

Agradecimientos

Me gustaría agradecer en primer lugar a mi tutora, Montserrat Bermudez Bausela, por su ayuda, correcciones y sugerencias.

A mi pareja y familia, que han sido el soporte moral y logístico para la realización del Máster y el presente trabajo.

A todo el equipo docente del Máster en Lingüística Inglesa Aplicada.

A las editoriales de libros de texto en inglés, porque nos facilitan enormemente el trabajo de recopilar textos y materiales de enseñanza, y en especial a las personas que amable y desinteresadamente me facilitaron información.

A mis compañeras profesoras de secundaria con las que comparto el día a día; muy aplicadas, muy de inglés, y muy de darle a la lengua. Mil gracias.

Abstract

The English variety being taught in Navarra (Spain) seems to be a British variety, due to the introduction of CLIL programmes in collaboration with the British Council and to the Anglo-American cultural references in the ELT coursebooks used. We suspect this might not be the best option for learners to get by in a globalized world in which English is increasingly used as a *lingua franca* so, in this research project, we decided to investigate if the vocabulary most frequently presented in the most popular ELT coursebooks in Navarra is the vocabulary speakers actually use. In order to do that, we have contrasted the frequency of said lexical items with their occurrence in a learner corpus (CLC) and in a native corpus (BNC) and analyzed these materials to determine the variety of English presented and the nature of the cultural references. After the exploration of the analysed data, we can interpret that the lexical terms presented in the coursebooks share similar frequencies with those in the corpora, but that the meaning taught in the coursebooks is not always the most frequently used in reality. In addition, we can conclude that the teaching of vocabulary could be improved if the materials were corpora-based, had more universal sociocultural elements and were not as *nativelike*.

Key words: variety, corpus linguistics, lingua franca, ELT materials, *nativelike*

Resumen

La variedad de lengua inglesa que se enseña en Navarra (España) parece ser una variedad británica, debido a la introducción de programas AICLE en colaboración con el *British Council* y a la prevalencia de referencias culturales anglo-americanas en los materiales utilizados. Ante la sospecha de que esta puede no ser la mejor opción de enseñanza para desenvolverse en un mundo globalizado en el que se utiliza cada vez más el inglés como lengua franca, en el presente trabajo se ha querido comprobar si el vocabulario que con más frecuencia se presenta en los materiales más comúnmente utilizados en Navarra es el que se utiliza en la realidad. Para ello se ha comparado la frecuencia de dichos términos con su ocurrencia real en un corpus de estudiantes de inglés (CLC) y en un corpus nativo (BNC) y se han analizado una serie de materiales para determinar la variedad del inglés presentado y la naturaleza de las referencias culturales. Tras explorar los datos analizados interpretamos que los términos léxicos presentados en los materiales comparten frecuencias similares en los corpus, pero que el significado que se enseña en los

materiales no es siempre el más utilizado en la realidad. Además, concluimos que la enseñanza del vocabulario podría beneficiarse si los materiales estuvieran basados en un corpus y tuvieran elementos de referencia socioculturales más universales y menos centrados en el *nativismo*.

Palabras clave: variedad, lingüística de corpus, lengua franca, materiales de inglés como lengua extranjera, *nativismo*

Table of contents

1. Introduction, hypothesis and objectives	9
1.1 Hypothesis	11
1.2 General objectives	12
2. State of the art	13
2.1 Corpus linguistics	13
2.2 Language Teaching	16
2.3 ELT Materials	19
2.4 Language Testing	23
2.5 English as a Lingua Franca	25
3. Methodology	31
3.1 Revision of the EFL materials used in public schools in Navarra	32
3.2 Sample of EFL textbooks	34
3.3 Data collection	36
3.4 Learner Corpus and Native Speaker Corpus	46
3.5 Results analysis	47
4. Discussion	52
5. Conclusion	55
6. Bibliographical references	58
7. Annexes	60
8. Resumen en español del TFM	68

Abreviaturas, siglas y símbolos

ALTE	Association of Language Testers in Europe
CEFR	Common European Framework of Reference for Languages
CLC	Cambridge Learner Corpus
CLIL	Content and Language Integrated Learning
COBUILD	Collins Birmingham University International Language Database
CPE	Certificate of Proficiency in English
CSE	Compulsory Secondary Education
EAP	English for Academic Purposes
EFL	English as a Foreign Language
EIL	English as an International Language
ELF	English as a Lingua Franca
ELFA	English as a Lingua Franca in Academic Settings
ELT	English Language Teaching
ENL	English as a Native Language
ESL	English as a Second Language
ESOL	English for Speakers of Other Languages
ESP	English for Specific Purposes
FCE	First Certificate in English
GA or GenAm	General American English
ICCI	International Corpus of Crosslinguistic Interlanguage
ICE	International Corpus of English
ICLE	International Corpus of Learner English
LINDSEI	Louvain International Database of Spoken English Interlanguage
LOCNESS	Louvain Corpus of Native English Essays

LTA	Language Testing Assessment
NNS	Non-native Speaker
NS	Native Speaker
PAI	Programa de Aprendizaje en Inglés
POS	Part of Speech
RP	Received Pronunciation
SLA	Second Language Acquisition
SUE	Successful User of English
VOICE	Vienna–Oxford International Corpus of English

1. Introduction, hypothesis and objectives

The present project aims to investigate the English varieties that are being taught in Compulsory Secondary Education (CSE) schools in the region of Navarra (Northern Spain). In order to do that, I will research into some of the lexical items that are presented in the most used EFL textbooks in the region and contrast them with the lexical items actually used by English speakers by carrying out a corpus-based research.

English has been established worldwide as an international language for communication (English as an International Language, EIL; English as a Lingua Franca, ELF) and international relations, and consequently, the teaching of English as a Foreign Language (EFL) also known as English as a Second Language (ESL), has taken an important place in teaching programmes in Spain in the last 25 years. This has led the education authorities to implement Content and Language Integrated Learning (CLIL) programmes starting from an early age, such as *Programa de Aprendizaje en Inglés (PAI)*¹ and a signed collaboration among the Ministry of Education of Spain, the British Council Foundation in Spain and the Autonomous regions in Spain, starting in 1996.²

Therefore, the English variety being taught aims to be a native British variety, even more so after the introduction of such CLIL programmes in some schools with the collaboration of the British Council; and the ultimate goal of EFL/ESL (both terms are used interchangeably) teaching is the acquisition of a *nativelike* English language and a deep cultural and socio cultural awareness of the British culture.

The idea of making students fluent in English with a native British speaker model in mind and widely skilled in the British culture and customs comes across as a great one. Nevertheless, as previously mentioned, English has become the international language of communication and of social and intercultural interaction, so English learners will be expected to communicate and interact with other learners from different cultural backgrounds. Consequently, as Anderson and Corbett (2009) suggest, it seems natural to question the status given to native-speaker performance as a model for language

¹ See <https://www.educacion.navarra.es/web/dpto/programas-de-aprendizaje-de-ingles-infantil-y-primaria> for further information about the Linguistic programmes in Navarra.

² See the following web sources for more information regarding the signed agreement in Navarra and in the Ministry of Education: <https://www.educacion.navarra.es/web/dpto/mec-british-council>
<http://www.educacionyfp.gob.es/va/mc/british-council/programa.html>

acquisition, and suggest a new model based on non-native speaker, ‘lingua franca’ interaction (Jenkins 2007).

Additionally, with recent developments that have somehow shaped the world differently, (Brexit, *Trumpism*, a worldwide health crisis as COVID-19 outbreak), the question of what native English variety and whether it is a native variety that we should aim for in the ESL teaching-learning process has taken a higher relevance. After all, in a globalised world that right now has people from a wide array of different nationalities connected through video call, social networks and the media, presenting native English as the goal and model for our ESL learners might not be the most adequate approach.

It is universally accepted that English is a widely used language today all around the world, but the relevant fact is that around 75% of its speakers use it as a second language³. This means that around three quarters of the current English speakers are non-native speakers (NNS), so they have learnt the language at some point in their lives or are still learning it.

Because of the prospects that a working knowledge of English can provide to people’s careers, many professionals feel the need to learn English for global communication and use it as ‘a contact language used among people who do not share a first language’ (Jenkins 2007: 1). As a consequence of this global use of English, the question of what native English variety and whether it is a native model that we should establish as the objective in the EFL/ESL teaching-learning process has become a key issue in EFL teaching research.

Research on ESL teaching and second language acquisition (SLA) started in the 20th C, focusing on different theories of acquisition and methodologies and still continues today. Research on English as a Lingua Franca (ELF) or English as an International Language (EIL) started around the year 2000 (see 2.5).

In addition, the intuition of EFL/ESL teachers that the contents presented in EFL textbooks do not match the reality in all parts of the world where English is the L1, has to

³ See the Interview and podcast with Jenkins in 2021
<https://www.english.com/blog/english-as-a-lingua-franca-a-podcast-with-jennifer-jenkins/>

be contrasted with corpus research and needs to be considered in the design of EFL/ESL materials, syllabus, methodology and English Language Testing. This is why Corpus Research is also a branch of linguistics in development, and it needs to be used to analyse and help design textbooks and materials for ESL teaching, whether the variety being taught was a native or a global one or one for specific or academic purposes (ESP/EAP). Corpus Linguistics applied to ESL teaching started also around the 2000s, with the development of ICTs, although Corpus-based Dictionaries and Grammars for English learners started to be developed in the 80s.

1.1 Hypothesis

The variety of English as a Foreign Language (EFL) being taught in the region of Navarra is somehow different from the native English varieties found in countries and regions where English is the first language (L1). These differences arise when we contrast the choice of lexical items and structures of native-speakers (NS) with that of NNS and observe the tight categorization of some structures in ELT materials, which is not as tight and turns out to be much more flexible in native English varieties. In addition, these native varieties portrayed in textbooks seem to be restricted exclusively to British and American ones (O'Keefe and McCarthy 2010: 414), with Cambridge ESOL examinations and CEFR (Common European Framework of Reference for Languages) levels receiving a lot of attention.

Additionally, it seems as an oversimplification and somehow naïve idea to present British English as a uniform variety, as there are quite a few regional varieties coexisting in the UK, with marked lexical and phonological differences among them. Received Pronunciation English (RP) seems to be the most universally recognisable and imitated type of British English accent, but is not necessarily the major variety in the UK. In the EFL teaching community, this simplification of British English as a uniform variety can be understood as a way to avoid overwhelming learners with distinguishing marked regional differences that add up to the lexical items they need to learn, so at basic levels this simplification would seem fair, and even the adoption of RP as a model to teach pronunciation might be justified. For the American English variety, General American (GA) is the umbrella term that describes the 'typical' American accent as opposed to RP

English, and is the variety that is modelled when teaching and learning an American English variety. So, if British and (North) American natives understand each other in spite of the regional phonetic and lexical differences, then, it comes as a logical consequence that native speakers should be able to adapt (in so-called accommodation) to the standard lexical or phonetic way of talking of non-native speakers and/or EFL learners too. This is a skill that does not come naturally to native speakers (of any L1) and generally requires some practice. To explore this idea would mean to extend beyond my research objectives, and due to time constraints I will not explore this further.

Going back to my previous argument, in my EFL teaching community we have observed that there is a tendency in EFL textbooks to oversimplify dialogues and texts, categorize structures too tightly and present vocabulary whose frequency does not always match the frequency found in native speakers' corpora or the way real language is used by speakers in native reality, regarding sociolinguistic and pragmatic features.

Even if the EFL taught and learned in Navarra differs in some aspects from the *nativelike* English variety, this EFL has some common features (lexical, grammatical and phonetic) that make it a valid language for global communication among EFL speakers with a different L1, in Spain and other non-English speaking countries.

1.2 General objectives

- To investigate what English variety is taught as a foreign language in schools in Navarra by looking at the ELT materials used.
- To investigate examples of production of EFL learners with a different L1.
- To compare and contrast the findings of the second objective with the variety being taught and with the language of native English speakers.

2. *State of the art*

2.1 **Corpus linguistics**

Corpus linguistics as it is known today started in the 1960s, and since then, its impact on our understanding of language has grown exponentially. For example, native speaker (NS) intuition about lexical meanings, lexical frequency, collocations, and the way grammar structures were used needed to be supported by corpus evidence.

Although the activity of looking at language, carrying out word counts, exploring collocations and grammatical structures is as old as the written Bible, these days Corpus linguistics has advanced in parallel to the development of technology and is deeply related to the advance of ICTs and their inclusion as part of the language study tools available. For example, in some research carried out by Frazier (2003), he showed that the traditional presentation of would-clauses for hypothetical conditions together with an if-clause is not as frequent in real language. Would-clauses are most frequently found without an if-clause, whereas in textbooks the relevance and space given to them as belonging together does not conform their frequency in the corpus data analysed. He provides evidence that ‘grammatical descriptions in traditional textbooks sometimes make claims about the grammatical co-text of features, and corpus studies can provide empirical testing of these claims’ as quoted by Conrad in O’Keefe and McCarthy (2010: 260).

