Theoretical Exploration of the Lexical Availability Task as a Tool for the Selection of Vocabulary to be Taught in Class

Leah Geoghegan*
Universidad de la Rioja, Spain

María Pilar Agustín Llach
Universidad de la Rioja, Spain

Abstract
Since its conception in the 1950s, the Lexical Availability Task (LAT) has steadily gained prominence as an assessment tool in language teaching and learning, perhaps due to the fact that it not only provides a useful indication of the lexical resources which are available to learners, but it also obtains this rich set of data in a rather economical way (Jiménez Catalán & Fitzpatrick, 2014). The LAT requires participants to write as many words as come to mind in response to a specific stimulus or cue-word within a time limit of 2 minutes. The responses obtained provide a lexical corpus of available words for that cue-word or centre of interest. While previous research has highlighted the value of using lexical availability (LA) as a tool for selecting target foreign language vocabulary (Ávila Muñoz, 2017), this article argues that this tool should not be used in isolation, but rather should be implemented appropriately alongside a number of other factors in order to best determine target vocabulary. A theoretical exploration of the LAT as a tool for the selection of vocabulary to be taught in class is thus provided, taking into consideration a number of key factors: the most readily available words in each field of natives of the age group, other similar groups, words from the school context, equivalent words of highly available words from L1 and the teaching context. Suggestions are made for incorporating LA as a tool for vocabulary selection in classrooms.

1 Introduction

In a review of recent research on vocabulary acquisition, teaching and assessment, Benigno and De Jong (2017) found that, despite rapid advancements in the field, there remains little consensus as to which and how many words are required to communicate effectively as proficiency increases. While generally there is agreement on what grammatical structures should be taught at different levels, the same cannot be said

---

*Corresponding author. Email: legeogw@unirioja.es

for teaching vocabulary: syllabi include word lists, but it is unlikely that the list created for one group of students will be the same as for another group at the same level (Cristina, 2010). This is perhaps because, as Richards pointed out in the early 1970s, "methods of selecting items for foreign language testing may reflect differing views of language, and of language teaching" (Richards, 1970, p. 87).

Vocabulary has often been selected based on its frequency in the language, how useful it is, and how difficult it is (Laufer & Nation, 2012). However, even with assistance from technology, which allows us to easily acquire authentic samples of how and how often words are used, approaches based on frequency and coverage often fall short of providing an optimal solution to vocabulary selection (Cristina, 2010). This is often due to the fact that words rarely have a single meaning; as Laufer and Nation (2012) point out, homonyms and multiword units cannot be distinguished and accounted for in word counting programs. In such instances, qualitative judgement by human intuition may need to be incorporated, something which cannot easily be replicated by modern computer programs (Martinez & Schmitt, 2012). This judgement more often than not comes from teachers and material writers, who can make informative decisions of word properties beyond frequency using their expertise of language learning textbooks and materials (He & Godfroid, 2019). However, other researchers have also stressed that students should also take part in the process, either by training them to analyse new vocabulary items (Barker, 2007) or having them select target vocabulary themselves (Choi & Ma, 2015).

Here, we deal with another instrument that might allow us to overcome the trichotomy: frequency-usefulness-difficulty, since it is based on a completely different dimension, namely, availability. Below, we offer a detailed account of the construct of lexical availability and the tool to delve into it: the lexical availability task. Before that, we will review the three criteria traditionally used to study and select vocabulary.

2 Revising Vocabulary Selection Criteria: Frequency, Usefulness, and Difficulty

Firstly, regarding frequency, there is a clear benefit in focusing efforts on learning the 2000 most frequent words in a language, as they have been shown to make up at least 80 percent of both spoken and written text (Read, 2004). Frequency is generally determined based on word frequency lists, which analyse the most common words in written or spoken corpora. Some of the most influential lists have been West's (1953) General Service List (2000 headwords based on the frequency figures of a 5,000,000-word written corpus) and, more recently, Nation's (2012) British National Corpus (BNC)/Corpus of Contemporary American English (COCA) word family lists (29-word family lists representing both British and North American English and based partly on spoken corpora). Regarding the latter, Webb and Sasao (2013) have remarked that these new lists are more representative of modern English and offer a better indication of how vocabulary is used by native speakers today. Secondly, the criterion of usefulness aims to address the understanding that some words are useful, even when they are not particularly frequent in the language, to particular learners who have specific needs (Laufer & Nation, 2012). An example of this is the Academic Word List created by Coxhead (2000), which provides 570 words frequent in academic discourse which are not already part of the General Service List. Notably, usefulness here adheres to the notion of frequency, but analyses words in specific contexts so as to provide those which are most frequent, and consequently most useful to learners with specific lexical requirements. Finally, the notion of difficulty, or learnability, has also been highlighted as an important factor in vocabulary selection, and may be interlingual or intralingual (He & Godfroid, 2019). While the former is based on interactions between the L1 and L2, the latter is concerned with interactions between known words and new words in the L2. Interlingual learning difficulty is thus when there is or is not formal or semantic congruence between the mother tongue and the target language (Lado, 1972). In other words, for instance cognates, which are easy to learn, can enhance learners' vocabulary in a relatively simple way (Laufer & Nation, 2012).

