

A CORPUS-BASED STUDY OF EXAMPLES AND EXAMPLE MARKERS IN THE SPEECH AND WRITING OF SPANISH EFL STUDENTS

Paula Rodríguez-Abruñeiras

Universidad de Santiago de Compostela, Spain

Abstract: *The use of examples is a crucial aspect of human communication. However, they have attracted relatively little attention in the literature. With the aim of filling this gap, the current paper explores the use of examples and example markers (EMs) by Spanish EFL students, drawing on data from the Santiago University Learner of English Corpus (SULEC). The analysis considers both semi-fixed formulas and fully grammaticalized EMs. Results indicate a limited use of these expressions, with learners predominantly relying on fully grammaticalized EMs. The use of grammaticalized EMs reveals notable patterns, with the most common forms found in the corpus – for example and like – being prevalent across proficiency levels, although advanced learners exhibit a wider set of such markers. Moreover, women EFL learners seem to use EMs more frequently than men. The analysis also uncovers disparities between written and spoken language, with a greater presence of EMs in the written medium. Further investigation into the types of sequences in which the EMs for example and for instance appear reveals distinct patterns: argumentation predominates in written texts, while selection is more common in spoken discourse.*

Key words: *example markers, exemplification, argumentation, selection, EFL, corpus-based analysis.*

1. INTRODUCTION

In their work on exemplification, Lee and Mlynář (2023:609) note that the use of examples is a pervasive social practice (cf. Lischinsky, 2008:243). This is in line with Siepmann's (2005:111) claim that markers used to introduce examples are one of the most prevalent types of discourse markers. Despite this ubiquity, the scholarly attention given to exemplification as a discourse strategy remains limited. As Siepmann (2005:111) points out, exemplification tends to be viewed as a minor textual operation, often overshadowed by more prominently studied rhetorical strategies, such as inference or justification. This perceived marginality is reflected in the relative scarcity of detailed analyses in both linguistic and pedagogical literature. Given this gap, it is reasonable to argue that there is a prevailing perception of examples as secondary devices in textual organization. However, this perspective underestimates the communicative value of examples, which play a crucial role in illustrating points, clarifying ideas, and supporting arguments in a variety of social contexts (Su et al., 2022:1; Na, 2023:284). The present study builds on this premise, proposing that exemplification deserves more sustained attention as a key strategy in meaning-making and discourse construction.

While it is true that the use of examples presents considerable challenges for both native and non-native speakers, a number of studies have shown that L2 learners encounter particular difficulties in using them appropriately (Hinkel, 2004:297; Siepmann, 2005:112; Su and Zhang, 2020:3; Su et al., 2021:122, 2022:2). Given the pivotal role of examples in discourse construction and the problems they present to language users, there is a pressing need for broader and more comprehensive research into their usage patterns, functions, and effects. Through a greater understanding of the complexities of how examples are used in discourse, we can provide valuable insights into how speakers navigate the challenges of communication and optimize their rhetorical strategies. Moreover, a greater understanding of the role of examples can inform pedagogical approaches aimed at enhancing language learners' proficiency in this important area of discourse. Hence, research in this area has the potential to advance both theoretical understanding and practical applications in the fields of linguistics and language learning (Lee and Mlynář, 2023:614).

Existing work on the use of examples tends to categorize all structures that feature examples under the umbrella term *exemplification*. However, authors such as Eggs and McElholm (2013) argue against this simplistic approach, instead delineating three distinct structures: exemplification, argumentation, and selection. Building on

To cite this article: Rodríguez-Abruñeiras, P. (2025). "A Corpus-Based study of examples and example markers in the speech and writing of Spanish EFL students". *Revista de Lingüística y Lenguas Aplicadas*, 20, 115-127. <https://doi.org/10.4995/rlyla.2025.21633>

