompt would appear. The procedure was then repeated until all prompts had been completed, and the learners then moved on to the next test.

3.3. Participants

The study at hand focuses on a subgroup of L3 learners with the aim of analysing the lexical availability of a homogenous group of heritage learners who speak the same L1 (see Section 2). Criteria for selection thus included: (1) having Arabic as a home language; and (2) having completed the LAT, or at least the prompts 'Food and Drink' and 'Love'. In addition, the same number of L2 non-heritage learners were selected at random in order to compare the two groups. This resulted in a total of 134 tenth-grade EFL learners: 67 L3 Spanish/Arabic learners, and 67 L2 Spanish learners.

3.4. Data Analysis

Prior to data analysis, the OPT was corrected, awarding one point for each correct answer for a total of 60 points. The LAT was then lemmatised according to the procedure outlined by Jiménez Catalán and Agustín Llach (2017). This procedure included (i) the correction of spelling errors; (ii) the deletion of unintelligible words and Spanish L1 words in the English data and English words in the

Spanish data; (iii) the repetition of the same word in the same prompt being counted only once; (iv) lemmatizing lexical phrases and compound words as one lexical unit and counting them as one word (e.g., *orange-juice*); (v) deleting proper nouns; and (vi) changing plural words into the singular form, except in cases where the word is always plural (e.g., *potatoes* to *potato*).

Once the tests were corrected and prepared, the data was analysed. In order to analyse the quantitative differences between the groups, independent samples t-tests were carried out using SPSS (Version 26), comparing the two groups in terms of their results of the OPT and the number of responses retrieved in the LAT for each of the two prompts. To analyse the qualitative differences between the groups' lexical availability, WordSmith Tools (Version 5) was used to create word lists for each prompt, for each group and for all students together. These lists were then analysed to determine the presence of Spanish/English cognates. Finally, to examine the effect of negative CLI, words which were removed during lemmatisation were analysed to determine whether these words were removed due to being proper nouns, a Spanish word or unintelligible.

4. Results and Discussion

4.1. Quantitative Differences in Language Level and Lexical Availability

The first research question asked whether there were quantitative differences in the lexical availability and language proficiency of L2 monolinguals (non-heritage learners) and L3 bilinguals (heritage learners) in a Spanish context (Table 1). In terms of language proficiency, results of the independent samples t-tests revealed a statistically significant differences between the two groups, with L2 learners ($M = 27.15$, $SD = 8.31$) outperforming the L3 learners ($M = 21.69$, $SD = 8.91$); $t = 3.64$, $p = <.001$. This suggests that the non-heritage learners had a higher level of English language proficiency than the heritage learners. In terms of lexical availability, however, there was

a clear difference depending on the prompt at hand. For 'Food and Drink', an advantage was again seen for the non-heritage learners who retrieved a statistically significant higher number of words (M = 11.72, SD = 4.29) than the heritage L3 learners (M = 8.85, SD = 4.48); t = 3.78, p = <.001. For the prompt 'Love', on the other hand, no statistically significant difference was observed between the L2 learners (M = 8.06, SD = 3.46) and the L3 learners (M = 7.52, SD = 4.90); t = 0.46, p = .464. This indicates that while non-heritage learners produce significantly more words in the lexical domain of 'Food and Drink', no such difference exists between the groups in the lexical category 'Love'.

Table 1: *Quantitative differences between non-heritage and heritage learners*

	M		*SD*		*t*	*df*	*p*
	L2	L3	L2	L3			
Language Proficiency	27.15	21.69	8.31	8.91	3.64	130	<.001
LA: Food and Drink	11.72	8.85	4.29	4.48	3.78	132	<.001
LA: Love	8.06	7.52	3.46	4.90	0.46	118.7	.464