3 Lexical Availability and the Lexical Availability Task

Recently, lexical availability (LA) has also been highlighted as a valuable tool for selecting target foreign language vocabulary (Ávila Muñoz, 2017). LA refers to "the words that people have in their minds and that emerge in response to cue words that stand for domains closely related to daily life such as 'Food and drink', 'Animals', 'Politics', or 'Poverty'" (Jiménez Catalán, 2014, p. v). From a theoretical and a methodological point of view, LA offers the possibility of creating thematic vocabulary lists covering wide arrays of topics from everyday topics to specialized knowledge areas. Its origins date back to France in the 1950s, where the first LA studies were carried out with the aim of selecting the words which should be taught to the federation of territories known as *Union Française* (López Morales, 2014). Originally, word selection was generally made based on the most frequent words in the language, which were assumed to be the most practical and useful. It was soon realised, however, that many commonly known words by native speakers, such as *t-shirt* or *teeth*, did not feature in these lists (Payne, 2016). Evidently, such words, despite their lack of frequency, could not be considered unimportant in daily language. LA consequently emerged as a way to make up for this issue in word frequency lists, placing greater emphasis on functional native-speaker vocabulary which may not necessarily appear frequently in the language. Michéa (1953), thus made a distinction between "frequent words" and "available words", a turning point which led to the first LA studies by Gougenheim et al. (1964) in France, Dimitrijévic (1969) in Scotland, Mackey (1971) in Canada, and López Morales (1973) in Puerto Rico. As is highlighted by Fernández Smith, Sánchez-Saus Laserna and Escoriza Morerar (2012), this work made important contributions to both the theoretical and methodological systematisation of LA research. So, indeed, LA studies have a pedagogical nature in origin, since their first and fundamental aim was to select the lexical items to be taught in the foreign language classroom (Gougenheim et al., 1964; Michéa, 1953).

Methodological homogeneity has come with the development of the lexical availability task (LAT), which is the primary data collection instrument used in LA research. It is generally administered as a paper-and-pencil questionnaire, whereby participants are presented with prompts (also referred to as "cue words", "centres of interest", or "semantic categories") and asked to write down any words that come to mind in response (Samper Hernández & Jiménez Catalán, 2014). Each selected prompt is presented on a different page with numbered lines, and participants are allowed two minutes for each prompt. Although the number of prompts used varies from study to study, most draw from the 16 categories proposed by Gougenheim et al. (1964) in the first LATs. As Šifrar Kalan (2015, p. 197) notes, these include: *Parts of the human body, Clothing, Parts of the house, House furniture, Food and drink, Objects on the table for the meal, The kitchen and its utensils, School furniture and materials, Heating and lighting, The city, The countryside, Means of transport, Farm and garden work, Animals, Games and entertainment,* and *Jobs and professions.* By using these same, or similar, prompts and following the same methodological approach, LA researchers have benefitted from cross-comparison of results even in a wide array of contexts. This systematicity has granted the LAT the status of a reliable task with consistent results, both in the L1 and the English as a Foreign Language (EFL) available lexicon (Canga Alonso, 2017). However, the prompts, semantic categories or lexical fields that can serve as a stimulus to retrieve lexical items is open-ended, which offers great potential of adaptability to different contexts, purposes, and informants. For instance, researchers working in Chile in the mathematical field have used the LAT to identify the vocabulary of their students, prospective mathematics teachers, with the prompts: *algebra, calculus, geometry,* and so on (e.g., Rojas Díaz et al., 2018). Similarly, Zambrano Matamala has focused on tapping into the vocabulary of pedagogy in future teachers looking at the fields *learning, motivation to learn, learning planning,* among others (e.g., Zambrano Matamala, 2021).

To date, the LAT has been approached from a number of different perspectives, including cognitive, related to lexical access and categorization; sociolinguistic, addressing different individual variables such as bilingualism, socioeconomic background and geographical origin; and pedagogical, related to language learning and teaching. Within the field of language learning, Fernández Smith et al. (2012,

p. 15) have highlighted the numerous advantages which LA brings with it: it allows us to analyse the vocabulary which students know, detect potential differences between learners with different mother tongues, observe learning phases, compare students' LA with that of native speakers, and provide editors with teaching material that is suitable for the different levels of teaching.

3.1. Lexical Availability and Frequency, Usefulness and Difficulty

Within a language teaching context, LA has clear potential in deciding which vocabulary is to be taught. However, it is also important to consider other factors alongside LA, in keeping with Laufer and Nation's (2012) criteria of frequency, usefulness and difficulty in vocabulary selection, mentioned above.

We suggest that, while frequency is evidently an important factor to consider in the vocabulary selection process, it could be best put to use alongside LA and the LA measure of LA index. This index, which refers to the response words, is based on a mathematical formula and provides a measure of the frequency of the lexical unit and its position in the list, whether it was produced first, second, and so on, allowing discrimination within the ranges of units of equal frequency (López Morales, 2014). This would allow us to differentiate between different words of equal frequency, providing a potentially more useful measure which also takes the availability of these commonly used words into consideration. We propose that this should be done by analysing the most readably available words in each field of natives of the age group or other similar groups. This would enable us to best determine the most appropriate vocabulary for students of different ages and cognitive abilities. In a series of studies with a slightly different focus, Milton and colleagues (Roghani & Milton, 2017; Milton & Alexiou, 2020) propose the use of a category generation task, similar to our LAT, to assess learners' vocabulary size. Responses to the category generation task are compared to frequency lists and, accordingly, productive vocabulary sizes based on those frequency lists are calculated. The procedure, however, is different from the comparison proposed in the present study. Additionally, we believe that LA offers a particularly beneficial way to select the vocabulary which is most useful to students. In particular, we propose that LA be used as a way to determine the most readily available words from the school context and the teaching context, so as to provide the most practical vocabulary. This approach would also be especially useful in Content and Language Integrated Learning (CLIL) settings, whereby content is taught through the vehicle of the foreign language. Canga Alonso (2017), for example, has highlighted the need in LA research to focus on prompts which are related to the subject at hand, e.g., using the prompt *Parts of the Body* with students who have studied Natural Sciences and Biology through the target language. This would not only be useful in assessing L2 speakers on their vocabulary acquisition in CLIL classes, but also as a means to determine the most appropriate and useful target vocabulary to be taught in these classes.