The first use of corpora for linguistic research was carried out in the area of Lexicography, with the Collins Birmingham University International Language Database (COBUILD) project. This project started in 1980 in the University of Birmingham, led by John Sinclair⁴. Up to 2007, it had produced 16 dictionaries and grammars, the Collins COBUILD Dictionary, and the Collins COBUILD Grammar Patterns series, and inspired the design of the *Lexical Syllabus* (Willis 1990 as seen in O’Keefe, McCarthy and Carter 2007). In the section of Language Teaching we will deal further with the *Lexical Syllabus*.

⁴ See <https://www.collinsdictionary.com/cobuild/> for further information about the history of the COBUILD project

The COBUILD project was also important for grammar research, and it established the concept of 'pattern' as an interface between lexis and grammar and today, major grammars of English are corpus-informed. (O'Keefe et al 2007) Biber, et al. was probably one of the most relevant ones, taking all examples and texts from the Longman Spoken and Written English corpus (LSWE), a corpus that represents four registers of English of the British and American variety.

The research on corpora and their use in the ESL classroom is a relatively new topic, starting in the 1990s. There are only a very few commercially available grammar and textbooks for ESL teaching based on corpus data. The ESL teaching community seems reluctant to apply the use of corpora to their practice, but this might be due to lack of training in the use of corpora in the classroom. In the Masters in Applied Linguistics there are a few subjects related to corpus linguistics, as well as in other postgraduate courses in other linguistic areas. However, as we will see, it is not frequent in ELT Teacher Training courses.

Some of the previous corpus-based research carried out about 'real' language and how teaching materials present this language has shown a gap between what is 'real' and what is presented as 'real' in textbooks. This has been discussed by applied linguists in the last thirty years, such as Breen 1985; Scotton and Bernsten 1988; Nostrand 1989; O'Connor Di Vito 1991; Glisan and Drescher 1993; Gilmore 2004; Anderson 2007 as seen in O'Keefe and McCarthy 2010: 403)

Additionally, we would need to consider what is understood as 'real' language. Most linguists, EFL/ESL teachers and general population would consider 'real' language that which is spoken by NSs (native speakers) as opposed to that language spoken by NNSs (non-native speakers), seemingly imperfect in its use of grammar and lexical items and with a marked accent. Again it is worth keeping in mind that there is not a single NS English variety, but textbooks and materials simplify this variety under an umbrella term of 'British English' or 'American English'. Following Jenkins (2007), it is necessary to consider the existence of English as a Lingua Franca (ELF) being used all over the world, which is as alive and as 'real' as any other privileged NS English variety. She defines ELF as 'a contact language used among people who do not share a first language, and is commonly understood to mean a second (or subsequent) language of its speakers'

(Jenkins 2007: 1). She advocates for the necessity of learners making an informed choice as to what variety to learn, depending on whether they will be interacting with NSs or other NNSs. ELF research is quite recent, as already mentioned in the introduction. There are only a few ELF corpora available, and they will be considered later on.

Going back to Corpus Linguistics, in any course on Corpus Linguistics, the building and design of a corpus will be one of the main topics. Nevertheless, the aspect that is of utmost relevance is providing the linguist with the necessary tools to be able to choose the most adequate corpora out of the available ones for their research purposes. That would include how to access corpora, search for information, and interpret the information that they return. Building a corpus can be an interesting research project, but for my purposes and due to time constraints, here we need to resort to already available corpora.

There are general learner corpora available, collected by universities, publishers and research groups that are useful for ESL research and this project. These are quite useful in order to compare learner production with comparable native corpora, or even compare learners from different L1s, categorize their mistakes and create specific materials for them. Several studies with learner corpora have been carried out, such as those of Granger (2003) in O’Keefe and McCarthy (2010). Granger considers that research into learner corpora ‘brings about exciting pedagogical perspectives in a wide range of areas of English language teaching (ELT) pedagogy: materials design, syllabus design, language testing, and classroom methodology’ as presented by Cheng in O’Keefe and McCarthy (2010: 319)

Some of the best-known learner corpora are ICLE (International Corpus of Learner English) and LINDSEI (Louvain International Database of Spoken English Interlanguage), both by *Université Catholique de Louvain*. The ICLE has samples of written language exclusively. The full version contains over 9,000 texts (more than 5 million words) written by learners from 26 different L1 backgrounds. It consists mainly of argumentative essays, while its spoken counterpart, LINDSEI, contains transcripts of spoken language elicited by informal interviews and by picture prompts.

In order for the data to be comparable, these learner corpora need to be contrasted with a similar native corpus which needs to be balanced in topics, genres and other variables such as age and time of publication. For example, LOCNESS (The Louvain Corpus of Native English Essays, made up of 324,000 words), contains essays (mostly argumentative) written by British and American students, post-secondary to university level and because of that it is often used for direct comparisons with the ICLE Corpus of non-native speaker writing because they represent the same genre and topics (O’Keefe and McCarthy 2010).

However, these learner corpora showcase the language of advanced learners. For data of younger learners and/or lower level learners, the CLC (Cambridge Learner Corpus) has become an invaluable tool. The Cambridge Learner Corpus is an 11,000-text corpus of learner responses to ESOL exams, compiled by Cambridge University and Cambridge Language Assessment. The uncoded version (without error tagging but POS tagging) is open and accessible to all users.⁵

In addition, the International Corpus of Crosslinguistic Interlanguage (ICCI), by Tokio University, collected data from lower-level learners in Austria, Hong Kong, Israel, Poland, Spain and Taiwan. When O’Keefe and McCarthy (2010) was published, ICCI was a project in progress whose aim was to collect the writing of primary school to pre-university learners of English across various proficiency levels and L1 backgrounds. However, I cannot find any new information about this project and/or corpus details.

2.2 Language Teaching

As previously mentioned, research on ESL teaching and second language acquisition (SLA) started in the 20th C, focusing on different theories of acquisition and teaching methodologies.

How languages are learnt, acquired and accessed is one of the main concerns of applied linguistics, both in L1 and L2 contexts. There are obvious differences between acquisition of L1 and L2, such as age, context, contact with the language, motivation, etc.

⁵ Access the Open CLC: https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fcup_open_clcu

L2 acquisition is what is more relevant to our purposes, and therefore has been subject to acquisition theories, different methodological approaches and research. We will provide here a short overview of the main methods used for L2 teaching, also popular for ESL teaching.

In the early 20th C languages were taught by the ‘Direct Method’, which made emphasis on listening and speaking and gave more importance to language exposure than the former ‘Grammar translation Method’ popular in the previous century. From then on, methods focused on other skills, such as the ‘Reading Method’ or ‘Audiolingualism’. Then, with the advent of the concept ‘communicative competence’ in the early 1970s (Hymes, 1972) a methodology based on communication and of a more dynamic nature was devised and popularized, basing L2 learning on the learning of communicative functions. This is what gave way to Communicative Language Teaching and ‘emphasized the use of language for meaningful communication’ (Littlewood, 1981 in Schmitt 2013). Meaningful communication in language teaching takes place when the interaction among learners is a real one (not ‘pretended’ dialogues), such as having to find information which is indeed unknown by one of the learners, asking for directions, asking for and giving personal information to a classmate in the first days of a language course or having to interact to find an agreement about a real-life issue. This meaningful communication was then taken out of the language class and used for communication in other non-linguistic areas. As Schmitt says (2013: 6),

Taken further, students could be taught some non-language-related subject, such as history or politics, in the L2. The assumption was that the learners would acquire the L2 simply by using it to learn the subject matter content, without the L2 being the focus of explicit instruction. Taking the communicative approach to its logical extreme, students could be enrolled in ‘immersion’ programmes where they attended primary or secondary schools which taught subject matter only in the L2.

These ‘immersion’ programmes became very popular in Spain, especially in the last twenty five years. The findings about these immersion programmes (also known as CLIL programmes) show that students generally acquire a good command of the L2, but that certain problems of accuracy still persist. In any case, research shows different and varied outcomes of these programmes, and their adoption worldwide has been so extended that it is only natural to receive contradictory findings. These programmes have been implemented in countries with different educational contexts and backgrounds, with varying degrees of political and local authorities intervention and with different

availability of human and/or material resources. In the case of the region of Navarra, the CLIL programme that was adopted in Primary and CSE became known as *Programa de Aprendizaje en Inglés PAI* together with the Ministry of Education of Spain and the British Council Foundation in Spain.

With the implementation of this programme, the question of which English variety was being taught came to the forefront. Unquestionably, the variety was a native one, and due to the signed collaboration with the British Council foundation, the proximity of Spain to the UK, and the UK being still active part of the Council of Europe, it seemed natural that the native variety chosen for ESL teaching and CLIL programmes was British English.

There have been some case studies about the direct use of corpora in the ESL classroom. Most teachers do not feel confident enough in their use of corpora and concordance tools in order to attempt this. An interesting article exploring this further is that of Gabrielatos (2005), which explores the use of corpora for ELT purposes. Firstly, the author presents the different types of corpus that can be found and to what measure they are useful for ELT teachers. Secondly, the author aims to demystify corpora and define their place within language teaching as a whole (2005: 2). We can see how corpora constitute the perfect tool for condensing language exposure, rather than creating a sample text for specifically ELT purposes in which a particular structure is replicated in a higher proportion than normal. Finally, the author introduces ideas to use corpora in the classroom in a flexible way, such as devising tasks to explore collocations and meaning of a set lexical item as well as to revise and critically examine grammar rules. These can be carried out in what is called “soft version” or teacher-centered approach or “hard version” or learner-centered approach.

However, as Schmitt declares, there is need for further research and experimentation on the use of corpora and concordance tools in the classroom. He is confident that ‘concrete evidence about how effective these methods are will only become apparent over time, once enough teachers have experimented with the use of corpora as reference sources and learning tools’ (Schmitt 2013: 294).

If corpora are not being used in the classroom yet, and the effects of their use are still to be investigated, this brings us to the necessity to revise ELT materials and see whether

they are corpora-based. The introduction of corpora-based classroom tasks might seem more feasible and somehow less challenging if ELT writers are using corpus-data to design their materials.

2.3 ELT Materials

The rapid implementation of CLIL programmes as a way to make L2 acquisition more immersive affected the design of ELT course materials and, to a lesser degree, Syllabus design. Most EFL/ESL textbooks used in CSE have specific references to the CEFR levels or Cambridge ESOL examinations. Grammatical structures and vocabulary are presented by means of written and recorded texts which are carefully designed to present the structures and the vocabulary required for the specific level and the year of the student. Some of these structures are presented descriptively, but the description is quite limited and sometimes even restrictive. They are inserted into texts to exemplify their use for the learner, and the result constitutes a condensation of such structures which is not found naturally in ‘real’ language. The same happens with target lexical items, which are replicated in texts in a higher proportion than normal compared with corpus data.

This is supported by studies such as those of Römer 2005; Cheng and Warren 2005. They show that the English presented in EFL teaching textbooks ‘are often decontextualized and lack an empirical basis’ when compared with English used in natural environments. (O’Keefe and McCarthy 2010)

To make things even harder, the syllabus of EFL for CSE in Spain is made of a series of descriptors of syntactic contents that learners should be taught and know after each year. ⁶The lexical contents are more flexible and vary from one material to the next. These contents present assumptions, intuitions about language use and frequencies that are not supported by the findings of Corpus Linguistics. As we saw before, Corpus Linguistics has been adopted as an invaluable tool to study language in general, but it has not taken off as a tool for EFL teaching, nor for the design of EFL materials and syllabus. This might be due to the fact that research on corpora and their use in the EFL classroom is a relatively new topic, starting in the 1990s.

⁶ See the official Syllabus of the Ministry of Education for First Foreign Language (English) p254 <https://www.boe.es/boe/dias/2015/01/03/pdfs/BOE-A-2015-37.pdf>

Most publishing houses of EFL teaching materials do not select the most frequent lexical items from corpora, but rather choose them according to their own semantic criteria. This provokes variability from one material to the next and the learning process is not facilitated. Of course, frequency is not the only criteria for selecting particular lexical items for an EFL course, as there are also other criteria such as familiarity, availability, and level of difficulty (O’Keefe and McCarthy 2010), but it should be taken into account, as when students are faced with ‘real’ language, that the most common words are the ones that are more likely to appear.

In an attempt to determine with a higher degree of certainty the criteria for lexical-item selection in the teaching materials, I contacted the three most popular publishers of ELT materials in Navarra. Their answers can be seen on section 3.2.

In a study from 2002, Biber and Reppen explored ‘the grammatical features to be included, the order of grammatical topics, and the vocabulary used to illustrate these topics’ in six EFL/ESL materials. O’Keefe and McCarthy (2010: 323) explain that:

[Biber and Reppen] highlight sharp discrepancies between the information found in grammar materials and the real-life language use that learners encounter. Based on their findings, they argue that ‘frequency should play a key role in the development of materials and in the choices that teachers make in language classrooms

Indeed, just a few of the textbooks and materials published for classroom use today are corpus-based. As Hughes explains in O’Keefe and McCarthy (2010: 402):

The slowness of commercial English Language Teaching (ELT) publishers and the teaching community to adopt corpus insights more enthusiastically is due to the powerful influence of traditional pedagogic grammars promoted by titles such as these [Murphy’s *English Grammar in Use* and Swan’s *Practical English Usage*] and also to other broader factors such as the tendency for a lack of communication between the teacher training community in applied linguistics and the research community.