In terms of the prompt 'Food and Drink', these results are inconsistent with previous studies by Agustín Llach (2023) and Fernández-Fontecha et al. (2024), which found similar results by L2 and L3 learners in terms of number of words produced. However, the results here may largely be attributed to language proficiency, given that the L2 group also had a higher proficiency level than the L3 group. This interpretation is based on suggestions by van Ginkel and van der Linden (1996) that students with a higher language level will be able to produce a higher number of words. However, in terms of the prompt 'Love', no statistically significant difference was observed, despite the higher language level of the L2 learners. In other words, the non-heritage learners' higher language proficiency did not result in them producing a higher number of words in this specific lexical domain. This finding may be attributed to differences in the type of prompt, as the former is a taxonomic prompt, whereas the latter is a slot-filler prompt. This is in keeping with recent

research by Agustín-Llach and Palapanidi (2024), which found clear differences in terms of the lexical availability and production of typical exemplars of native speakers and learners with different language level proficiencies, particularly in terms of slot-filler categories as opposed to taxonomic categories. This important difference between the type of prompt at hand will be addressed further in the qualitative analysis below.

4.2. Positive and Negative Cross-Linguistic Influence

The second research question asked whether there were qualitative differences in terms of CLI, and whether this influence was positive and/or negative. To address positive CLI, word lists were created for each of the two prompts at hand, outlining words produced uniquely by each group as well as words produced by both groups. These lists were then analysed to determine the presence of Spanish/English cognates, and a percentage was calculated for each list to determine how many cognates each of the three lists contained. Table 2 and Table 3 below provide an example of the word lists, indicating the first ten words on each list. Cognates on the lists are highlighted in grey, and the total percentage of cognates are indicated at the bottom. For example, for the prompt 'Food and Drink', L2 learners produced a total 74 words which had not been retrieved by the L3 learners. Of these, 22 were Spanish/English cognates, such as *bottle* or *bowl*, which accounted for 29% of the total words produced.

Table 2: *Positive Cross Linguistic Influence in the Prompt 'Food & Drink'*

L2 Only	L3 Only	Both
AVOCADO	ALCOHOLICDRINK	ALCOHOL
BAKEDCHICKEN	BAR	APPLE
BEAN	BEET	APPLEJUICE
BERRY	BLAND	APPLEPIE
BIRTHDAY	BROCCOLI	BACON
BLUEBERRY	CAFE	BANANA
BOTTLE	CAULIFLOWER	BEEF
BOWL	CHEESECAKE	BEER
BREAKFAST	CHEW	BISCUIT
BUBBLEGUM	CHICKENNUGGET	BREAD
…	…	…
22/74 (29%)	23/53 (43%)	42/90 (46%)

Note: Shaded columns indicate that the word is a Spanish/English cognate; figures in the final row indicate the number of cognates out of the total number of words on the list.

Table 3: *Positive Cross Linguistic Influence in the Prompt 'Love'*

L2 Only	L3 Only	Both
ACT OF SERVICE	AFFAIR	AFFECTION
A LOT	AMAZING	BABY
ANGEL	ANIMAL	BEAUTIFUL
ANNIVERSARY	AUNT	BEST FRIEND
ARROW	BALLOON	BESTIE
ATTENTION	BLUE	BOYFRIEND
BEACH	BOMBASTIC	BRIDE
BEACHWITHYOUR GIRLFRIEND	BOND	CANDY
BE KIND	BOOK	CARD
BELIEVE	BOY	CHOCOLATE
…	…	…
30/109 (28%)	25/94 (27%)	22/68 (32%)

Note: Shaded columns indicate that the word is a Spanish/English

cognate; figures in the final row indicate the number of cognates out of the total number of words on the list.

As shown, for the prompt 'Food and Drink', there was a clear difference between the number of cognates in the unique words by L2 learners (29%) and the ones by L3 learners (43%). It appears that, for this prompt, heritage learners have a higher percentage of Spanish/English cognates in their unique words than the non-heritage learners. Shared words also contained a higher percentage of cognates (46%), indicating that positive CLI was more frequent in the more common words produced by both groups of learners. On the other hand, less of a difference was observed between the two groups for the prompt 'Love', with cognates accounting for 27% to 28% of the unique words for each group, and a slightly higher percentage of 32% in their shared words. This suggests that both groups rely on cognates to a similar degree in this category, and again, that there is a difference in terms of the lexical availability prompt at hand.