To the effect of difficulty, we suggest that equivalent, highly available words from the students' L1 also be the focus in vocabulary. This would allow students to draw on their linguistic knowledge in their first language, increasing their confidence and motivation in the target language (Laufer & Nation, 2012). An additional benefit of the LAT is that it provides researchers with different lexical measures that give a multi-perspective picture of the learner's lexical productions. First, as discussed above, it allows us to obtain a LA index for each lexical unit under analysis. Additionally, another 4 indexes are generally used in LA studies: total number of lexicons, related to the number of participants; mean response length, which refers to the mean number of tokens produced by the sample; number of types or of different words that appear in the responses sample; and cohesion index, which marks the number of coincidental responses within the respondents per each semantic field or prompt. It is, thus, an index pertaining to the prompt or centre of interest.

3.2. LA and Vocabulary Selection for the Classroom

If studies on LA have mainly concentrated on finding out the available lexicons of native speakers for different purposes, whether didactic, sociolinguistic or cognitive, the appearance of research studies

looking into FL learners' lexical productions have given the LA trend a wholly new perspective. This has given a special significance to LA findings to be applied in the FL classroom. Two main lines of research can thus be distinguished: research conducted with native speakers and studies that look into foreign language learners' available lexicons. The first of the research trends has identified the available thematic vocabulary of native speakers with different characteristics such as age, gender, professional background, or geographical origin. This has allowed the creation of objective vocabulary lists which may be used in the foreign language classroom. Accordingly, vocabulary to be taught in the FL class is selected under the premise that the vocabulary available to native speakers of similar characteristics is the desirable one. In this sense, Benítez Pérez (2009) states that the vocabulary that FL learners have to learn must be a faithful reflection of the vocabulary used by the native community. Bartol Hernández (2010) also highlights how comparison of native lists with learner lists allows us to detect errors in the learning process, in particular those due to L1 influence.

Different empirical studies have set to the task of elaborating vocabulary lists for the classroom based on native responses to the LAT (see Tomé Cornejo, 2015, p. 340 for a thorough account of these). For instance, Izquierdo Gil (2005) elaborated lists of basic vocabulary for French adolescent learners of Spanish as a foreign language. To this purpose, she selected the 20 most available words for the traditional 16 centres of interest of several available lexicons in Spain and other Spanish speaking countries. Following a slightly different methodology in setting the threshold for word selection (words lexical availability over 0.1), Bombarelli (2005) also created her own list of basic vocabulary for Spanish FL learners based on native accounts. She found that 128 words she had selected did not appear in the A1/A2 materials she analysed. Wingeyer (2014), for his part, explores the native responses to other innovative fields such as *the town*, *the countryside*, *popular myths*, *insults*, or *things to do to have fun at weekend nights*, to thus build a proposal of vocabulary teaching for adolescents. Hidalgo Gallardo (2017) goes one step further and establishes a threefold comparison among 1) native available lexicons, 2) teaching materials, and 3) learners' lexicons, with the aim of exploring learners' lexical gaps and finding ways to remedy them. The second main trend pointed out above refers to studies which look into the available lexicon of FL learners. These studies have different objectives (Tomé Cornejo, 2015, p. 338) such as a) comparing native and learner lexical availability (e.g., Carcedo Ginzález, 2002; Herreros Marcilla, 2014), b) comparing learners' different lexicons (e.g., Gallego Gallego, 2014; Šifrar Kalan, 2009), c) comparing learners' available lexicon with their actual production of vocabulary in communicative tasks (e.g., Frey Pereyra, 2007), and d) comparing lexicons with teaching materials (e.g., Fernández-Merino Gutiérrez, 2011; Jiménez Berrio, 2012). Determining the available lexicons of FL learners can have not only the descriptive objective of exploring lexicon organization (e.g., Ferreira & Echeverría, 2010), and designing teaching approaches accordingly, but it can also help appraise learners' lexical knowledge as a kind of vocabulary test (Hidalgo Gallardo, 2017; Palapanidi, 2019), figuring out learners' communicability within their group (Salcedo & del Valle, 2013) or exploring categorization behaviours of learners (Hernández-Muñoz, Izura & Ellis, 2006).

4 The Study

In order to illustrate the usefulness of the LAT as a tool for the selection of vocabulary to be taught in class, this section provides an example of how LA can be used with pedagogical purposes. It aims in particular to demonstrate three key strengths of using LA to select vocabulary to be taught in the classroom in relation to the criteria of frequency, usefulness and difficulty. Firstly, LA can enable us to home in on the most useful vocabulary in frequency lists, by combining the most readily available and most frequent vocabulary. Secondly, LA can be used to better determine target vocabulary in certain lexical domains, highlighting vocabulary which may not appear on the typical frequency lists used. Finally, LA provides a convenient way to verify equivalent, highly available words from the students' L1 which can be used to suitably adjust the difficulty of the selected vocabulary. To this effect, we provide

an analysis of two LA prompts alongside a target vocabulary list, in order to demonstrate the convergence and divergence of a sample group of students' LA and the vocabulary they are expected to have acquired within a particular domain, and how similar this vocabulary is to the equivalent in their L1.