What Hughes mentions about the lack of communication is also relevant, as she points out to the problem that we had foreseen in the Corpus Linguistics section: there is no real contact or interaction between the ELT training and ELT research. Courses in ELT do not present training in Corpus Linguistics, and this lack of training might be why the use of corpora is disregarded by the ELT community. The communication among teachers and researchers should be two-way. Literature on Corpus Linguistics and how to use corpora in the EFL classroom is not scarce, and the teacher with an interest will certainly be able

to find it. However, there are no data about the real use of corpora by EFL teachers in Spain, apart from the occasional case study. So, if corpora are being used by teachers, their feedback is not reaching the researchers.

Three decades back, the *Collins COBUILD English Course* series was written by David Willis. It was a lexical syllabus based on the COBUILD (Collins Birmingham University International Language Database) project, collected by John Sinclair. This could have been the start of a successful integration of corpus linguistics and ELT by means of the creation of corpus-based ESL teaching materials. However, this idea did not really took off, the series had a short popularity, and the integration of corpus data has become an exception rather than the norm. Nowadays, other than the *Touchstone* series, ELT materials are not based on corpus data (see Walsh in O’Keefe and McCarthy 2010: 334 and McCarten, same).

The *Touchstone* series is a coursebook by Cambridge for adults and young adults, and presents American English as its target. Some other books of the same publishing house declare to be corpus-informed (rather than corpus-based). Let’s look at these distinctions. Following McCarten in O’Keefe and McCarthy (2010), she considers three approaches when it comes to the writing of course materials and their relationship to existing corpus. The first one is the ‘corpus-driven’ approach, where ‘the corpus provides the basis of the description of language usage without recourse to previously held beliefs’ (2010: 415); then, there is the ‘corpus-based’ approach, in which ‘the corpus provides examples for pre-existing rules’ (2010: 416). Finally, she presents the third approach advocated by McCarthy (1998) the “corpus-informed”, which borrows from both approaches and is suggested as an alternative possibility.’ (2010: 416)

Then, considering all the previous ideas, it is clear that 1) ELT materials today are not corpus-based, even if a few of them might be corpus-informed, and do not always reflect the ‘reality’ of the actual use of language and 2) EFL teachers lack the training or the willingness to use corpora in the classroom, as the prevalence of ELT materials from popular publishing houses would not make them think that the contents are not reflecting the linguistic reality accurately enough.

This distance between ELT material design, research and teaching seems to be growing wider apart. However, in her chapter in O’Keefe and McCarthy (2010), McCarten presents a selection of the areas where coursebook design can be informed by corpora quite easily. She describes vocabulary syllabus such as the Syllabus of the *COBUILD English Course* by Willis, determined by word frequency and what is known as a ‘common core vocabulary’. The basic idea of the common core vocabulary is that the most frequent 700 words in English all occur at least 650 times each in the Main Corpus (this core vocabulary constitutes Level 1 in the Course). Then, the rest of the most frequent 2,500 words occur 120 times each in the Main Corpus (and are taught in the 3 subsequent levels). These lists together constitute a 10% of the corpus. The remaining 90% of the corpus are rare words, with frequencies below 50, and constitute the ‘Reserve Corpus’.

As Willis explains in his book *The Lexical Syllabus*,

The commonest patterns in English occur again and again with the commonest words in English. If we are to provide learners with language experience which offers exposure to the most useful patterns of the language, we might well begin by researching the most useful words in the language. (Willis, 1990: 45)

However, we should not forget that frequency alone is not reason enough to select which lexical items to be taught and which not. For example, as McCarten reminds us, colour names are not of a similar frequency, with nouns such as ‘black’ or ‘white’ being much more common than ‘orange’. However, reason and common sense would suggest teaching the colours together. The same applies for the days of the week or the seasons.

Another further possibility to inform a course syllabus on corpora is to use wordlists. There are Academic Wordlists, aimed at teachers and learners of ESP/EAP. Similarly, Oxford has created their own vocabulary lists, the Oxford 3,000 and 5,000 for EFL/ESL learners at any level. The lists present the most important 3,000 and 5,000 words respectively, based on corpora of written and spoken tests.

Similarly, a corpus can help us find common collocations, multi-word units, clusters and lexical bundles that are of a high frequency. Some of these items are quite frequent, as in the form of greetings and set phrases that are widely used but not taught at an elementary

level. Finally, a corpus can also help the material designer to look at the context in which certain structures are used and select the most appropriate ones for teaching.

2.4 Language Testing

Language Testing and Assessment (LTA) is also an essential part of ELT worldwide. Since it started, back in the early 20th C with the Cambridge ESOL's Certificate of Proficiency, many are the institutions that have offered proficiency language tests.⁷

In turn, many institutions have also required a certain proficiency level for their prospective employees or students. Testing and assessment is something which is regularly happening locally in ESL classes. It is by testing that teachers can inform learners' performance by comparing their performances to those corpus data corresponding to the performance of other learners in their level, specified at CEFR level or not. So we can assert that language testing is happening non-stop locally, nationally and internationally.

In order to write a test that measures learners' performance at a specific level, it is important to establish what the level indicators are and what learners of a language can do at a certain level. This is done by analysing learner data and learner corpora, such as CLC and ICLE.

It is interesting to note that for language testing the corpora that are used are learner corpora. By classifying the gender, motivations, ages, nationalities, demographic variable, hours of input, instruction and practice, learner corpora are created with samples of their performance. It is by comparing different learner corpora that we will find the structures and vocabulary most common to the level corresponding to that number of input hours, time of practice and tuition, so language testers will be informed about the essential structures and vocabulary that have to be found at that level. To put it in other words, 'language testers therefore need to use corpus data to identify the linguistic exponents of a particular proficiency level which can only be done reliably if a learner's

⁷ See <https://www.cambridgeassessment.org.uk/about-us/who-we-are/our-heritage/> for historical information regarding Cambridge ESOL examinations.

level is correctly identified and recorded in a corpus' (O'Keefe and McCarthy 2010: 636-7)

The English Profile Project makes use of learner data. This project was developing Reference Level Descriptors for English, in collaboration with the Council of Europe. This was done by using the CLC as a starting point and then collecting written and spoken learner data from international teams, from classrooms and other settings, in which every variable was recorded (background information, educational context, mother tongue, self-perceived and externally assessed proficiency level) with the aim of balancing the corpus and create hypothesis about what the key features that define a level are. (O'Keefe and McCarthy 2010). Similarly, the Association of Language Testers in Europe (ALTE) created a reference-scale with descriptors of *can-dos* aligned with each of the six CEFR levels.⁸

Native corpora are used in test writing, but to ensure validity and reliability. NS corpora provide the texts to be used at examinations. These texts can be used in the original form, adapted, or as a model for production tasks, for example. This ensures that the texts presented in tests are real, relevant to the world outside the testing situation and with natural frequencies of lexical items and structures. There has been some research on the impact of native text adaptation for tests, such as that of Hughes (see Barker's chapter in O'Keefe and McCarthy 2010). The author investigated the impact that editing of authentic texts had in FCE reading texts with native corpus frequencies to find that 'the simplified texts differ from authentic texts in a number of ways but there were no significant differences in the abstractness and ambiguity between the two groups of texts' (O'Keefe and McCarthy 2010: 641). This encourages the continuation of native corpora as a source for test texts, although further research is welcome.

Finally, we should consider the status of the native speaker for both ESL language teaching and testing. Native data corpora are not useful to test learners, as we cannot ask natives to perform at a specific level exclusively. But what we can do indeed is to compare (proficient) learner performance with native one as a way to inform, orient and

⁸ See Council of Europe website and <https://www.coe.int/en/web/common-european-framework-reference-languages/the-cefr-descriptors> and English Profile Project website <https://www.englishprofile.org/>

improve proficient production (lexical and grammatical structures, prosodic and pragmatic features). However, it is natural to expect that the same task performed by different native speakers will differ greatly from one another. This characteristic seems to be more person-related and context-related than native language dependent.

Some researchers, including Cook (1999), Widdowson (2000), Seidlhofer (2004, 2005), Jenkins (2006),

question whether learners should be judged against the models and norms extracted from native-speaker corpora, and advocate instead those of English as a Lingua Franca (ELF), or World Englishes (Kirkpatrick 2007). They argue that realistic models for learners are to be found in non-native-speaker corpora (O'Keefe and McCarthy 2010: 415)

We understand that judging a ESL learner against a NS is not fair, and proposing a native speaker model for production tasks might be challenging for the learner, who might see NS standards as unattainable. Indeed, this is the case for some proficient learners, who even after years of studying ESL and in possession of C2 Proficiency level, feel that they do not mix with the natives because they will always be marked in terms of accent or sociocultural differences. Jenkins talks about this in her book *English a Lingua Franca: Attitude and Identity* (2007). The idea of exploring a ELF model or an EIL model against which to assess learners internationally is worth considering. We will look at that in the next section.

2.5 English as a Lingua Franca

English has become the international language for worldwide communication. No matter the field or register, from the top academic domain, going through the professional and educational domain to the domain of marketing, sports, tourism and commerce, people from different L1 will choose English as the language of communication.

The English variety used as a means of communication among people from different mother tongues has been labeled differently: World (Standard) English (Crystal 2003), International English (Rubdy and Saraceni 2006), Global English (Phillipson 2007). For Jenkins (2007) and Seidlhofer (2004), English as a Lingua Franca (ELF) is the preferred term, as using other terms might be misleading in the sense that it 'suggests that there is one clearly distinguishable, codified and unitary variety called *International English*'

(Seidlhofer 2004: 210, as seen in Jenkins 2007). As Kirkpatrick defines, ‘a Lingua Franca is the common language used by people of different language backgrounds to communicate with each other’ (Kirkpatrick 2007: 7), and this applies to English and how it is used all around the world, both in ENL countries (countries of the Inner Circle – countries where English is spoken as a native language, see Kachru 1985 in Jenkins 2007) and ESL/EFL countries (countries of the Outer Circle, ex-colonies; and countries of the Expanding Circle; *ibid.*) when the people communicating happen to be NNSs of English.

Additionally, ELF is a very different concept from both ESL and EFL, whose goal Jenkin acknowledges to be the acquisition of English as a Native Language (ENL).

Therefore, if we consider for a minute who our English learners in CSE will be likely communicating with in the future, we might need to change our approach. Brexit has made it difficult right now for EU citizens to freely go study and work in the UK. Even commercial relationships between the UK and mainland Europe are more complex, thus interactions will be restricted. Of course the possibility of travelling to the UK for studying or working still exists, but it will certainly not be as convenient as it was a few years back. Additionally, in the USA, immigration policies have not been very indulgent for citizens from other countries to establish there. In fact, with Trump’s government, some nationalities saw their possibilities of entrance to the USA even more restricted. To explore all of this would mean to extend beyond my research purposes, but the point which we are trying to make is clear: our learners will probably interact more with other NNSs than with NSs, so restricting ESL teaching to a *nativelike* model implies restricting their future prospects.

Research on International English and ELF is quite recent, but it has produced interesting resources. The International Corpus of English (ICE) was created in order to compare English varieties around the world. The informants are NSs with their own national or regional English variety. This corpus deals with varieties spoken in the former British colonies and other World English varieties.

There are only two corpora that deal with interactions of NNs where English is used as a lingua franca as the speakers do not share an L1 or a common cultural background. One

is the Vienna-Orford International Corpus of English (VOICE) and the other is the corpus of English as a Lingua Franca in Academic Settings (ELFA) (see Lee in O’Keefe and McCarthy 2010). VOICE records one million words of natural spontaneous interactions among NNSs. ELFA deals with academic interchanges, and also consists of one million words.

Research on the VOICE corpus identified a series of lexico-grammatical differences between NS English and ELF (Seidlhofer 2004 in O’Keefe et al. 2007). For example, dropping the 3rd person singular -s in the Present Simple, omitting definite and indefinite articles, inserting prepositions, using the universal question tag ‘isn’t it/innit’ and conversing count to uncount nouns or vice versa. Although most ESL/EFL teachers would classify these as mistakes, Seidlhofer admits, based on her research, that they seem to cause no problems for successful communication in ELF.

Jenkins’ research, on the other hand, is more concerned with the phonological aspects of ELF. She found that some phonological features which are common to NS varieties of English were not necessary in ELF interactions for successful communication. She mentions examples such as the absences of certain weak forms, the distinction between voiced and voiceless ‘th’ with its substitution with /t/, /f/ or /s/ for voiceless /θ/ and /d/ or /z/ for /ð/ (see Jenkins in Rubdy and Saraceni 2006 and Jenkins 2007). She defines the phonological features that determine intelligibility as ‘Lingua Franca Core’, and vehemently reminds critical voices that hers is not a model as it ‘respects both ELF learners’ right to choose whether or not they adopt it and the diversity of their accents’ (Rubdy and Saraceni 2006: 36). In addition, she goes on to say that ‘it is the near universal focus on RP and GA in Expanding Circle English language education which is an imposition - one which I am endeavouring to reverse through my ‘core’ approach’ and that:

NS accents are not only sociolinguistically inappropriate for communication in which NSs are rarely involved, but also psycholinguistically and socio-psychologically unachievable for the majority of adolescent and adult learners (ibid.)