In terms of the negative CLI in the participants' lexical availability, Table 4 below shows the words which were excluded, because they were either proper nouns, a Spanish word, or unintelligible.

Table 4: *Negative Cross Linguistic Influence*

		Proper Nouns	Spanish	Unintelligible	Total
L2	Food & Drink	13	2	3	18
	Love	0	1	0	1
L3	Food & Drink	19	3	0	22
	Love	4	2	0	6

As shown, the results were rather similar for both groups: there was a far greater number of removed words in the prompt 'Food and Drink' than the prompt 'Love' for both L2 and L3 learners (18 versus 1 and 22 versus 6, respectively). However, most of the removed words consisted of proper nouns in the prompt 'Food and Drink'. This was

generally due to the fact the learners included brand names in this category, such as the names of different soft drinks. Words written in Spanish were observed only in three cases in the prompt 'Food and Drink' for L2 learners, and not at all for L3 learners in either prompt. This suggests practically no observable negative CLI for either group in these two lexical domains.

In summary, there appears to be a difference between the two groups in terms of positive CLI for the prompt 'Food and Drink', with L3 learners producing a higher percentage of Spanish/English cognates. This finding is consistent with suggestions by Otwinowska (2016) that bilingual learners are better equipped to make cross-linguistic comparisons and benefit from cognateness effects. However, it appears that this advantage is only observed in the taxonomic prompt 'Food and Drink', as no difference is observed in the open, slot-filler prompt 'Love'. In addition, despite the higher presence of cognates in the words retrieved by heritage learners in the prompt 'Food and Drink', the quantitative analysis revealed that the non-heritage learners still produced a higher number of words in this prompt than the heritage learners. In other words, reliance on cognates did not lead to a quantitative advantage in this category.

5. Conclusions

In conclusion, results firstly revealed that L2 non-heritage learners performed better on the OPT and the prompt 'Food and Drink' than the L3 heritage learners, while there was no statistically significant difference for the prompt 'Love'. In addition, a clearly higher percentage of cognates was observed for L3 learners' unique words in the prompt 'Food and Drink' whereas little difference was found for the prompt 'Love'. Furthermore, very little negative CLI was found in terms of the presence of Spanish words in the English lists. This indicates that learners in both groups are successful in supressing negative CLI.

While these results offer some novel findings regarding the

differences between heritage and non-heritage learners' lexical availability and CLI, several limitations should be addressed in future research. Firstly, one of the novelties of the present study is that it included heritage learners who shared the same L1, as opposed to previous research which has used heterogenous samples of learners with multiple different L1s. However, this study so far has included only an analysis of the Spanish/English cognates. If we are to analyse these learners' vocabulary, it is imperative that we also take into account these learners' L1, Arabic, and also provide an analysis of the presence of Arabic cognates in the data. This would allow us to understand whether there is CLI not only from the learners' L2, but also their L1. In addition, the study provides an analysis of only two of the 13 prompts in the data. There is evidently a need to further explore the other prompts to determine whether similar findings would be found in other taxonomic and slot-filler prompts. It would also be very beneficial to explore and compare the four prompts which learners completed in both English and Spanish (e.g., 'Love' vs. 'Amor/Amar') to determine whether there is CLI in the participants' Spanish responses, and how these responses vary in English.

Despite these limitations, the present study offers some provisional insights into a under researched area, highlighting clearly important issues for further research. These findings should benefit secondary school EFL teachers who are increasingly working in multilingual classrooms with not only non-heritage Spanish speakers, but also numerous heritage learners with varying linguistic profiles. In particular, the results should encourage such teachers to adopt multilingual strategies in the classroom and tailor instruction to leverage the linguistic strengths of bilingual students. Awareness of the needs of these learners and whether the acquisition of cognates can indeed support their acquisition of EFL is an important consideration for these stakeholders nationwide.