4.1. Instruments

The LAT that is used here contains the tokens produced by a group of 99 Spanish learners of English as a foreign language in 10th grade of compulsory secondary education (age 14-15). Informants had to complete a LAT responding to several prompts such as town, countryside, hobbies, professions and food and drink. Prompts were selected from those traditionally used so as to allow for comparison with other studies. The students had learned English as a foreign language at school since kindergarten as part of the school curriculum and had reached an A2 level of English on the Common European Framework of Reference (CEFR). The target vocabulary list that has been chosen for the current illustration is thus the A2 Key and A2 Key for Schools Vocabulary List developed by Cambridge English. This list draws on vocabulary from the Council of Europe's Waystage 1990 specification (van Ek & Trim, 1991) and other vocabulary, including vocabulary which corpora have shown to be highly frequent. It covers both receptive and productive vocabulary which is appropriate to the A2 level on the CEFR. In addition to providing an alphabetical list of target vocabulary, this list provides a range of topic lists, which include vocabulary related to a specific theme. For the purposes of this analysis, two of these topics were chosen, namely Food and Drink and Town and City. These were selected given that these are lexical domains which were assessed in the abovementioned LAT and provide one area which is relatively closed (Food and Drink) and one which is relatively open (Town and City). In addition, Food and Drink contained a total of 105 target words, while Town included just 20. While it can be deemed relatively easy to decide which words or lexical items can be categorized under food and drink and thus, coincidences among responses are to be expected; it is much harder to delimit the category town [and city], so fewer coincidences and more heterogeneity among responses can be counted on. For the purposes of the analysis, multiword units were merged into one unit, as is common practice in LA, e.g., "ice cream" to "icecream", "main course" to "maincourse", "bus station" to "busstation", etc.

4.2. Procedure for Analysis

The analysis was carried out using Dispogen, a program created by Echeverria, Urzúa and Figueroa (2005) which allows the calculation of LA values for each word in response to each prompt under analysis (Samper Hernández & Jiménez Catalán, 2014). Additionally, Dispogen provides relative frequency data, convergence and divergence data, and allows for individual and group comparisons of data. From the information provided by Dispogen, two ratios were calculated as part of the analysis: firstly, the number of target words (i.e., words that appear in the A2 Key and A2 Key for Schools Vocabulary List developed by Cambridge English) produced in the LAT out of the total number of words on the target vocabulary list; and secondly, the number of target words produced in the LAT out of the total number of words produced in the LAT. This gave us a ratio for a) the target words produced / total target words (TWP/TW) and for b) the target words produced / words produced for the two prompts under analysis (TWP/WP) on a scale of 0 to 1. For example, the target vocabulary list Town contained 20 items. A student who produced 7 words that appear on this list would receive a TWP/TW index of 7/20 = 0.35. This student produced 22 words in total, and so would receive a TWP/WP index of 7/22 = 0.31. It should also be noted that when calculating the number of target words produced, multiword units containing words that featured on the target list were also included. For example, the word "market" appeared on the target list, but "supermarket" did not. However, given the predominance of the word "supermarket" in the words produced in the LAT, it was included as a target word in the analysis. Indices were calculated both for each individual participant and for the entire group.

5 Results

As previously indicated, the aim of the analysis was threefold: to demonstrate how LA allows us to determine the most useful vocabulary in frequency lists, how it can be used to better select target vocabulary in certain lexical domains, and how it may be used to determine equivalent, highly available words from the students' L1. The following three sections provide the results of the analysis, dealing with each aim in turn.

5.1. Frequency: The most readily available vocabulary in frequency lists

One of the main benefits of the LAT is that it allows us to determine the most frequent available words that are produced for each prompt. By analysing LA alongside frequency lists, we can determine the vocabulary that is not only frequent in the target language, but also that which is the most readily available. The results of the analysis indicated there were no target words on either list that did not appear in the words produced by the participants. In other words, participants do produce vocabulary which has been suggested to be the most frequent and which is to be targeted in EFL classes. On closer examination, we can then determine which available vocabulary is the most common. Table 1 provides an overview of words which appeared in the LAT a minimum of ten times, in order of frequency and LA Index. For example, in the prompt *Town,* the token "car" was most common, having been produced by a total of 55 participants. In addition, it had the highest LA index, indicating that the word was produced not only by numerous participants, but also that it appeared earlier in the words produced, i.e., it was one of the first words that came to mind.

Table 1
Words Produced in the LAT at Least Ten Times

Town			Food and Drink		
Token	**Total**	**LA Index**	**Token**	**Total**	**LA Index**
car	55	0.442	water	84	0.777
school	62	0.409	apple	62	0.514
house	52	0.398	tomato	55	0.447
park	58	0.389	hamburger	50	0.447
street	42	0.336	fish	43	0.359
shop	42	0.300	banana	43	0.351
hospital	38	0.265	orange	43	0.345
shoppingcentre	33	0.235	potato	42	0.343
people	34	0.233	pizza	40	0.341
city	22	0.203	milk	39	0.319
bus	27	0.200	meat	35	0.297
building	23	0.197	spaghetti	33	0.291
cinema	25	0.185	salad	32	0.279
highschool	31	0.176	bread	29	0.240
university	28	0.161	chip	27	0.233
supermarket	26	0.158	rice	27	0.233
tree	21	0.135	chicken	27	0.221
restaurant	22	0.128	lettuce	25	0.204
road	18	0.126	orangejuice	24	0.203
busstation	19	0.123	watermelon	23	0.183
townhall	17	0.119	strawberry	23	0.182
skyscraper	12	0.093	juice	22	0.181
dog	15	0.092	vegetable	22	0.180
library	15	0.089	wine	22	0.179

|					
trainstation	13	0.078	macaroni	21	0.178
bakery	13	0.077	coffee	21	0.168
village	9	0.070	steak	20	0.163
bank	11	0.070	lemon	20	0.155
train	11	0.065	pear	20	0.154
underground	10	0.064	chocolate	19	0.150
theatre	8	0.062	cake	18	0.145
swimmingpool	12	0.061	cheese	18	0.137
taxi	10	0.057	soup	17	0.134
			sausage	17	0.133
			carrot	17	0.132
			melon	17	0.131
			beer	16	0.125
			icecream	16	0.121
			egg	16	0.118
			pineapple	14	0.111
			fruit	14	0.109
			onion	13	0.106
			bean	13	0.103
			sugar	12	0.085
			pasta	11	0.087
			tea	11	0.084
			biscuit	10	0.078
			sandwich	10	0.078