Similarly, Prodromou (2005), who collected a corpus of spontaneous speech of proficient non-native users of English between 2000 and 2003, refers to the idiomatic aspect of a language as something that is easy for NSs, but challenging and unnecessary for NNSs. Idioms are embedded into culture and tradition, and they do not make part of the NNS

mental lexicon. Using them would hinder communication in ELF, and actually, proficient NNSs manage to communicate efficiently by strategic use of their resources without resorting to idiomaticity or metaphor, and their discourse is more transparent than that of some NSs. (Prodromou 2005 in O’Keefe et al. 2007)

Considering these aspects and the idea that our learners might not have in mind blending in with NS of English in their near future, we should consider the status of the NSs. It is clear by now that a *nativelike* English variety should not generally be the model for our learners. Then, what or who should the model speaker be, or rather, the focus? Prodromou (2003) considers that the focus should be the ‘Successful User of English’ (SUE). Other authors admit that learners may prefer to approximate native-speaker norms without losing their ‘cultural integrity’ (Timmis 2005 as seen in O’Keefe and McCarthy 2010: 415)

In Prodromou’s chapter (see Rubdy and Saraceni 2006) he presents the findings of his research, which, consisted in exploring the roles of idiomaticity in NNS discourse as well as to describe varieties of EIL/EFL as used by the participants, who defined themselves not as ‘learners’ of English, but as regular users. This allowed him to further explore the concept of the SUE. SUEs are seen as expert users, whether native or non-native. In his research, he focused on the differences between successful native-speaker users and successful non-native-speaker users. Successful non-native speakers have a flawless command of grammar and are rarely lost for words as long as these refer to individual lexical items, used in a more or less referential manner. The mistakes made, if any, tend to be in the context of idiom use and collocation. He is concerned with the use of idiomaticity, as we mentioned before, and found out that successful non-natives hardly use any ‘pure’ idioms and are rarely creative in that sense, unlike successful native users who exploit pure idioms for creative effects. Non-natives use a variety of collocations and lexical phrases, but generally transparent and non-metaphorical. (Rubdy and Saraceni 2006: 63). He concludes that the speech of NS can be distinguished from the speech of non-native SUEs by the existence of hedging and expressions such as ‘sort of’, ‘you know’, ‘I mean’, and phrasal verbs, which are more common in NSs. All in all, ‘SUEs are highly successful L2 communicators but who will achieve this goal by strategic use of their resources in ways different from those of native speakers’. (O’Keefe 2007: 84).

Seidlhofer considers that the preferences of learners and users of English as to which variety of English they need and want should be taken into account and that ‘some awareness of the global roles of English should be achieved by all English users in the Inner, Outer and Expanding Circles alike’ (Seidlhofer in Rubdy and Saraceni 2006: 48).

Unfortunately, it is not up to learners in CSE to choose the variety they want, but we teachers could inform them about the different Englishes existing so they can decide about their future English language training based on their plans for the future. It is important to make learners aware of the existence of that ‘expert user’ or SUE model, to use Prodromou’s term, and look up to them rather than attempt to reach that unattainable native speaker model.

After looking at these arguments it seems clear so far that a native speaker model should not be the aim of our ESL/EFL teaching classes. However, this seems to contradict the implementation, as previously mentioned, of the CLIL teaching programmes signed among the ministry of Spanish Education and the British Council. Additionally, the idea that native English should be the goal of ELT is quite extended among people making decisions in our institutions and the public opinion in general. Then, the learners and teachers, who should be the ones who have a say about how their own learning and teaching should be, are rarely taken into account when it comes to deciding which model of English to learn and teach.

Kirkpatrick acknowledges that the pressure of publishers and international English language teaching institutions promoting a NS variety of English is not without commercial interests. The ELT materials and testing materials they create are based on NS models and they will push for that, making us believe that native is better and ‘they will suggest that native-speaker teachers are somehow innately superior as language teachers’ (Kirkpatrick in Rubdy and Saraceni 2006: 71). Kirkpatrick goes on to assert that:

The insistence on a native-speaker norm diminishes local teachers of English and undermines their self-confidence and self-respect. At the same time, the very advantages that they can bring as teachers are disregarded. (idib.)

Kirkpatrick talks about the existence of three models of English, the native-speaker model, the local nativized variety and the Lingua Franca model. He also argues that ELF

can ‘free speakers from what they might feel to be the cultural straitjackets of their L1s’ (Rubdy and Saraceni 2006: 80). This means that when using ELF, pragmatic convention of the L1 can be dropped, if that is what the speaker wishes, as communication is the main focus.

Finally, he concludes by saying that a lingua franca model is the most reasonable one when the learners’ major reason to speak English is to communicate with NNSs. However, he considers that there is not yet an adequate description of lingua franca models and that research should go on that way. Until then, learners and teachers will have to continue learning and teaching nativized or native-speaker models. We could add that this is more the case when the pedagogical decisions regarding ESL/EFL syllabus are not for the teachers or learners to make, as it happens in CSE in Spain.

McKay (McKay in Rubdy and Saraceni 2006: 127) considers that English in EIL no longer belongs to a single culture and therefore it is necessary to be culturally aware of the diversity of contexts in which EIL is taught and used. She goes on to suggest that, in terms of ELT materials, the traditional use of Western cultural content in ELT texts needs to be examined.

In the next section, we will present the methodology we will use to revise and analyze the ESL/EFL materials used in Navarra to see which English variety they present, whether their contents are corpus-based and whether these match with ‘real’ English.

3. Methodology

1. Revision and analysis of the EFL teaching materials being used in public schools in Navarra in 2nd year of CSE this current year 2020/21:

- Search on the schools' website what coursebook they use for 2nd CSE and write down the titles.
- Count the total number of times each coursebook is used and annotate the publishing house, country of edition and English variety they claim to teach, whether there is any reference to external ESOL examinations and if so, which one, whether there is a reference to CEFR levels (European Framework).
- Select the 3 most popular coursebooks for in-depth analysis.

2. Sample of EFL textbooks:

- Take a sample of a lexical field found in all of the EFL textbooks, and record the words presented in a word list.
- Take a sample of an activity dealing with receptive skills to see how context is presented and analyze it qualitatively.
- Observe the cultural pages of the textbook (if any), to find which English-speaking countries are presented as a model.

3. Data collection:

- Create a spreadsheet to collect and categorize all these data.

4. Revision and exploration of learner corpora with different L1s and a native English corpus comparing a selection of lexical elements:

- Use the Open Cambridge Learner Corpus.
- Search for the frequencies of the words in the word list recorded above in the learners' productions.
- Use an Open Corpus of English. A good balanced option is the BNC for the British variety.

5. Results analysis

3.1 Revision of the EFL materials used in public schools in

Navarra

In the region of Navarra there are currently 48 public schools that offer CSE, that is Compulsory Secondary Education from 12 to 16 years old. With the implementation of immersion and CLIL programmes in Primary Education, there were increasing numbers of learners whose L2 competence was above the average compared with the learners not enrolled in immersion programmes. CSE schools had to cater for those learners and provided a continuation to said programmes. Currently, there are three possible scenarios in Secondary Education. First, what is known as ‘Bilingual Sections’ (*Secciones Bilingües SSBB*), which integrates language and content using English or other languages as the instruction language in other non-linguistic areas. There might be selection criteria for accepting students into the programme, such as an exam or a minimum mark in English in Primary school.

Secondly, what is known as ‘Plurilingual Secondary’ (*Secundaria Plurilingüe*), which gives continuation to the programmes of PAI (*Programa de Aprendizaje en Inglés*), British Council, and other French or German language programmes started by the learners in Primary Education. The entry requirement is that the student comes from a school with the programme and that they have attended it. The methodology used is CLIL.

The British Council programme, within what is understood as ‘Plurilingual Secondary’, is offered by four schools in Navarra, and the treatment offered is different, as their syllabus is integrated with that of the British Council syllabus and the English level aimed at is higher, by looking at their coursebooks materials. The number of schools that offer ‘Bilingual Section’ programmes or ‘Plurilingual Secondary’ together is 31 out of the total 48.⁹

Finally, there is the scenario in which learners are not attending any linguistic programme. In spite of this, it is worth mentioning that in the region of Navarra the linguistic situation is very rich. Most students, apart from the subject of EFL, will have Basque (as a language subject in model A, or as a vehicle language in model D) or

⁹ See [Annex 1](#) for full list of schools in Navarra.

French (as a language subject or as a vehicle language) , so English is just another language they have to learn. This is the case of 17 schools. ¹⁰

The schools can freely choose the coursebooks to be used in their classes. There are generally differences among them and also within the same school, as it is common practice to choose a different coursebook for the bilingual programmes and the non-bilingual programmes.

In order to test our hypothesis, we need to look at the materials used in the different schools and for the different linguistic programmes. Therefore, the following steps have been taken: we first conducted a search on the schools' websites for the coursebook being used for 2nd CSE and wrote down the titles. This information was easy to obtain from the school websites. However, a few schools do not specify what coursebooks they are using or the information is terribly outdated. After going through the list of all the schools, we counted the total number of times each coursebook is used and annotated the publishing house, English variety they claim to teach, whether there is any reference to external ESOL examinations and if so, which one, and whether there is a reference to CEFR levels (European Framework) (See Table 1).

COURSEBOOKS ¹¹	Number of schools using it	Publisher	English Variety	Reference to ESOL examinations	Reference to CEFR level
A 2	2	Burlington	Not specified	no	yes
AEU 2	6	Burlington	Not specified	yes	yes
ARE 2	5	Burlington	Not specified	yes	yes
ATA 2	2	Burlington	Not specified	yes	yes
B B1	1	Macmillan	British	no	yes
BIE A2	1	Burlington	International?	yes	yes
BIE B1	2	Burlington	International?	yes	yes
BIE B1+	2	Burlington	International?	yes	yes
C 2	2	Cambridge	Not specified	yes	yes
CPSS	1	Cambridge	British	yes	yes
D 2	2	Oxford	Not specified	no	yes
EU 2	1	Burlington	Not specified	no	yes
M 2	9	Oxford	Not specified	no	yes

¹⁰ See <https://www.educacion.navarra.es/web/dpto/secciones-bilingues> for information on Linguistic programmes in CSE.

¹¹ See [Annex 2](#) for full coursebook titles per school and coursebook count.

NEU 2	11	Burlington	Not specified	no	yes
NEU 3	1	Burlington	Not specified	no	yes
P 2	2	Macmillan	Not specified	no	yes
RE 2	4	Burlington	Not specified	no	yes
TA 2	4	Burlington	Not specified	no	yes

Table 1

3.2 Sample of EFL textbooks

Once we know what coursebooks are being used, we need to take a sample, based on popularity. We are taking the four most popular coursebooks for in-depth analysis. However, there are a few things that we would need to address before that, by looking at the numbers. It seems that the popularity of Burlington Books in Navarra is soaring. Established in 1994, this publisher, ‘became the first publisher specialising in English textbooks to publish materials specifically designed for the education system in Spain’¹². Their success is based on a good sales team, and the reedition of their most popular coursebooks with an addition in the title, such as ‘New’ or ‘Advanced’. After this publisher, there is a draw in coursebook use, and three of the other major ELT publishers in Spain, Cambridge, Macmillan and Oxford are used twice each.

For analysis we will take the most popular coursebook, *New English in Use 2*, together with its sister series *English in Use 2* and *Advanced English in Use 2*. The contents presented are the same, and in the case of the ‘Advanced’ there is an extra section aimed at exam preparation of KET, PET and FCE and integrated language skills development. We will also take the second most popular, Oxford’s *Mosaic 2* and the third most popular, *Advanced Real English 2* together with its sister series *Real English 2*.

It is interesting to note that by searching on the publishers websites, there is not much information available about each of the coursebooks in terms of English variety and whether the coursebook is corpus-informed. Only Macmillan and Cambridge specify that the English variety is British, but only for one of their two coursebooks used. Then, Burlington Books has a series titled *Burlington International English*, but it is not clear

¹² <https://www.burlingtonbooks.com/Spain/Page.aspx?PageID=382&zoneIndex=0&subMenuIndex=2>

whether the title is making reference to the type of English taught or it is just the chosen name for a title. Regarding whether the materials are corpus-based, only Cambridge asserts that their coursebooks are corpus-informed. It was necessary to clarify this, so we sent an email to the other main publishing houses used in the ELT materials in Navarra (Oxford, Burlington Books and Macmillan). We asked three questions:

- In order to elaborate the texts for ELT materials and select the target vocabulary for each didactic unit, do you take a corpus-based or corpus-informed approach?
- If that is the case, which corpus is used and for what ELT materials?
- If that is not the case, what criteria are followed in order to select the target lexical items in each unit?