Acknowledgements

This study was partially funded by research project PID2022-137337NB-C21, funded by MICIU/AEI/10.13039/501100011033/ and ERDF/UE.

References

Agustín Llach, M. P. (2023). Mapping the mental lexicon of EFL learners: A network approach. *Revista de lingüística y lenguas aplicadas, 18*, 1-17. https://doi.org/10.4995/rlyla.2023.18769

Agustín-Llach, M. P. (2019). CLI in lexical accessibility. In M. J. Gutierrez-Mangado, M. Martínez-Adrián & F. Gallardo-del-Puerto (Eds.), *Cross-linguistic influence: From empirical evidence to classroom practice* (pp. 49-63). Springer International Publishing. 10.1007/978-3-030-22066-2_3

Agustín-Llach, M. P., & Palapanidi, K. (2024). Apples, tomatoes, and health: comparison of structural characteristics of the L1 and L2 mental lexicons. *Applied Psycholinguistics, 45*(5), 873-896. https://doi.org/10.1017/S0142716424000328

Berthele, R., & Udry, I. (2022). Multilingual boost vs. cognitive abilities: Testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism, 19*(1), 142-161. 10.1080/14790718.2019.1632315

Bogulski, C. A., & Kroll, J. F., (in preparation). Vocabulary acquisition and inhibitory control: A paradox of bilingualism or two sides of the same coin?

Canga Alonso, A. (2017). Spanish L1 speakers' and EFL learners' available lexicon. *Anuario de estudios filológicos, 40*, 5-23. 10.17398/2660-7301.40.5

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism, 7*(1), 71-87. 10.1177/13670069030070010501

Costa, A., Caramazza, A. & Sebastian-Gallés, N. (2000). The cognate

facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(5). 1283–1296. 10.1037//0278-7393.26.5.1283

Dimitrijévic, N. (1969). *Lexical availability*. Julius Gross Verlag.

Fernández-Fontecha, A., Jiménez Catalán, R. M., & Ryan, J. (2024). Lexical production and organisation in L2 EFL and L3 EFL learners: A distributional semantic analysis of verbal fluency. *International Journal of Multilingualism, 21*(1), 60-77. 10.1080/14790718.2021.2016770

Gabrys-Barker, D. (2006). The interaction of languages in the lexical search of multilingual language users. In J. Arabski (Ed.), *Cross-linguistic influences in the second language lexicon* (pp. 144–166). Multilingual Matters. https://doi.org/10.21832/9781853598579-015

Gougenheim, G., R. Michéa, P. Rivenc & A. Sauvageot (1964). L'élaboration du Français Fondamental (1er degré). *Étude sur l'élaboration d'un vocabulaire et d'une grammaire de base*. Paris: Didier

Jarvis, S. (2009). Lexical transfer. In Pavlenko, A. (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 99–124). Multilingual Matters. https://doi.org/10.21832/9781847691262-007

Jiménez Catalán, R. M., & Agustín Llach, M. P. (2017). CLIL or Time? Lexical profiles of CLIL and non-CLIL EFL learners. *System, 66*, 87-99. https://doi.org/10.1016/j.system.2017.03.016

Jiménez Catalán, R. M., & Fernández Fontecha, A. (2019). Lexical Availability Output in L2 and L3 EFL Learners: Is There a Difference? *English Language Teaching, 12*(2), 77-87. 10.5539/elt.v12n2p77

Jiménez -Catalán, R. M., & Fitzpatrick, T. (2014). Frequency profiles of efl learners' lexical availability. In R. M. Jiménez Catalán (Ed.), *Lexical Availability in English and Spanish as a Second Language* (pp. 83-100). Springer. https://doi.org/10.1007/978-94-007-7158-1_6

Kaushanskaya, M., & Marian, V. (2007). Age-of-acquisition effects in the development of a bilingual advantage for word learning. In *Proceedings of the Boston University Conference on Language Development*.