As shown, use of the LAT allows us to complement frequency measures. This could be done with the target group of students in their native language at the beginning of the course, so as to establish which vocabulary they have available to them. These results could then be used to determine a more tailored vocabulary list, allowing teachers to select vocabulary which is not only frequent, but also which is highly available in a given language.

5.2. Usefulness: Indices for target words and words produced

A second benefit of the LAT is that, by focusing on particular lexical domains, it allows us to determine the vocabulary which will be most useful to the students in certain situations. This lets us select vocabulary that is highly available within a certain area, but which does not appear on frequency lists. The second part of the analysis thus aimed to show how LA can be used to determine the most appropriate vocabulary within a certain lexical domain, and highlight how the vocabulary in many areas may be particularly useful to the student, even though it may be disregarded when frequency alone is used to select vocabulary. To this effect, the following results demonstrate the relationship between the target words produced and the total target words or total words produced for the two prompts under analysis.

Figure 1 shows the average TWP/TW and TWP/WP for each prompt. As shown, in the case of the prompt Town, indices for TWP/TW and TWP/WP were relatively similar: 0.13 and 0.18 respectively. However, for the prompt Food and Drink, greater disparity was found, with a TWP/TW index of 0.11 and a TWP/WP index of 0.82. In addition, the TWP/TW index was relatively similar in both prompts, while greater differences were found in the TWP/WP index of each prompt.

Figure 1
Average TWP/TW and TWP/WP
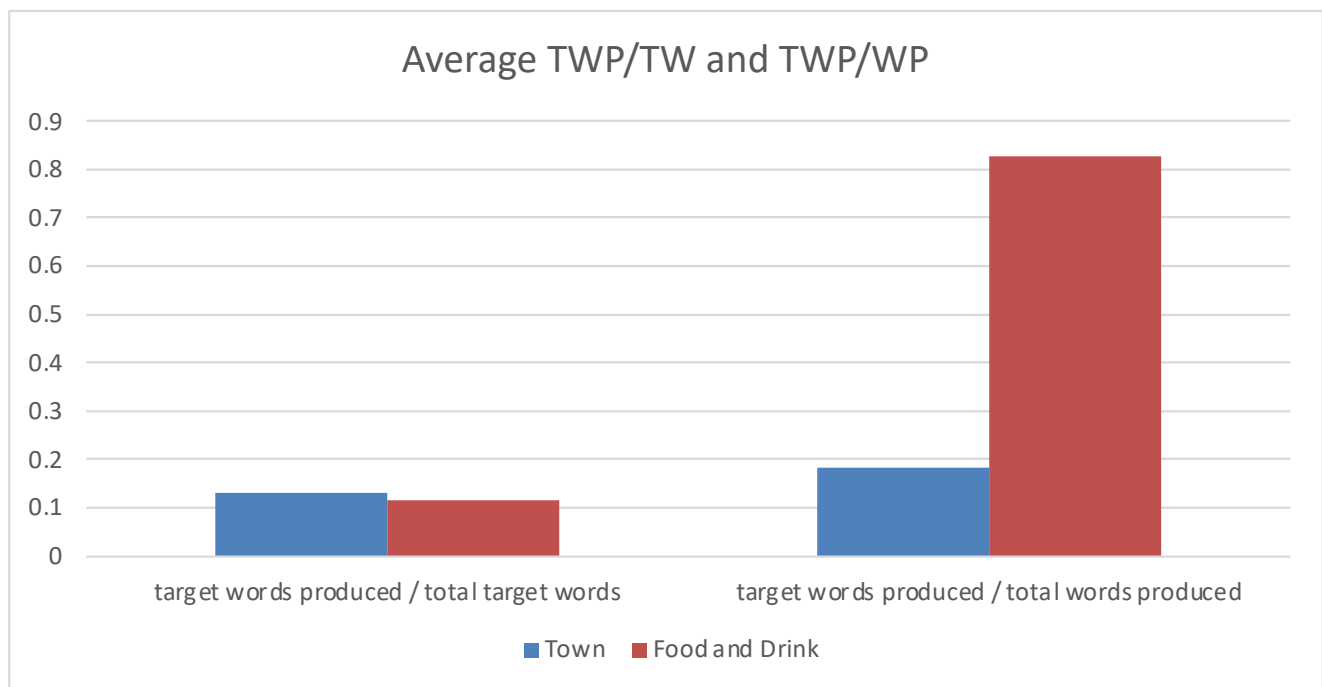
This indicates that for both prompts, the ratio of target words produced out of the total number of target words was similar, while differences were found in the ratio of target words produced out of the total number of words produced. In the case of *Town*, target words made up rather few of the total words produced, indicating that participants produced other words which did not appear on the target list. On the other hand, in *Food and Drink*, target words made up a large proportion of the words produced by participants. This can be explained by the fact that while *Town* is a relatively open domain, leading to greater variety in the responses, *Food and Drink* is relatively closed, resulting in the production of words that appear on the specific target word list. These differences offer interesting insights for the selection of target vocabulary in the classroom. In the case of open domains such as *Town*, use of the LAT allows us to determine the vocabulary that is most readily available to students, which evidently is not necessarily vocabulary that appears on the frequency lists which are so often used. This could clearly be applied to a wide range of other prompts, e.g., those relating to CLIL subjects or other areas which may be deemed to be useful to the students. This would provide teachers with a convenient way to determine the most useful vocabulary within the given domain and to ensure that it is not excluded simply because it does not appear on the frequency lists.