Oxford University Press gave the following answer to the questions:

- The syllabus (English as a First Foreign Language in CSE)
- CEFRL
- Oxford 3000 y Oxford 5000

Burlington Books offered a detailed answer to the questions, which in summary would read as follows: they use the official syllabus by the Ministry of Education, which prescribes in detail the linguistic contents for each year. As some topics and lexical fields are seen in different years, a careful selection is required. For this selection they use references as *www.EnglishProfile.org* for *CEFR*, and online dictionaries such as Cambridge, which assign CEFR levels to lexical terms. They continue to state that these levels can only be taken as an orientation, as we frequently find ‘difficult’ English words (from a monolingual perspective) associated to a high-proficiency level that are cognates for Spanish English learners, so therefore not that difficult but certainly of lower frequency in English. This factor is not always contemplated in the corpora.

For the time being, the email that was addressed to the other publishing house, Macmillan, has not received an informative answer, although a quick answer was received informing of its being forwarded to the head of editors.

We then proceeded to explore, hands-on, the ELT material samples. Exploring them closely, we can conclude that they are based on a syllabus that constitutes a mix of a situational syllabus but based around structural patterns and lexical topics. The didactic

units in all of the coursebooks are organized around topics based on ‘real’ life situations, supposed to be appealing for teenagers in 2nd year CSE. A selection of target lexical items is presented, and then, they are presented in context by means of written or oral texts. Some of these texts have cultural references to NS countries or people. Few oral comprehension texts present NNS; neither present NS of a non-standard variety. Following the presentation, a structural grammar pattern (or two) is presented as well, and then the learners are left to practice with the vocabulary and structures. The final aim of each unit is to produce a written or oral text. However, production is not free, as the meanings need to be constructed around the structures and the vocabulary that the learners were taught.

3.3 Data collection

We now proceed to examine the vocabulary contents that are presented in each of the coursebooks. These contents are presented under topics that generally correspond to a semantic field, such as *The family*, *Clothes*, *Places around town*, etc. The titles of the topics are self-explanatory, and by looking at them teachers can easily be informed about the lexical field that will be dealt with in that section. However, we can find topics as *Adjectives*, *Verbs* and *Plurals* that are somehow confusing, as they do not make reference to a semantic field but to a part of speech and therefore do not facilitate the teachers’ job as to give information regarding the type of adjectives or verbs that are presented (only one coursebook adds a subtitle).

The titles given for the vocabulary sections are shown in the table below:

English in Use 2	Real English 2	Mosaic 2
Activities	Adjectives [1]	Adjectives: feelings
Adjectives [1]	Adjectives [2]	Adjectives: feelings and qualities
Adjectives [2]	Animals	Clothes
Adjectives [3]	Appliances and musical instruments	Collective nouns
Adjectives [4]	Clothes	Containers
Animals	Competitions	Everyday items
Crime	Emotions	Holidays
Experiences	Family	Jobs
Fashion	Food	Kitchen gadgets

Fitness	Free time activities	Life stages
Geography	Jobs	Like, prefer, can't stand
Life events	Life events	Materials
Nutrition	Parts of the body	Money
People and crime	Places around town	Plurals
Places around town	Prepositions of Place	Street objects
School	School	The body
Sport	Things we do	The environment
The family	Transport	TV programmes
The house	Travel Items	Verbs [1]
The weather	Types of music	Verbs [2]
Vehicles	Verbs [1]	Verbs and prepositions of movement
Verbs [1]	Verbs [2]	Verbs: lifestyle
Verbs [2]	Weather	

Table 2

There are few topics that are presented in all three coursebooks. Topics such as *Clothes*, *Animals*, *(The) Family*, *School*, *(The) Weather*, *Vehicles/Transport*, *Activities/Free time activities*, *Jobs*, occur in just two of the course books. Initially, there would seem that there is no agreement as to what lexical contents to teach in 2nd year CSE, because there is no agreement in the topics, increasing the confusion of the teacher and the learner. (Discuss)

In order to make a quantitative analysis about the lexical terms that are taught the most frequently in 2nd CSE, we will choose a topic that is present in all three coursebooks and write down all the target lexical items that are presented for the student to learn. These words, as we already mentioned, are presented – preferably in context – in the didactic units by their appearance in the texts, dialogues and grammar activities in the coursebook. Quite often, the texts are adapted to present higher occurrences of the target lexical items, resulting in a higher lexical density of the target terms, very typical of sample text writing for ELT purposes in which a particular structure or lexical item is replicated in a higher proportion than normal (Gabrielatos 2005). However, we will not focus on that or attempt to explore lexical density on the texts presented in the coursebooks, but on the terms introduced in the wordlist as target lexical words. The objective is that by considering the most frequent words in the most frequent topics in the most frequently used coursebook in Navarra, we will find those words that are quite similar to what we could call a ‘core vocabulary’ for students in 2nd year CSE, and which students will be most likely to

encounter in their interactions in English, be them online, face-to-face, fact or fiction, in written or spoken mediums, and in examinations.

Next, we will observe and count which of these target words occur in all three samples of the coursebooks, in two of them, or if it is just in one. The only transparent topic present in all the coursebooks is that of *Life events/Life stages*. (See Table 2). The other two common topics, of a more darker nature, make reference to *Adjectives* and *Verbs*. We will start by counting the words of the topic of *Life events/Life stages*. (See Table 3)

Present in coursebooks (min. 1, max. 3) ¹³	Life events/life stages
3	get a job
3	get married
3	have a child/children
3	move to/move/move house
2	(be) born
2	die
2	fall in love
2	get divorced
2	go to university
1	baby
1	become
1	belong to
1	buy a house
1	child
1	elderly
1	graduate from
1	grow up
1	have a boyfriend/girlfriend
1	join

¹³ For details as to which coursebook presents each of the terms, see [Annex 3](#)

1	learn to drive
1	leave home
1	meet
1	middle-aged
1	receive awards
1	retire
1	study
1	teenager
1	toddler
1	young adult

Table 3

After that, we will write down and count the samples of the *Adjectives* topics. We should point out that the number of items is not balanced, with one of the coursebooks presenting 4 *Adjectives* sections and the other two coursebooks presenting 2 sections each (See Table 4). One of the latter includes a subtitle to clarify that the adjectives dealt with are feelings and qualities. Nevertheless, we will proceed to write down all of them and count, again, the ones that are present in all three coursebooks, in two of them or just in one.

Present in coursebooks (min. 1, max. 3)	Adjectives
3	calm
3	clever
3	funny
2	angry
2	brave
2	cheap
2	colourful
2	curly
2	dark
2	dull
2	expensive

<i>Cont.</i> Present in coursebooks (min. 1, max. 3)	Adjectives
1	hard
1	heavy
1	height
1	helpful
1	honest
1	hungry
1	ideal
1	impatient
1	isolated
1	lazy
1	lonely

2	fat
2	friendly
2	good-looking
2	light
2	lively
2	long
2	old
2	peaceful
2	plain
2	pleasant
2	rude
2	serious
2	shy
2	straight
2	strong
2	surprised
2	weak
2	worried
2	young
1	adventurous
1	afraid
1	asleep
1	athletic
1	attractive
1	beautiful
1	bored
1	bright
1	busy
1	casual
1	comfortable
1	cool
1	creative
1	crowded
1	elaborate

1	lovely
1	mean
1	medium
1	modern
1	nasty
1	nervous
1	noisy
1	old-fashioned
1	ordinary
1	patient
1	polite
1	positive
1	practical
1	pretty
1	proud
1	relaxed
1	relaxing
1	ridiculous
1	risky
1	rocky
1	rough
1	round
1	safe
1	sandy
1	scared
1	second-hand
1	shady
1	short
1	sick
1	silly
1	small
1	soft
1	sophisticated
1	tall

1	elegant
1	embarrassed
1	energetic
1	excited
1	exhausted
1	extraordinary
1	fair
1	fashionable
1	formal
1	frightened
1	frightening
1	generous

1	terrible
1	thin
1	thirsty
1	tired
1	trendy
1	unattractive
1	unique
1	unpleasant
1	unusual
1	upset
1	wonderful

Table 4

The final section of the data collection consists in observation of the cultural pages of the coursebook to find out which English speaking countries are presented as a model. We will also take a sample of an activity dealing with receptive skills in order to analyze how context is presented and analyse it qualitatively.

In the coursebook *English in Use 2* (with its ‘New’ and ‘Advanced’ counterparts), there are some references to the most commonly presented English varieties in the vocabulary sections. The section is called *Everyday English* and it presents a contrast of British/US words related to where we live in the first unit, British/US words related to fashion and shopping and British/US words related to food items.

In the section that is presented as *CLIL*, there are a series of references to cultural aspects and traditions. For example, the texts present the following topics:

- Multi-generational family, India and Africa
- Making champions in China
- Working children and the law: Britain
- JRR Tolkien: Literature
- Amish: USA and Canada
- Unusual Uniforms: The Queens' Guards, the Yeomen Warders, Scotsmen

In the coursebook *Real English 2* there is no specific reference to different English varieties. Together with its sister series, *Advanced Real English 2* it presents a section called *Real World Extra* where different cultural events, traditions, literature and competitions are presented, based on different English speaking countries from Inner and Outer Circles. The topics of the texts presented are:

- Extraordinary Schools: USA, England, Australia, in Victorian times
- Nature in the city: USA, New Zealand, England, South Africa
- Little Women: Literature
- Man-made wonders: Australia, South Africa, USA, UK
- World Sporting Events: England, Scotland, USA
- Traditional Musical Instruments: Australia, India, Scotland, the Beatles
- Myths, Legends and Folktales: Australian, British
- In a section called *Real life* there is a family tree of the British Royal Family

In the coursebook *Mosaic 2* there is no reference to different varieties of English. There is a section of *Cultural Pages* where the common header is *Around the world*. The texts have the following titles:

- Unusual lessons in India and Scotland
- An important discovery in the North Atlantic
- A school science fair in South Africa
- Money traditions in the Solomon Islands
- A talent show in the UK
- Talking about a dangerous job in India
- A legend from New Zealand
- Neighbourhoods in Canada, Wales and Ghana
- An Australian Festival

In the image below, there is a sample of an activity of oral comprehension in one of the coursebooks. (Marks and Addison 2010). It is a Listening activity with 3 exercises, a short question of multiple choice, a question consisting of a gap fill, and two open questions.

Listening



- 4. Michelle and Jamie are each presenting a biography in class. Listen to their presentations. Which people from the list below don't they mention in their presentations?**

Michelle:

1. Michael Douglas
2. Kirk Douglas
3. Michael Keaton
4. Catherine Zeta-Jones

Jamie:

5. Jamie's mother
6. Jamie's grandmother
7. Jamie's grandfather
8. Jamie's father



- 5. Listen again. Copy and complete the sentences.**

1. Michael Douglas was born in
2. Douglas won an Oscar in
3. Douglas has got the same birthday as his
4. Marie is years old.
5. Marie met her husband at a
6. Marie and her husband lived in Ontario on a

- 6. Can you remember the following details?**

1. How many children have Michael Douglas and Catherine Zeta-Jones got?
2. How old was Marie when she met her husband?

After listening to the recording, we can conclude that the accent is British and the context is not real, but created specifically for the audio recording. There is a high density of the lexical terms that were presented as a target, those entitled *Life events*. There are two cultural references; one to some actors that the students probably do not know about, and the second one to Ontario (Canada). This is the listening script, with the target words underlined:

Teacher: OK, Michelle let's hear your report.

Michelle: OK, umm – my project's about my favourite actor, Michael Douglas. He's the son of another famous actor from the 50s and 60s, Kirk Douglas. Michael Douglas was born in 1944 in New Jersey in the USA. He studied at Eaglebrook School in Massachusetts. He decided to be an actor when he was a teenager, but his father didn't want him to become an actor. His best films are *Romancing the Stone*, *The American President* and *Wall Street*. He won an Oscar for his performance in *Wall Street* in 1987. Michael Douglas met Catherine Zeta-Jones, a famous and beautiful actress, in 1998 at a film festival in France and they got married in 2000. They've both got the same birthday, 25th September. They ve got two children.

Teacher: Thank you, Michelle – That was very interesting. OK, Jamie, it's your turn.

Jamie: Err, my project isn't about a famous person, it's about my grandmother, Marie. She's my mother's mother. Grandma Marie is 62 years old. She **was born** here, in Montreal, Canada. She **met** my grandfather at a party when she was 20 years old. My grandfather came to the party with another girl, but when he saw Marie, he **left** the other girl and danced with Marie all night. When my grandparents **got married** a year later, Grandma **moved** to my grandfather's farm in Ontario. There were cows on the farm and my grandfather taught Grandma Marie how to milk cows. She also **studied** farming at the Ontario Agricultural College so she really learned how to manage a farm. When my grandfather **died** five years ago, she decided to manage the farm alone.

This other example of an activity in the same coursebook, presents a Speaking activity for oral production. Learners are expected to produce the target words from the topic and to produce accurate grammatical structures and practice the pronunciation of the past simple regular verbs. The model dialogue presented is one between teacher and learner, and there is a cultural reference to Charles Chaplin.

Speaking

7. Brian researched a famous person for a school project. His teacher is asking him questions. Copy and complete their dialogue with the words below. Then practise the dialogue with a partner.

did he have • born • die • get married • job

Teacher: Where was Charlie Chaplin ¹ ?

Brian: In London.

Teacher: What was Charlie Chaplin's ² ?

Brian: He was an actor and a film maker.

Teacher: When did he ³ ?