Correspondence author: paula.rodriguez.abruñeiras@usc.es



this framework, the main aim of the present study is to fill in the gap in the literature by examining how Spanish EFL students use examples. To this end, data has been drawn from the *Santiago University Learner of English Corpus* (SULEC; Palacios-Martínez, 2002-2022), which contains spoken and written material from the language production of EFL learners who are mainly native speakers of Spanish or Galician (although speakers with other L1 backgrounds are also found in the corpus, such as Catalan, Euskera and Portuguese, among others). The corpus data include information on informants' level of English and sex, which also allows for a degree of sociolinguistic analysis. The study, then, seeks to address the following research questions:

- RQ1. What lexico-grammatical features are used by Spanish learners of English to introduce examples?
- RQ2. What is the incidence of EMs across proficiency levels (pre-intermediate, intermediate and advanced)?
- RQ3. Are there disparities in the use of EMs according to learners' sex?
- RQ4. How does text type (spoken vs. written) affect the use of EMs?
- RQ5. How are the EMs *for example* and *for instance* used in terms of exemplification, argumentation and selection?

The paper is structured as follows. In Section 2 we look at the use of examples as a discourse strategy, (semi) fixed formulas which may be used to introduce examples, and the types of sequences in which examples can be embedded. Section 3 discusses some methodological issues relating to the corpus material and the procedures adopted in the study. After this, the analysis of the corpus material is presented in Section 4: the use of semi-fixed formulas vs. grammaticalized markers to introduce examples, together with a more fine-grained analysis of grammaticalized markers that takes sociolinguistic and textual features into account. A concluding section then outlines the main findings of the study.

2. EXAMPLES AND EXAMPLE MARKERS

2.1. The use of examples as a discourse strategy

Examples are discourse tools that can be used to facilitate understanding and communication by granting epistemic access to specific attributes of the subject matter they exemplify (Triki, 2021; Mann, 2024). They introduce “a fact, incident, quotation, etc. that illustrates, or forms a particular case of, the situation described in the previous discourse”, which is why “in any context imaginable, the use of the [example] will invariably fit the definition” (Siepmann, 2005:99-100). Examples do not explain or describe the previous statement they refer to; instead, they present specific cases in point relating to that previous statement which are in fact “much richer in features than the general rules they are supposed to illustrate” (Lischinsky, 2008:244). Because of their specificity and vividness, examples tend to be more effective than base-rate information (Biggsby et al., 2019; Krämer and Peter, 2020; Ash et al., 2023; Sellnow-Richmond et al., 2024). They evoke both the entirety of a concept and one of its constituent elements (Siepmann, 2005:113). By doing so, examples hold a pivotal role in conveying the intended meanings of speakers and writers, thus ensuring the successful exchange of information and interaction (Triki, 2021). They are, then, “a necessary rhetorical tool for clarity and persuasion” (Triki, 2021:2).

The sequences in which examples are usually inserted (though not exclusively) consist of three units: a first unit with a broader referent (*general element* or GE; *exemplified unit* in Triki's 2021 terminology), a second, more specific unit which provides an example of the GE (*exemplifying element* or EE), and a connector which links the two units and usually comes between them (*example marker* or EM). Whereas this threefold structure is the prototypical one, there are instances in which either the EM or the GE may be omitted, as will be explained in the sections below.

Triki (2021) has pointed out that the EE may either provide an illustration of the GE (it may be a single item or a list of items which belong to the same class as the GE), or it may be an inference of a real or hypothetical example. This idea of examples being either factual or speculative, given or potential, is not novel. Indeed, in *Rhetoric* II, 20 Aristotle classifies examples as a type of oratorical argument (alongside enthymemes) and distinguished between those examples which speakers recount from the past (real, factual) and those which they invent in order to support and illustrate their arguments (Siepmann, 2005:112).