Klein, E. C. (1995). Second versus third language acquisition: is there a difference? *Language Learning, 45*(3), 419-465. https://doi.org/10.1111/j.1467-1770.1995.tb00448.x

Kroll, J. F., Bogulski, C. A., & McClain, R. (2012). Psycholinguistic perspectives on second language learning and bilingualism: The course and consequence of cross-language competition. *Linguistic Approaches to Bilingualism, 2*, 1–24. 10.1075/lab.2.1.01kro

López Morales, H. (1973). *Disponibilidad léxica de los escolares de San Juan*. San Juan de Puerto Rico: Universidad de Puerto Rico.

López Morales, H. L. (2014). Lexical availability studies. In R. M. Jiménez Catalán (Ed.), *Lexical Availability in English and Spanish as a Second Language* (pp. 1-11). Springer.

Lorenz, E., Rahbari, S., Schackow, U., & Siemund, P. (2020). Does bilingualism correlate with or predict higher proficiency in L3 English? A contrastive study of monolingual and bilingual learners. *Journal of Multilingual Theories and Practices, 1*(2), 185-217. 10.1558/jmtp.15517

Mackey, W. F. (1971). *Le vocabulaire disponible du français*. Didier.

Michéa, R. (1953). Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage. *Les Langues Modernes, 47*, 338–344.

Ministerio de Educación, Formación Profesional y Deportes (2025). Enseñanzas no universitarias. Alumnado matriculado. Curso 2023-2024. Datos avance. Retrieved from: https://www.educacionfpydeportes.gob.es/servicios-al-ciudadano/estadisticas/no-universitaria/alumnado/matriculado/2023-2024-rd.html

Molnar, T. (2008). Second language versus third language vocabulary acquisition: A comparison of the English lexical competence of monolingual and bilingual students. *Toronto Working Papers in Linguistics (TWPL), 33*, 1-16.

Muñoz, C. (2020). Cognate recognition by young multilingual language learners: The role of age and exposure. In C. Bardel & L. Sánchez (Eds.), *Third language acquisition: Age, proficiency and multilingualism* (pp. 145–165). Language Science Press. 10.5281/zenodo.4138743

Otwinowska, A. (2016). *Cognate Vocabulary in Language Acquisition and Use*. Multilingual Matters. https://doi.org/10.21832/9781783094394

Samper Hernández, M., & Jiménez Catalán, R. M. (2014). Researching Lexical Availability in L2: Some Methodological Issues. In R. M. Jiménez Catalán (Ed.), *Lexical Availability in English and Spanish as a Second Language* (pp. 189-205). Springer. https://doi.org/10.1007/978-94-007-7158-1_11

Sánchez-Saus Laserna, M. S. (2024). Lexical Availability and Foreign Language Teaching: Main Contributions of a Growing Field. *Theory and Practice of Second Language Acquisition, 1*(10), 1-20.

Siemund, P., & Lechner, S. (2015). Transfer effects in the acquisition of English as an additional language by bilingual children in Germany. *Transfer effects in multilingual language development, 4*(1), 147. 10.1075/hsld.4.07sie

Siemund, P., & Mueller, J. T. (2020). Are Multilinguals the Better Academic ELF Users? In A. Mauranen & S. Vetchinnikova (Eds.) *Language Change: The Impact of English as a Lingua Franca* (pp. 234-266). Cambridge University Press. https://doi.org/10.1017/9781108675000.013

van Ginkel, C. I., & van der Linden, E. H. (1996). Word associations in foreign language learning and foreign language loss. In K. Sajavaara & C. Fairweather (Eds.), *Approaches to second language acquisition* (pp. 25-35). University of Jyväskylä.

First version received: January, 2025
Final version accepted: September, 2025