Brian: Chaplin got married four times. He got married to his last wife, Oona, in 1943.

Teacher: How many children ⁴ ?

Brian: He had 11 children.

Teacher: When did he ⁵ ?

Brian: In 1977. He was 88 years old.

8. Use questions from Exercises 3 and 7 to ask and answer questions about your partner's grandparents or great-grandparents.

Where was your grandmother born?

She was born in Argentina.

Say It Right!
Listen and repeat.

1. moved /d/
2. watched /t/
3. wanted /ɪd/

The last example presents the target words in context in another coursebook (Marks and Addison 2014) and encourages students to connect the personality definitions with the real life people in the pictures. These personalities have verified origins, but the British cultural point of view is the one that comes out on top, with references to The Beatles, JK Rowling, Prince William, Henry VIII and Egyptian Pharaoh Tutankhamun, whose tomb

was precisely discovered by a British explorer after years of a British-founded expedition campaign in Egypt, former British colony. The only non-British personality here is Antoni Gaudi. Nevertheless, it should be reminded that this is just a sample of an activity in the coursebooks dealing with the target vocabulary studied, and this British-biased tendency might just be a matter of coincidence.

Vocabulary

- ▶ **Listen and repeat the words and phrases in colour in the quiz below. Then match the people to their biographical information.**

Who Are They?

- 1** This famous musician sang and played the guitar in a very popular band – the Beatles. He received many awards for his songs during his lifetime. He got married twice – the second time to Yoko Ono. They had a child in 1975. It was his second son.
- 2** This boy grew up in Egypt many years ago. When he was nine years old, he became the Pharaoh. He was the Pharaoh for only ten years and then he died.
- 3** This British author comes from Wales. When she finished university, she got a job as a teacher. In 1990, she moved to Portugal. She created the character Harry Potter.
- 4** This famous Spanish architect was born in Reus, Catalonia. He graduated from the Barcelona School of Architecture in 1878. He designed and created the Sagrada Família, but he didn't finish it.
- 5** This British prince went to university in Scotland. There, he met and fell in love with his wife, Kate Middleton. He joined the army in 2006. His son, Prince George, was born in 2013.
- 6** This British king lived from 1491 to 1547. He belonged to the Roman Catholic Church. In those days, Roman Catholics didn't get divorced, but this king did. He got divorced twice, and married six times!

▶ Answers, page 58



3.4 Learner Corpus and Native Speaker Corpus

In this section we are going to explore a learner corpora, more precisely the Open Cambridge Learner Corpus (Uncoded) to search for the frequencies of the words in the word list extracted from the coursebooks in the previous section. We expect that some of these terms will be quite common, and due to time constraints we will only search for the lexical terms that appear in all three coursebooks and those that appear only in two of them, adding to a total of 39 lexical items.

After looking up the occurrences in the Open CLC, we will compare them with the occurrences in the BNC to revise and explore a native English corpus comparing the same selection of lexical items most frequent in the coursebooks. Initially, we considered the possibility of removing some of the domains from the corpus that represent different genres. However, in the end it was decided not to do so, as we expect that the vocabulary that ELT materials teach for general EFL are the lexical items that the learners will find the most in real life, be it spoken, written informative, academic or fiction. See the results on Table 5 below.

Coursebook's presence (min. 1, max. 3)	Coursebooks' most frequent target words	CLC Hits per million	CLC percent of whole corpus	BNC Hits per million	BNC percent of whole corpus
3	get a job	26.27	0.00300	5.52	0.00055
3	get married	44.19	0.00440	10.39	0.00100
3	have a child/children	22.09	0.00226	7.26	0.00073
3	move to/move/move house	223.92	0.02266	378.39	0.03797
3	calm	34.93	0.00350	12.65	0.00130
3	clever	22.99	0.00230	21.02	0.00210
3	funny	86.28	0.00860	40.03	0.00400
2	(be) born	70.16	0.00700	48.5	0.00490
2	die	128.68	0.01300	195.01	0.02000
2	fall in love	25.38	0.00250	7.7	0.00077
2	get divorced	3.88	0.00039	0.47	0.00005
2	go to university	13.44	0.00130	1.19	0.00012
2	angry	55.53	0.00560	35.57	0.00360
2	brave	13.73	0.00140	14.24	0.00140
2	cheap	151.37	0.01500	57.05	0.00570

2	colourful	22.39	0.00220	9.7	0.00097
2	curly	2.09	0.00021	3.46	0.00035
2	dark	68.67	0.00690	96.35	0.00960
2	dull	24.18	0.00240	15.48	0.00150
2	expensive	217.35	0.02200	50.85	0.00510
2	fat	42.99	0.00430	27.26	0.00270
2	friendly	194.66	0.01900	33.93	0.00340
2	good-looking (also non-hyphenated)	9.55	0.00095	4.57	0.00046
	good looking	4.18	0.00042	1.3	0.00013
2	light	41.5	0.00420	63.87	0.00640
2	lively	57.03	0.00570	13.05	0.00130
2	long	454.12	0.04500	363.53	0.03600
2	old	723.42	0.07200	515.7	0.05200
2	peaceful	29.86	0.00300	14.19	0.00140
2	plain	7.17	0.00072	25.64	0.00260
2	pleasant	87.48	0.00870	23.04	0.00230
2	rude	30.15	0.00300	8.4	0.00084
2	serious	117.63	0.01200	106.13	0.01100
2	shy	21.8	0.00220	7.93	0.00079
2	straight	8.36	0.00084	32.07	0.00320
2	strong	171.97	0.01700	169.74	0.01700
2	surprised	5.08	0.00051	5.4	0.00054
2	weak	30.15	0.00300	39.24	0.00390
2	worried	10.45	0.00100	6.6	0.00066
2	young	581.3	0.05800	314.75	0.03100

Table 5

3.5 Results analysis

In this section we will proceed to present figures to compare, contrast and analyse the data obtained from the corpora searches. In the Open CLC, 'have' is the second most used verb (after 'be'), with a frequency of 40,069 and 'get' is the 10th most used verb, with a frequency of 6,517. The frequency of 'move' goes down to display a frequency of 691 occurrences. However, the search offers interesting insights, as 'have a child/children' is not a very frequent collocation. Of the two collocations with 'get', 'get

married’ is more frequent than ‘get a job’, but still they are less frequent than the use of ‘move’ as a verb meaning ‘to relocate’. By briefly considering the data we have to confirm that most of the vocabulary items taught to learners by the coursebooks seem quite common in this corpus, and indeed the learners show these are part of their active vocabulary as they use it in their production.

However, if we compare the normalized frequency per million in both corpora, we can really start to see some differences in frequency of occurrence.

Open CLC	Hits per million
old	723.42
young	581.3
long	454.12
move to/move/move house	223.92
expensive	217.35
friendly	194.66
strong	171.97
cheap	151.37
die	128.68
serious	117.63
pleasant	87.48
funny	86.28
(be) born	70.16
dark	68.67
lively	57.03
angry	55.53
get married	44.19
fat	42.99
light	41.5
calm	34.93

Table 6

BNC	Hits per million
old	515.7
move to/move/move house	378.39
long	363.53
young	314.75
die	195.01
strong	169.74
serious	106.13
dark	96.35
light	63.87
cheap	57.05
expensive	50.85
(be) born	48.5
funny	40.03
weak	39.24
angry	35.57
friendly	33.93
straight	32.07
fat	27.26
plain	25.64
pleasant	23.04

Table 7

The top 20 words from the coursebooks in the Open CLC can be seen on Table 6, on the two left columns, organized by higher to lower frequency. Most of these words are indeed words coming from two of the coursebooks or the three of them, so they are taught quite a lot. That means that coursebook writers consider these words to be quite frequent and useful for learners, and indeed learners make use of them, as the corpus evidences.

Let's now consider the two right columns on Table 7, corresponding to the top 20 words from the coursebooks but this time in the BNC corpus. Their frequency is also ordered from higher to lower. In the top 4 we can see that there is almost a perfect match with the BNC data. Both in the Open CLC and in the BNC the most frequent words of the ones explored are 'old', 'young', 'long' and 'move/move to/move house'. The rest of the most frequent words are almost exactly the same in both corpora. There are 16 matches in total counting the first 4 already mentioned, although with slight variabilities in the order or appearance. See Table 8 for a numbered comparison side to side.

	Open CLC
1	old
2	young
3	long
4	move to/move/move house
5	expensive
6	friendly
7	strong
8	cheap
9	die
10	serious
11	pleasant
12	funny
13	(be) born
14	dark
15	lively
16	angry
17	get married

	BNC
1	old
2	move to/move/move house
3	long
4	young
5	die
6	strong
7	serious
8	dark
9	light
10	cheap
11	expensive
12	(be) born
13	funny
14	weak
15	angry
16	friendly
17	straight

18	fat
19	light
20	calm

18	fat
19	plain
20	pleasant

Table 8

As mentioned before, in the top 20 list that we analysed, there are 17 matches between the Open CLC and the BNC. The three terms that are unique to the Open CLC (but still were fairly frequent in the coursebooks and therefore were part of the 39 items research in the corpora), are ‘lively’, ‘get married’ and ‘calm’. The three terms in the top 20 that are exclusive to the BNC are ‘weak’, ‘straight’ and ‘plain’. ‘Straight’ in the sense that was used in the coursebooks in that particular vocabulary section referred to the sense of ‘not curved or wavy’, so the actual instances of production in which that adjective can be used are quite restricted. However, in the BNC, the senses of ‘straight’ were more varied, meaning ‘properly positioned so as to be level, upright, or symmetrical’, ‘in proper order or condition or ‘not evasive; honest, clear and logical’.

To finish this section, it is necessary to go back to our exploration of the three most popular ELT coursebooks in Navarra and look again at the words which were presented in all of them, under the lexical topics of ‘Life Events/Stages’ and ‘Adjectives’. These were the words common to all three coursebooks:

move to/move/move house
funny
get married
get a job
have a child/children
calm
clever

As we saw earlier, ‘move to/move/move house’ is in top 5 position in both the Open CLC and BNC. However, this data has to be taken with caution. ‘Move’ is a very polysemous verb in itself, and some of the samples from the keyword search do not mean ‘to relocate or change place of residence’, which is the sense that the word is given in the coursebooks, but it can also mean ‘to make progress’ or simply ‘to go in a specified direction’.

‘Funny’ is an adjective that is in the top 20 in both corpora. It is quite frequent then. However, in the coursebooks, this term was presented in the sense of ‘causing laughter or amusement’, when it has a much more common sense ‘difficult to understand, strange or curious’. It is interesting to point out that this adjective is frequently confused by learners with ‘fun’, and that the latter sense of the adjective is acquired much later, which seems to be at least as common as the first one in the BNC. We would need to look further into the concordances to test this.

‘Get married’ is present in the top 20 in the Open CLC only with 44.19 occurrences per million. In the BNC it falls in 27th position, with just 10.39 occurrences per million. ‘Calm’ is only present in the Open CLC in 20th position, with 34.93 occurrences per million whereas it is not in the top 20 in the BNC, where it falls to position 26th with a 12.65 incidence per million.

The remaining lexical terms, ‘get a job’, ‘have a child/children’ and ‘clever’, which were introduced in all three coursebooks and therefore would be expected to be much more common in actual use, are not happening in the top 20 in neither the Open CLC or the BNC, although they are quite close.

		Open CLC
4	move to/move/move house	223.92
12	funny	86.28
17	get married	44.19
20	calm	34.93
24	get a job	26.27
27	clever	22.99
29	have children	22.09

		BNC
2	move to/move/move house	378.39
13	funny	40.03
21	clever	21.02
26	calm	12.65
27	get married	10.39
32	have children	7.26
33	get a job	5.52

Colour code:

Both in CLC and BNC
 Only in top-20 in CLC
 Below top-20 in BNC
 Below top-20 in both CLC and BNC

4. Discussion

We have seen that the frequency of the words analysed is quite similar in both corpora, with the same four terms in the first four positions. Comparing the top-20 words, there are only slight differences of frequency. And three of the most frequently taught words in the coursebooks did not make it to the top-20 in any of the corpora analyzed.

We are fully aware that the context where corpora samples are taken is very different, in the case of the Open CLC and the BNC. There is nothing similar in the context of a corpus informed by learners' production in ESOL examinations and the context of a native corpus, where 'real' life is portrayed by the informants. By examining the omissions in the BNC which are present in the Open CLC and vice versa, it would seem that the learners that have informed the CLC are more positive in their choice of adjectives, and mention 'getting married', 'lively' and 'calm' a lot more than it is mentioned in 'real' native texts. The native speakers that have informed the BNC, seem to be a little more pessimistic in their choice of adjectives, showing a higher use of 'weak', 'straight', and 'plain'. This can just be explained by context, as speakers whose text samples form part of the CLC were taking an examination, the real-life context is set aside and an ideal context is imagined, whereas in the BNC, the contributions are real and therefore context has a direct effect on them.

As O'Keefe et al. (2007) argue, the examples in ELT should be corpora-based in order to model and exemplify the language for learners. These models should be 'authentic rather than contrived', and for contrived they understand those models of language that are specifically created to illustrate a particular selection of lexical items or language form.