5.3. Difficulty: Equivalent, highly available words from the students' L1

A final benefit of the LAT is that it provides a convenient way to determine which vocabulary may be the easiest to acquire. If this is known, teachers can better allocate time spent learning vocabulary, giving more attention to more challenging vocabulary which students may not acquire by themselves, or incorporating vocabulary that may be easier in order to increase students' motivation. Table 2 provides a brief illustration of this application using the data discussed above, providing the ten most common words in each prompt in English alongside their Spanish translation and highlighting the presence of cognates.

Table 2
Cognates in the Ten Most Common Words

Town		Food and Drink	
English	**Spanish**	**English**	**Spanish**
car	coche	water	agua
school	colegio	apple	manzana
house	casa	tomato	tomate
park	parque	hamburger	hamburguesa
street	calle	fish	pescado
shop	tienda	banana	banana
hospital	hospital	orange	naranja
shoppingcentre	centrocomercial	potato	patata
people	gente	pizza	pizza
city	ciudad	milk	leche

As shown, results indicate the presence of a number of Spanish-English cognates, including "park" and "hospital" in the prompt *Town* and "tomato", "hamburger", "banana" and "pizza" in the prompt *Food and Drink*. As mentioned above, if teachers incorporated the use of a LAT in the students' L1 at the beginning of a course, this would provide a convenient way to determine the most readily available vocabulary in relation to a number of different areas. As shown here, the results could also be analysed in terms of cognates, so as to anticipate areas of difficulty that the students may encounter. For example, cognates in the prompt *Town* include "park" and "hospital", which will evidently be easier for students to acquire. On the other hand, words such as "house" or "street" which bear no resemblance to their Spanish counterpart will clearly require more effort to learn.

6 Discussion and Conclusion

In this illustrative study, we have tried to show how the LAT can be a useful tool to tap into students' vocabulary and appraise their lexical competence against a previously compiled list of vocabulary appropriate for A2 level EFL students, which is recommended for the EFL classroom. As is noted in the A2 Key and A2 Key for Schools Vocabulary List, the list is not an exhaustive list of all words which appear on the KET and KET for Schools question papers, and so candidates should not limit their vocabulary study to the list alone. Our exploration of students' LA has provided us with a way to complement such a vocabulary list, thereby filling in the gaps that students may be missing in their lexicon. As noted, this can be done with a particular focus on three issues: frequency, usefulness and difficulty.

Firstly, regarding frequency, we have highlighted how LA can be incorporated and used alongside frequency lists to complement these measures. As demonstrated in the analysis, this can be done by analysing the LA index of the results, thereby selecting vocabulary which is not only produced by numerous participants but which appears earlier in the words produced. Teaching words with high lexical availability indexes covers not only lexico-statistical criteria, but also communicability criteria, as well as prototypicality, since these words are key elements in establishing successful communication (e.g., Salcedo & del Valle, 2013). Given the simplicity and lack of time needed to carry out the task, we suggest that the task could very manageably be incorporated into the FL learners' curriculum, assessing students' vocabulary at the start of the term to better cater to their vocabulary needs. In a very comprehensive review paper, Paredes García (2015) highlights that the LAT can have several applications in the FL classroom, starting with the elaboration of native vocabulary lists within the thematic areas purported by the CEFR (2001) pertaining to the personal, public and professional or educative domains. These native

lists provide the framework against which to compare learners' needs and accomplishments (e.g., Bartol Hernández, 2010; Paredes García, 2014), and therefore they could be the basis for teaching materials and programs. Additionally, they can serve as reference to evaluate previously existing programs and materials, as well as learners' lexical competence (e.g., Hidalgo Gallardo 2017). Sánchez-Saus Laserna (2011) also compares native and learners' vocabulary lists to look into the variables that influence the process of vocabulary acquisition in the FL, such as age, L1 background or L2 proficiency.

Secondly, regarding usefulness, the analysis provided an example of how TWP/TW and TWP/WP ratios can be used to assess LA data. Using the examples provided, it is clear that not all lexical domains are created equal: some prompts are more likely to see students produce vocabulary which is commonly found on typical frequency lists. However, we have seen that the analysis of other domains, such as *town*, allows us to determine other vocabulary which may not be targeted when focusing solely on frequency lists. This information is of vital importance when students are required to learn context-specific vocabulary, such as that related to CLIL classes. For example, vocabulary related to Natural Science may not be particularly frequent in the language at hand but is inevitably important to students who are studying this subject through their target language. By incorporating LA, teachers can better target the thematic vocabulary which is most useful to their students, and rather than selecting this in an arbitrary manner, LA can provide a convenient and time efficient way to choose what words will be most useful.

Finally, previous research has stressed the important role that cognates play in language learning, allowing learners to increase their vocabulary in a rather simple way (Laufer & Nation, 2012). It is important that teachers encourage the learning and teaching of cognates because they can be an important head start in the FL, are easy to learn, and can increase learners' vocabulary to a great extent. Similarly, teaching students to use cognates can be another interesting strategy to learn vocabulary. Our analysis has shown how cognate analysis can be incorporated into LA studies as a means to further refine selected vocabulary and thus allow teachers to make educated decisions regarding the relative difficulty of the chosen vocabulary.