2.2. Grammaticalized and non-grammaticalized formulas to provide examples

Examples can be introduced by means of two main strategies: a grammaticalized EM or a semi-fixed expression (i.e. a collocation with a certain degree of variability; see Siepmann, 2005; Paquot, 2007, 2008; Rodríguez-Abrueñas, 2019a). In Siepmann's (2005) terminology, these would be first and second level discourse

markers, respectively. Although we will focus here largely on grammaticalized (i.e. first level) EMs, the latter will also be considered. Despite authors talking about the use of “obligatory” markers to introduce examples (Meyer, 1992:77), there is a third possibility, in which no explicit expression is used to introduce the EE, although this is far more unusual. Such “zero connectors” (Siepmann, 2005:113) are indicated by a pause in speech or some form of punctuation mark in writing (see also Rodríguez-Abruñeiras, 2019a:72), as in (1) below, where a forward slash has been added to indicate the point in which the GE ends and the EE begins. In example (2), two EEs are given implicitly: on the one hand, the colon introduces a series of items that serve as instances illustrating the GE “a range of assets”; on the other hand, an EE is found between brackets, presenting a list of examples included within “data files”.

1. Daughters without Japanese names talked about how they wanted to have them and how they made sure that their children would. / D20 reported: When we got older and we met our friends that were hapa [half Japanese, half white], they'd have Japanese middle names. We'd go 'Why didn't you give us one of those too?' We were really wanting one. Me and my sister really wanted a Japanese or a middle Japanese name. (Siepmann, 2005:113)
2. These components typically include a range of assets: data files (LP records, COBOL programs, word processing documents, etc.), various pieces of durable hardware, and training, or human capital. (Lischinsky, 2008:254)

In this study, implicit examples such as those in (1) and (2) will not be considered. Rather, we will concentrate on instances where there is some explicit indication of the inclusion of the EE within the GE. Let us turn first to fully established EMs.

Quirk et al. (1985:1316) identify the following as present-day English EMs: *including*, *included*, *for example*, *for instance*, *e.g.*, *say* and *such as*. Meyer (1992:77) adds *like* to this list but classifies *including* and *included* as markers of particularization rather than exemplification. In addition, the *Oxford English Dictionary* attributes an exemplifying function to *as*, which suggests that this form may also be considered part of the set. In previous studies (see Rodríguez-Abruñeiras, 2017, 2019a) we proposed a classification of these forms taking semantic criteria into consideration. This list is reproduced below:

- Neutral EMs: *for example*, *for instance* and *e.g.* These forms introduce the example without adding any specific nuance. *For example* is considered to be the most common and prototypical of all EMs (Siepmann, 2005:89).
- Hypothetical EMs: *say*. The example presented by means of *say* often serves as a supposition or hypothesis: it is presented as a hypothetical illustration of the general explanation, but with no claim that it is in fact a definitive example of that GE.
- Comparative EMs: *like*, *such as* and *as*. These forms originated as comparative forms, a function which they continue to have at present, which is what makes them similar.
- Focalizing EMs: *including* and *included*. These markers imbue the example which they introduce with some emphasis; in other words, examples are presented as intentionally selected to provide a degree of salience. For this reason, Meyer (1992:76-77) classifies them as markers of particularization.

The second option discussed above is to introduce examples by means of semi-fixed expressions with some variability (also known as *phrase-frames* or *p-frames*; see Garner, 2016; Tan and Römer, 2022). Siepmann (2005:118-139) identifies seven such strategies:

- Imperatives, hortatives and functionally equivalent constructions: imperatives containing the forms *take*, *see* or *consider*; *let us + look/take*; *let me + give offer/give*; etc.
- Noun phrases with the head *example*: *for example*, *here is an example*, *as an example*, etc.
- Exemplificatory infinitive clauses: *to take but one example*, *to give a few examples*, *to cite another example*, *to name/mention but two examples*, *to demonstrate this*, *to illustrate with a simple example*, etc.
- Clauses built around verbs denoting ‘illustration’ or ‘demonstration’: *a brief example will serve to illustrate this*, *the following example demonstrates that*, *two examples will suffice*, *some examples may serve to clarify these points*, *an example or two should make clear these all-important distinctions*, *a numerical example helps illuminate the issue*, etc.