When we looked at the samples of the activities that we presented, we can see that they have been created for that linguistic situation. The Listening comprehension presents most of the target words of the *Life stages* topic and it is lab-recorded. The context presented is a familiar context for the learners, a class situation where two students give a presentation of a famous person. The first student gives a speech about her favourite actor, Michael Douglas. This cultural reference will not be understood by most students in 2nd CSE today. That would mean that some of them disconnect out of lack of interest before they listen to the first 20 words. The second presenter in the Listening gives a

speech about his grandmother. This is more interesting to listen to. There are no cultural references that might get lost in translation, and some students might find this more interesting too.

This Listening comprehension could be a good option to present the target vocabulary in a condensed way, but it would require changing the cultural references to some other ones more relevant for teenagers today. However, the Speaking activity that is presented in the next page, is not a good idea. We have again a teacher-student interaction in a familiar context for the learners, but it is so artificial that it is barely recognisable as a realistic classroom interaction. I think that a better conversational example could have been found in a corpus, as we have seen these words are quite frequent in actual use by NS. In addition, it is an oversimplified type of interaction, and it could not qualify as a spontaneous natural conversation anywhere in real life, in a native or non-native context.

Finally, we will discuss the variety of English taught in the coursebooks. As we initially thought, the variety being taught aims to be a British variety, although none of the coursebooks analysed states it clearly. There are references in one of them to the British and US varieties from a lexical point of view. All of them mention CEFR of languages as a level-indicator and those with the 'Advanced' addition to the title include exam practice for Cambridge ESOL examinations, a very popular examination in Europe for EFL learners, as we know.

The coursebooks seemingly make the effort to present a varied array of cultural references, but the prevalence of a British tendency is undeniable. As we established in the hypothesis, the native varieties portrayed in textbooks seem to be restricted exclusively to British and American ones. However, the cultural references in the ELT materials studied are mostly British, as well as the accents in the oral texts (Listenings). Surprisingly, we were wrong in considering that the American English variety and culture would be portrayed, as there are neither many cultural references about the USA, nor General American accent.

In one of the coursebooks there are references to many countries of the outer circle (ex-British colonies, see Kachru 1985 in Jenkins 2007), but curiously enough, there is no reference to the USA at all. The fact that the coursebooks are edited in Europe might

justify this tendency to portray more UK-based themes. However, there is no need to exclusively select English-speaking countries to portray a historical or cultural fact of interest to learners. As we saw, English can be the medium to present any kind of cultural content, anywhere in the world, not necessarily Native English speaking countries or countries from the outer circle. The use of real texts of an ELF variety, what Cortazzi and Jin (1999) term *international target culture materials* (in Rubdy and Saraceni 2006: 128) could be the perfect medium to introduce this variety and present cultural content from different peoples and places in the world in a non-native English context, and not the prevailing Western European tradition. These texts could include samples of different NNSs using varied lexical terms, portraying different accents and phonological variation and using cross-cultural pragmatic rules.

Some of the difficulties faced in this project were time constraints. We could have benefited from further research and extended the investigation to other regions in Spain and even other countries in Europe to see whether the coursebooks being used belong to the same four popular ELT publishers.

Additionally, it would have been interesting to build a wordlist including all the target lexical items for all the coursebooks in order to analyze all possible matches, and not limiting it to the three only common topics of 'Adjective' and 'Life events/stages'. With this wordlist, we could compare it fully to the BNC and see whether the most frequent words are those actually being taught. We could also compare it with the existing wordlists for learners Oxford 3000 and 5000.

Finally, this research could be extended by introducing other corpora for comparison, such as the Corpus of Contemporary American English (COCA), to see if there are any significant differences of frequency of use of the words and maybe find more information about the English variety taught in the coursebooks (which in most of them is not specified) and introducing an ELF corpus, such as the VOICE.

5. Conclusion

We have revised a selection of EFL teaching materials used in the region of Navarra with the objective of investigating the English varieties that are taught in the region of Navarra. This has been carried out by looking at the most frequent lexical items presented in the most popular EFL coursebooks in the region and contrasting them with the lexical items used by learners and with those actually used by native English speakers by means of concordance searches in the Open CLC and BNC respectively.

English is a worldwide language used for communication among individuals who do not share a common L1, so it has been established as a *lingua franca* of sorts where it is used to communicate among NNS of English. This *lingua franca* differs in certain aspects with the (sometimes) prescriptive structural patterns offered in EFL coursebooks. But similarly, the EFL that is taught by means of the coursebooks explored is a little different to the one used by ‘real’ native speakers.

The major ELT material publishers seem exclusively concerned with two varieties of English, British English and American English. However, there is no such thing as a single British English, as the varieties spoken in the UK are quite varied and rich in their diversity of accents and lexical variation. Therefore, the idea of making students fluent in English with a native speaker model in mind and offering exclusively spoken models with British accent is not a reflection of the actual linguistic variety of the UK and much less of the English spoken in the world. Additionally, this might seem discouraging for some learners, who feel that their accent is always going to be marked against what is considered the ‘native speaker’ standards.

Furthermore, ELT materials have a tendency to oversimplify dialogues and texts, categorize structures too prescriptively and present vocabulary in receptive texts with altered frequencies of lexical items or an artificial abundance of a specific grammar structure. All of this, embedded with a lack of observance of pragmatic features, sociolinguistic and cultural aspects, and sometimes with the popular cultural references of the time that go out of fashion quickly and dialogues that disregard natural speech (with its natural errors, false starts, hesitations and repetitions).

In order to investigate the simplification of texts, the cultural references to people and places where English is spoken and the alteration of texts to comply with the language

said to be taught in ELT materials, we have looked at the same selection of EFL coursebooks and taken a sample of three activities dealing with receptive skills (Listening and Reading) and one dealing with a productive skill (Speaking). The model texts were quite effective in their showcasing of the target lexical words and grammatical structures of the didactic units, but the contextual and cultural references were weak and uninteresting for teenagers as well as too forced and/or artificial.

All of the ETL materials used in Navarra make reference to CEFR levels of English in the European Framework and most of them make reference to Cambridge ESOL examinations and offer practice for them. This suggests how important it is for learners today to get a certificate in English due to the positive enhancement they offer to a *curriculum vitae*. However, we should consider the possibility of introducing examinations that judge learners' ability not comparing it to a native model, but rather comparing it to an ELF model for communicating with people from non-English speaking countries. It is important to remember that the ability to communicate goes beyond the native language, and is something related to a personal characteristic rather than nationality. Our model for testing learners should be those 'expert users' or 'successful users' of English (see section 2.5), who are not necessarily NS, but can communicate in all contexts, with control of pragmatics and aware of cultural aspects and differences.

In addition, we should reinforce the importance of Corpus Linguistics and SLA/ELT research working together. The previous issues found in ELT materials could be remedied if those materials were corpus-informed (or based). The same as dictionaries started to take their data from corpora three decades ago, now it is about time that ELT publishers took their data from corpora as well. The problem here would be to decide what corpora to use. There is no single answer. Undoubtedly, for learners that are learning to communicate, be autonomous and live in a globalized world not necessarily full of NS of English, the idea of an ELF corpus and ELF teaching is the perfect one. Three quarters of the English speakers today worldwide are speakers of English as a foreign language. So, why should we base our teaching materials on exclusive and restrictive native speaker (British) models based on intuitions? This way we could establish 'learner language'/EIL as the aim for universal English communication using corpora as a source for learning materials.

Another reason to support ELF as an acceptable language of communication among people from a whole selection of L1s, not just NS of English, is that ELF is richer, as the communication possibilities widen when we compare the number of NS of English to that of speakers of English as an L2. Unless students are settling down in countries from the Inner Circle where mirroring the local accent and being acquainted with the local customs might be essential for successful socialisation, students are more likely to be interacting with people from the expanding circle (countries where English is learnt and spoken). Therefore, if the goal of the learner is to settle down in an English speaking country, the *nativelike* English would be their preference, and therefore taught as a model of the use of grammar patterns that include chunks and hedging, lexicon making use of idiomaticity and lexical creativity, and phonological features distinguishing phonetic pairs beyond the ‘Lingua Franca Core’. However, presenting *nativelike* English as the only model and goal for EFL learners might not be the most adequate approach, as most surely in the English speaking country where they take up residence they will interact daily with NNS, as themselves.

In my teaching experience, I have seen that some EFL learners in CSE show rejection towards anything British or American (Brexit, Trump, etc); which might be one more reason to aim for ELF. The opposite is also true: good learners in upper secondary education show signs of not feeling able to mix with NS when they visit the UK, the USA or Ireland. They feel that their English, their grammar, vocabulary, or accent is not good enough and that the gap between foreign and native is insurmountable, no matter how proficient they might be.

As we tend to a more globalized world with international relations (Covid-19 permitting), the importance of English as a Lingua Franca will grow, and our attitudes will have to adapt to it. This is a field that could benefit from further exploration, together with the work carried out on ELT and Corpus Linguistics. Last but not least, this research has to reach the teaching community. Academic investigation is a magnificent tool for development, but if it does not get through to EFL teachers, it remains in ‘Academic limbo’. I would like to advocate for minimising the gap between linguistics studies and EFL teachers; that is why I took this Masters in Applied Linguistics. Conversely, researchers could help teachers immensely if they made their findings available to the teachers.

6. Bibliographical references

- Aijmer, K. (ed) (2009). *Corpora and Language Teaching*. Amsterdam: John Benjamins.
- Anderson, W., and Corbett, J. (2009) *Exploring English with Online Corpora*. 2nd ed. London: Palgrave.
- Biber, D. Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. London: Longman.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, U.K: Press Syndicate of the University of Cambridge.
- Frazier, S. (2003). A Corpus Analysis of Would-Clauses Without Adjacent If-Clauses. *TESOL Quarterly*, 37(3), 443-466.
- Gablasova, Dana, Brezina, Vaclav, & McEnery, Tony. (2017). Exploring Learner Language Through Corpora: Comparing and Interpreting Corpus Frequency Information. *Language Learning*, 67(S1), 130-154.
- Gabrielatos, C. (2005). *Corpora and Language Teaching: Just a fling or wedding bells?* Lancaster University
- Gobierno de España. 2020. BOE. RESOLUCIÓN de 3 de julio de 2020, de la Secretaría General Técnica, por la que se publica el Convenio con la Delegación en España de la Fundación British Council y la Comunidad Foral de Navarra, para la realización de proyectos curriculares integrados y actividades educativas conjuntas. Retrieved from <https://www.boe.es/boe/dias/2020/07/13/pdfs/BOE-A-2020-7774.pdf>
- Gobierno de Navarra. 2019. BON RESOLUCIÓN 8/2019, de 24 de enero, del Director General de Educación, por la que se establecen las bases para el desarrollo de los programas en lenguas extranjeras en Educación Secundaria: Secciones Bilingües y Secundaria Plurilingüe y por la que se convoca el programa de Secciones Bilingües de inglés, francés y alemán para la incorporación de nuevos centros públicos. Retrieved from <https://bon.navarra.es/es/anuncio/-/texto/2019/26/4/>
- Jenkins, J. (2006). Current Perspectives on Teaching World Englishes and English as a Lingua Franca. *TESOL Quarterly*, 40(1), 157-181. doi:10.2307/40264515

- Jenkins, J. (2007) *English as a Lingua Franca: Attitude and Identity*. Oxford: Oxford University Press
- Jenkins, J. (2009). *English as a lingua franca: Interpretations and attitudes*. *World Englishes*, 28(2), 200-207.
- Kirkpatrick (2007). *World Englishes : Implications for international communication and English language teaching*. Cambridge: Cambridge University Press
- Marks, L. and Addison, C. (2010) *Real English ESO 2 Student's Book*. Limassol (Cyprus): Burlington Books
- Marks, L. and Addison, C. (2014) *English in Use ESO 2 Student's Book*. Limassol (Cyprus): Burlington Books
- Meunier, F. (2002) ‘The Pedagogical Value of Native and Learner Corpora in EFL Grammar Teaching’, in S. Granger, J. Hung and S. Petch-Tyson (eds) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam: John Benjamins. 119–41.
- O’Keeffe, A. and McCarthy, M. (eds.) (2010) *The Routledge Handbook of Corpus Linguistics*. 1st ed. London; New York, NY: Routledge. Routledge Handbooks in Applied Linguistics.
- O’Keeffe, A., McCarthy, M. and Carter, R. (2007) *From Corpus to Classroom : Language Use and Language Teaching*. Cambridge: Cambridge University Press.
- OpenCLC (v1). 2017. Distributed by Lexical Computing Limited on behalf of Cambridge University Press and Cambridge English Language Assessment.
- Pelteret, C. (2014) *Mosaic 2 Student's Book*. Oxford: Oxford University Press
- Rubdy, R., & Saraceni, M. (eds.) (2006). *English in the world: Global rules, global roles*.
- Schmitt, N. (2013). *An introduction to applied linguistics (2nd ed.)*. London: Hodder Education.
- Willis, D. (1990). *The Lexical Syllabus: A New Approach to Language Teaching*. London: Collins COBUILD.