Our study has shown a great overlap between learners' responses and their expected word knowledge. This might point to adequate teaching programs. Further research should focus on the potential of the LAT as a word association task of the semantic fluency type which can help gain insights into the development of vocabulary in bilingual and multilingual learners. For instance, it can provide information related to how words are linked, how they are stored in the mind and what strategies learners use to retrieve them, such as semantic relatedness, e.g., fruit-vegetable, categorial membership, e.g., fruit-apple, or formal similarity, e.g., carpark-carwash. Additionally, the LAT can prove particularly helpful for topic areas where corpora do not exist or where corpora are likely to be too specialized, for instance, a corpus of Natural Science texts to be used in schools.

References

Ávila Muñoz, A. M. (2017). The available lexicon: A tool for selecting appropriate vocabulary to teach a foreign language. *Iranian Journal of Language Teaching Research*, *5*(1), 71-91.

Barker, D. (2007). A personalized approach to analyzing 'cost'and 'benefit'in vocabulary selection. *System*, *35*(4), 523-533.

Bartol Hernández, J. A. (2010). Disponibilidad léxica y selección del vocabulario. In R. M. Castañer Martín and V. Lagüéns Gracia (Eds.), *De moneda nunca usada. Estudios dedicados a José M.ª Enguita Utrilla* (pp. 85-108). Institución Fernando El Católico.

Benigno, V., & de Jong, J. (2017). Developing the GSE Vocabulary. *Global scale of English research series*. Retrieved from https://prodengcom.s3.amazonaws.com/GSE-Vocab.pdf

Benítez Pérez, P. (2009). ¿Qué vocabulario hay que enseñar en las clases de español como lengua extranjera? In L. Miquel and N. Sans (Eds.), *Monográficos MarcoELE: Didáctica del español como lengua extranjera. Expolingua 1994* (pp. 9-12). Fundación Actilibre.

Bombarelli, Á. (2005). La disponibilidad léxica como herramienta didáctica: una propuesta de selección del vocabulario para un nivel umbral de ELE. [Master's thesis, Universidad de Salamanca].

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

Carcedo Ginzález, A. (2002). Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición). Universidad de Turku.

Choi, M. L., & Ma, Q. (2015). Realising personalised vocabulary learning in the Hong Kong context via a personalised curriculum featuring 'student-selected vocabulary'. *Language and Education*, *29*(1), 62-78.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238.

Cristina, A. (2010). Vocabulary and language teaching. *Universitatea din Oradea Facultatea de Stiinte Economice. Jel Code Y*, *8*, 170-174.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior Research Methods*, *51*(3), 987-1006.

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

Echeverría, M., Urzúa, P., & Figueroa, I. (2005). *Dispogen II. Programa computacional para el análisis de la disponibilidad léxica.* Concepción, Chile, Universidad de Concepción.

Fernández Smith, G., Sánchez-Saus Laserna, M., & Escoriza Morerar, L. (2012). Studies on lexical availability: The current situation and some future prospects. In B. Eizaga Rebollar (Ed.), *Studies in Linguistics and Cognition* (pp. 35-57). Peter Lang.

Fernández-Merino Gutiérrez, P. V. (2011). Disponibilidad léxica de inmigrantes: Propuesta para una necesidad. *Language and Migration*, *3*(2), 83-105.

Ferreira, R., & Echeverría, M. S. (2010). Redes semánticas en el léxico disponible de inglés L1 e inglés LE. *Onomázein*, (21), 133-153.

Frey Pereyra, M. L. H. (2007). Disponibilidad léxica y escritura del ELE: propuesta de comparación de dos corpus. *Interlingüística*, (17), 366-373.

Gallego Gallego, D. J. (2014). Léxico disponible de estudiantes de español como lengua extranjera en la Comunidad de Madrid. [Doctoral dissertation, Universidad de Alcalá].

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, France: Didier.

He, X., & Godfroid, A. (2019). Choosing words to teach: A novel method for vocabulary selection and its practical application. *TESOL Quarterly*, *53*(2), 348-371.

Hernández-Muñoz, N., Izura, C., & Ellis, A. W. (2006). Cognitive aspects of lexical availability. *European Journal of Cognitive Psychology*, *18*(5), 730-755.

Herreros Marcilla, M. (2014). La disponibilidad léxica como propuesta de investigación docente. Análisis en alumnos inmigrantes no hispanohablantes. Paper presented at the *XXV Congreso Internacional de ASELE,* 17-20 September 2014, Universidad Carlos III, Madrid. Retrieved from https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/25/25_0459.pdf

Hidalgo Gallardo, M. (2017). La Disponibilidad Léxica como Herramienta de Evaluación Interventiva del Vocabulario en ELE: Aplicación en un Contexto Universitario Sinohablante. *SinoELE*, (16), 43-63.

Izquierdo Gil, M. (2005). *La selección de léxico en la enseñanza del español como lengua extranjera. Su aplicación al nivel elemental en estudiantes francófonos* [Doctoral dissertation, Universitat de València].

Jiménez Berrio, F. (2012). Aprendizaje integrado de contenidos y español para inmigrantes desde la perspectiva de la disponibilidad léxica: enseñanza del vocabulario. In R. Breeze, F. Jiménez Berrio, C. Llamas Saíz, C. Martínez Pasamar & C. Tabernero Sala (Eds.), *Teaching Approaches to CLIL/ Propuestas docentes en AICLE* (pp. 149-163). Servicio de Publicaciones de la Universidad de Navarra.

Jiménez Catalán, R. M. (2014). Preface. In R. M. Jiménez Catalán (Ed.), *Lexical availability in English and Spanish as a second language* (pp. v-ix). Springer.

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.

Lado, R. (1972). Patterns of difficulty. In K. Croft (Ed.), *Readings on English as a second language: For teachers and teacher-trainees* (pp. 277–291). Winthrop Publishers.

Laufer, B., & Nation, I. S. P. (2012). Vocabulary. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 163– 176). Routledge.

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

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

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied linguistics*, *33*(3), 299-320.

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

Milton, J. & Alexiou, T. (2020). Vocabulary size assessment: Assessing the vocabulary needs of learners in relation to their CEFR goals. In Dodigovic, M. & Augustin Llach, M.P. (Eds.) *Vocabulary in curriculum planning: needs, strategies and tools* (pp. 9 – 28). Palgrave/Macmillan.

Nation, I. S. P. (2012). *The BNC/COCA word family lists*. Retrieved from: http://www.victoria.ac.nz/lals/ about/staff/paul-nation

Palapanidi, K. (2019). Manifestaciones de clusters y switches en el léxico disponible de aprendices griegos de ELE en diferentes niveles lingüísticos. *MarcoELE: Revista de Didáctica Español Lengua Extranjera*, *28*. Retrieved from https://www.redalyc.org/journal/921/92157659004/html/

Paredes García, F. (2014). A vueltas con la selección de "centros de interés" en los estudios de disponibilidad léxica: para una propuesta renovadora a propósito de la disponibilidad léxica en ELE. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, *16*, 54-59.

Paredes García, F. (2015). Disponibilidad Léxica y enseñanza de ELE: el léxico disponible como fuente curricular y como recurso en el aula. *Linred: Lingüística en la Red*, *13*. Retrieved from https://ebuah.uah.es/dspace/bitstream/handle/10017/25257/Paredes_Disponibilidad_LR__2015_ 13.pdf?sequence=1&isAllowed=y

Payne, C. (2016). Lexical availability. *English Teaching Professional*, (102), 18-20.

Read, J. (2004). Vocabulary knowledge be defined. *Vocabulary in a second language: Selection, acquisition, and testing*, *10*, 209-227.

Richards, J. C. (1970). A psycholinguistic measure of vocabulary selection. *IRAL*, 8(2), 87-102.

Roghani, S. & Milton, J. (2017). Using category generation tasks to estimate productive vocabulary size in a foreign language. *TESOL International Journal 12* (1), 143-159.

Rojas Díaz, D., Zambrano Matamala, C., Salcedo Lagos, P., & Friz Carrillo, M. (2018). Un Enfoque de Reconocimiento de Patrones para el Análisis de Disponibilidad Léxica en Estudiantes de Pedagogía

en Matemática. *Estudios filológicos*, 62, 333-358.

Salcedo, P., & del Valle, M. (2013). Disponibilidad Léxica Matemática en Estudiantes de Enseñanza Media de Concepción, Chile, *Atenas*, *1*(21), 1-16. Retrieved from https://www.redalyc.org/articulo.oa?id=478048957001

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.

Sánchez-Saus Laserna, M. M. (2011). Bases semánticas para el estudio de los centros de interés del léxico disponible. Disponibilidad léxica de informantes extranjeros en las universidades andaluzas [Doctoral dissertation, Universidad de Cádiz].

Šifrar Kalan, M. (2009). Disponibilidad léxica en español como lengua extranjera: el cotejo de las investigaciones en Eslovenia, Salamanca y Finlandia. *Verba hispánica*, *17*(1), 165-182.

Šifrar Kalan, M. (2015). Lexical Availability and L2 Vocabulary Acquisition. *Journal of Foreign Language Teaching and Applied Linguistics*, *2*(2), 191-200.

Tomé Cornejo, C. (2015). *Léxico disponible. Procesamiento y aplicación a la enseñanza de ELE* [Doctoral dissertation, Universidad de Salamanca].

Van Ek, J. A., & Trim, J. L. M. (1991). *Waystage 1990: Council of Europe.* Cambridge University Press.

Webb, S. A., & Sasao, Y. (2013). New directions in vocabulary testing. *RELC Journal*, *44*(3), 263-277.

West, M. (1953). *A general service list of English words*. Longman, Green and Co.

Wingeyer, H. R. (2014). Temas de disponibilidad léxica de la región NEA de Argentina: mitos populares, la ciudad, el campo, insultos y diversiones nocturnas. Aportes para la elaboración de programaciones interculturales y de materiales en la enseñanza de e/le. *Signos ELE: Revista de Español como Lengua Extranjera, 8.* Retrieved from http://p3.usal.edu.ar/index.php/ele/article/view/2318.

Zambrano Matamala, C. (2021). Un estudio de la disponibilidad léxica en el ámbito de la autorregulación del aprendizaje en la formación inicial docente. *Lingüística y Literatura*, *42*(79), 11-33.

Leah Geoghegan is a language teacher and researcher with a PhD in English Philology (Universidad de la Rioja). She holds an MA in Theoretical and Applied Linguistics and an MA in Translation Studies. She has published chapters in the areas of second language acquisition, English as a lingua franca and study abroad.

Dr María Pilar Agustín Llach is a full Professor (English Philology, Universidad de La Rioja) and MPhil in Spanish Applied Linguistics. Among her research interest is vocabulary acquisition and teaching concerning age, gender, proficiency level, L1 influence, learning context (CLIL vs. non-CLIL), lexical transfer, and the instruments used to measure vocabulary knowledge are included in her research agenda.