7. Annexes

Annex 1

DATOS CENTROS CURSO ESCOLAR 2020-2021								
Nombre	A - Sección Bilingüe (Inglés)	A - Secundaria Pluri	D - Sección Bilingüe (Inglés)	D - Secundaria Pluri	G - Sección Bilingüe (Inglés)	G - Secundaria Pluri	Modelo A	Modelo A/G - British
IES Altsasu BHI	No	No	No	No	No	No	Sí	No
IES Barañain Alaiz BHI	No	No	No	No	No	No	No	No
IES Barañain BHI	Sí	Sí	No	No	Sí	Sí	Sí	No
IES Bera Toki-Ona BHI	Sí	No	Sí	No	No	No	Sí	No
IES Burlada Askatasuna BHI	No	No	No	No	No	No	No	No
IES Burlada Ibaialde BHI	Sí	Sí	No	No	Sí	Sí	Sí	No
IES Corella Alhama	No	No	No	No	Sí	No	No	No
IES Estella Tierra Estella BHI	Sí	Sí	Sí	Sí	Sí	Sí	Sí	No
IES Leitza Amazabal BHI	No	No	No	No	No	No	No	No
IES Lekaroz BHI	No	Sí	No	Sí	No	No	Sí	No
IES Lodosa Pablo Sarasate BHI	Sí	No	No	No	Sí	No	Sí	No
IES Marcilla Marqués de Villena BHI	Sí	No	No	No	Sí	No	Sí	No
IES Pam. Basoko BHI	Sí	No	No	No	Sí	No	Sí	Sí
IES Pam. Biurdana BHI	No	No	No	No	No	No	No	No
IES Pam. Eunáte BHI	No	No	Sí	No	No	No	No	No
IES Pam. Iturrama BHI	No	No	No	No	No	No	No	No
IES Pam. Julio Caro Baroja BHI	Sí	Sí	No	No	Sí	Sí	Sí	No
IES Pam. Mendillorri BHI	Sí	Sí	No	Sí	Sí	Sí	Sí	No
IES Pam. Navarro Villoslada BHI	Sí	Sí	No	No	Sí	Sí	Sí	No
IES Pam. Padre Moret Irubide BHI	Sí	No	No	No	Sí	No	Sí	Sí
IES Pam. Plaza de la Cruz BHI	Sí	No	No	No	Sí	No	Sí	No
IES Peralta Ribera del Arga BHI	Sí	No	No	No	Sí	No	Sí	No

IES San Adrián Ega BHI	Sí	No	No	No	Sí	No	Sí	No
IES Sangüesa Sierra de Leyre BHI	No	Sí	No	No	No	Sí	Sí	No
IES Sarriguren BHI	No	Sí	No	Sí	No	Sí	Sí	No
IES Tafalla Sancho III el Mayor BHI	No	No	No	No	No	No	Sí	No
IES Tudela Benjamín BHI	Sí	Sí	No	No	Sí	Sí	Sí	Sí
IES Tudela Valle Ebro BHI	Sí	Sí	No	No	Sí	Sí	Sí	No
IES Zizur BHI	Sí	Sí	Sí	Sí	Sí	Sí	Sí	No
IESO Aoiz DBHI	No	No	No	No	No	No	Sí	No
IESO Azagra Reyno de Navarra	No	No	No	No	Sí	No	No	No
IESO Berriozar DBHI	No	Sí	No	No	No	Sí	Sí	No
IESO Carcastillo Valle del Aragón DBHI	Sí	No	No	No	Sí	No	Sí	No
IESO Castejón DBHI	No	No	No	No	No	No	Sí	No
IESO Cintruénigo La Paz DBHI	Sí	Sí	No	No	Sí	Sí	Sí	No
IESO Cortes Bardenas Reales DBHI	Sí	No	No	No	Sí	No	Sí	No
IESO Doneztebe Mendaur DBHI	No	No	No	No	No	No	No	No
IESO Garralda DBHI	No	No	No	No	No	No	No	No
IESO Larraintzar DBHI	No	No	No	No	No	No	No	No
IESO Mendavia Joaquín Romera DBHI	No	No	No	No	No	No	Sí	No
IESO Noáin Elortzibar DBHI	Sí	Sí	No	No	Sí	Sí	Sí	No
IESO Ochagavia DBHI	No	No	No	No	No	No	Sí	No
IESO Pam. Iñaki Ochoa de Olza DBHI	Sí	Sí	No	No	Sí	Sí	Sí	No
IESO Pam. Iparralde DBHI	No	No	No	Sí	No	No	No	No
IESO Ribaforada El Cierzo	No	No	No	No	No	No	Sí	No
IESO Roncal DBHI	No	No	No	No	No	No	Sí	No
IESO Viana Del Camino DBHI	No	No	No	No	No	No	Sí	No
IESO Villava P. Atarrabia DBHI	No	No	No	No	No	No	Sí	Sí

Annex 2

DATOS LIBROS DE TEXTO CURSO ESCOLAR 2020-2021			
Nombre	Modelos A, D, G	Sección Bilingüe o Secundaria Pluri	British
IES Altsasu BHI	no ¹⁴	no	
IES Barañain Alaiz BHI	Advanced Real English 2		
IES Barañain BHI	New English in Use 2	New English in Use 2	
IES Bera Toki-Ona BHI	no	no	
IES Burlada Askatasuna BHI	Advanced Real English 2		
IES Burlada Ibaialde BHI	no	no	
IES Corella Alhama	no	no	
IES Estella Tierra Estella BHI	Action 2		
IES Leitza Amazabal BHI	New English in Use 2		
IES Lekaroz BHI	Mosaic 2	Mosaic 2	
IES Lodosa Pablo Sarasate BHI	no	no	
IES Marcilla Marqués de Villena BHI	New English in Use 2	Advanced English in Use 2	
IES Pam. Basoko BHI	Pulse 2	New English in Use 2	Burlington International English B1+
IES Pam. Biurdana BHI	New English in Use 2		
IES Pam. Eunáte BHI	Dynamic 2		
IES Pam. Iturrama BHI	Mosaic 2		
IES Pam. Julio Caro Baroja BHI	Think Ahead 2	Think Ahead 2	
IES Pam. Mendillorri BHI	New English in Use 2 Mosaic 2	New English in Use 3	
IES Pam. Navarro Villoslada BHI	Pulse 2	Beyond B1	
IES Pam. Padre Moret Irubide BHI	Real English 2	Real English 2	Advanced Real English 2
IES Pam. Plaza de la Cruz BHI	Real English 2	Advanced Real English 2	
IES Peralta Ribera del Arga BHI	New English in Use 2		
IES San Adrián Ega BHI	New English in Use 2	New English in Use 2	
IES Sangüesa Sierra de Leyre BHI	New English in Use 2	New English in Use 2	
IES Sarriguren BHI	Collaborate 2	Collaborate 2	
IES Tafalla Sancho III el Mayor BHI	English in Use 2		
IES Tudela Benjamín BHI	Mosaic 2	Burlington International English	Complete PET for Spanish Speakers

¹⁴ Information not available on the schools' websites

		B1	
IES Tudela Valle Ebro BHI	Advanced English in Use 2	Burlington International English B1	
IES Zizur BHI	Real English 2	Advanced Think Ahead 2	
IESO Aoiz DBHI	no	no	
IESO Azagra Reyno de Navarra	Mosaic 2	Burlington International B1	
IESO Berriozar DBHI	Mosaic 2	Advanced Think Ahead 2	
IESO Carcastillo Valle del Aragón DBHI	Think Ahead 2	Think Ahead 2	
IESO Castejón DBHI	Mosaic 2		
IESO Cintruénigo La Paz DBHI	Advanced English in Use 2	Burlington International English A2	
IESO Cortes Bardenas Reales DBHI	New English in Use 2	Advanced English in Use 2	
IESO Doneztebe Mendaur DBHI	Dynamic 2		
IESO Garralda DBHI	no	no	
IESO Larrainzar DBHI	New English in Use 2		
IESO Mendavia Joaquín Romera DBHI	Mosaic 2		
IESO Noáin Elortzibar DBHI	Action 2	Advanced Real English 2	
IESO Ochagavía DBHI	no	no	
IESO Pam. Iñaki Ochoa de Olza DBHI	Advanced English in Use 2	Advanced English in Use 2	
IESO Pam. Iparralde DBHI	New English in Use 2		
IESO Ribaforada El Cierzo	no	no	
IESO Roncal DBHI	no	no	
IESO Viana Del Camino DBHI	Mosaic 2		
IESO Villava P. Atarrabia DBHI	New English in Use 2	New English in Use 2	Burlington International English B1+

COURSEBOOKS	Number of schools using it
Action 2	2
Advanced English in Use 2	6
Advanced Real English 2	5
Advanced Think Ahead 2	2
Beyond B1	1
Burlington International English A2	1
Burlington International English B1	2
Burlington International English B1+	2
Collaborate 2	2
Complete PET for Spanish Speakers	1
Dynamic 2	2
English in Use 2	1
Mosaic 2	9
New English in Use 2	11
New English in Use 3	1
Pulse 2	2
Real English 2	4
Think Ahead 2	4

Annex 3

ENGLISH IN USE 2	REAL ENGLISH 2	MOSAIC 2
LIFE EVENTS	LIFE EVENTS	LIFE STAGES
belong to	become	baby
(be) born	(be) born	buy a house
die	die	child
fall in love	fall in love	elderly
get a job	get a job	get a job
get divorced	get divorced	get married
get married	get married	go to university
go to university	have children	have a boyfriend/girlfriend
graduate from	meet	have children
grow up	move	learn to drive
have a child	study	leave home
join		middle-aged
move to		move house
receive awards		retire
		teenager
		toddler
		young adult
ADJECTIVES 1-4	ADJECTIVES 1-2	ADJECTIVES: feelings and qualities
good-looking	dark	afraid
adventurous	tall	angry
beautiful	short	asleep
strong	clever	bored
curly	thin	brave
old	young	calm
straight	lazy	clever
clever	old	embarrassed
small	friendly	energetic
young	fair	excited
medium height	good-looking	friendly
brave	serious	funny
round	calm	helpful

funny	funny	hungry
athletic	curly	impatient
light	straight	lonely
fat	long	mean
weak	fat	nervous
creative	attractive	patient
honest	bright	polite
generous	cheap	positive
shy	dull	relaxed
long	elaborate	rude
dark	expensive	scared
casual	hard	serious
cheap	heavy	shy
colourful	light	sick
comfortable	ordinary	surprised
cool	plain	thirsty
expensive	soft	tired
fashionable	strong	worried
formal	unattractive	
lovely	unusual	
modern	weak	
old-fashioned	busy	
plain	colourful	
practical	crowded	
pretty	lively	
second-hand	noisy	
sophisticated	peaceful	
trendy	pleasant	
calm	rocky	
dull	sandy	
elegant	shady	
extraordinary		
frightening		
ideal		
isolated		
lively		
peaceful		

pleasant		
relaxing		
rough		
safe		
terrible		
wonderful		
angry		
exhausted		
frightened		
nasty		
proud		
ridiculous		
risky		
rude		
silly		
surprised		
unique		
unpleasant		
upset		
worried		

8. Resumen en español del TFM

Este trabajo se ha centrado en torno a la hipótesis de que la lengua inglesa que se enseña en Navarra es una variedad de carácter anglo-americano, más concretamente una variedad británica debido a la colaboración del gobierno de Navarra con Programas British y se ha basado en la literatura existente sobre lingüística de corpus, la enseñanza del inglés como lengua extranjera, evaluación lingüística, elaboración de materiales y programas didácticos y el inglés como lengua franca. Todo ello con el fin de testar la segunda parte de la hipótesis, que mantiene que las variedades británicas y americanas y su prevalencia cultural no tienen por qué ser las más idóneas para la enseñanza en educación secundaria obligatoria, teniendo en cuenta la mayor globalización a la que aspiramos y que la mayor parte de las personas que utilizan el inglés como lengua de comunicación diaria no son hablantes nativos de estas variedades. Hemos averiguado cuáles son los materiales de enseñanza más utilizados en Navarra para analizar la variedad de la lengua que presentan, la naturaleza de las referencias culturales y extraído los términos léxicos que más se enseñan en dichos materiales. De este modo, hemos podido constatar que la mayoría de referencias culturales son británicas, así como el acento de los textos orales. Además, hemos podido comparar de manera cuantitativa la frecuencia de los términos léxicos en un corpus de estudiantes de inglés (CLC) y en un corpus nativo (BNC) y determinar que si bien las frecuencias de uso son similares, el significado más utilizado por los nativos no es el primero que se enseña en los materiales, e igualmente el contexto de la situación de los informantes juega un papel importante en la selección de vocabulario. En el trabajo, mencionamos las dificultades encontradas y posibles ideas para ampliar la investigación. Concluimos que la enseñanza del inglés se vería beneficiada si los materiales presentaran elementos de referencia socioculturales más universales, estuvieran informados por *corpus* y menos centrados en el *nativismo*. Y además, presentar el modelo nativo británico o americano como única forma válida para los estudiantes de inglés no siempre puede resultar idónea. Por ello, sería interesante considerar las variedades del inglés de lengua franca y la existencia del ‘usuario experto’ como modelo de un hablante de éxito alcanzable y válido universalmente.