- Copular clauses: *a frequent example is, a case in point is, the easiest illustration is*, etc.
- Verb-noun collocations of the type *example + provide: a good example is provided by, at this point an example should be given, another such example is afforded by*, etc.
- Exemplifiers introducing subordinate clauses: *as strikingly exemplified, as vividly illustrated*, etc.

The expressions listed above, although somewhat recurrent, are less frequent than EMs. Moreover, they may exhibit internal variability, such as the potential for adjective insertion to modify the noun *example* or variation in subject-verb agreement. In essence, they represent varying degrees of fixedness and would thus occupy different positions along Michaelis' (2017) "idiomaticity continuum" (see Figure 1 below): grammaticalized EMs would be situated on the far left, while other recurrent yet more variable expressions would be found on the far right.

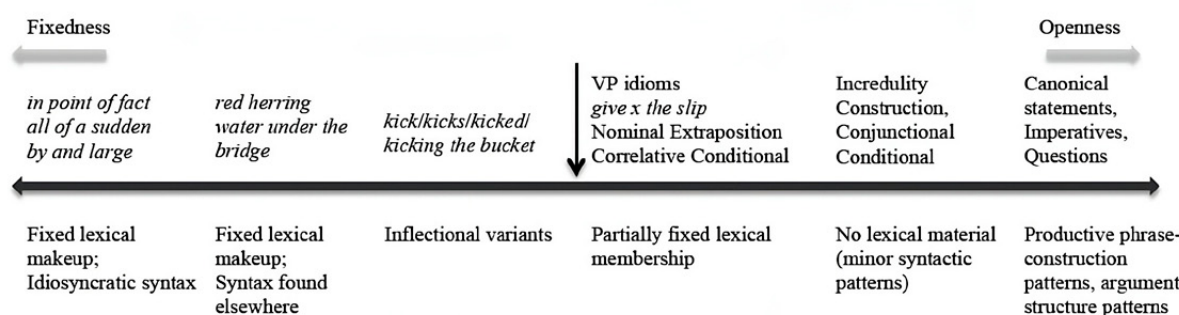


Figure 1. The idiomaticity continuum (Michaelis 2017).

In her study on exemplification in learner writing, Paquot (2008:110, in line with Granger, 1998:156) concludes that non-native learners resort to grammaticalized EMs more often than native speakers, for two main reasons: first, most languages (at least those in her study) have a direct counterpart to these connectors, which makes translation easier for learners. Second, teaching materials and writing guides frequently dedicate sections specifically to adverbial connectors, underscoring their importance in discourse functions. To this we might add that non-native speakers may rely on formulaic expressions due to shortfalls in vocabulary, hence they find fixed formulas easier to identify and integrate, which fosters a sense of comfort in communication (see Rodríguez-Abruñeiras, 2019b).

2.3. Exemplification, selection and argumentation

So far we have been avoiding the term *exemplification* as an inclusive term. For authors such as Quirk et al. (1985), Meyer (1992) and Halliday and Matthiessen (2014), among others, this is a rather restrictive term that can be applied to two-fold sequences which consist of (especially nominal) GE and EE standing in an appositional relationship (a non-prototypical one, in that only partial coreferentiality exists between them). However, examples may be used in sequences which do not have a prototypical appositional context. In this vein, as mentioned in the introductory section, Eggs and McElholm (2013) distinguish three types of sequences which can be used to give examples: exemplification, argumentation, and selection (cf. Rodríguez-Abruñeiras, 2019b, 2021). In exemplification (see example (3) below) and selection (see (4) and (5)), the units involved are complete communicative units which operate on the level of the sentence (they may be phrases, clauses, or sentences). Such structures are categorized by Eggs and McElholm (2013:12) as descriptive.

3. Boyce Avenue has played shows with Secondhand Serenade in important venues, *for example* the Hammersmith Apollo.
4. Boyce Avenue has *for example* played shows with Secondhand Serenade in the Hammersmith Apollo.
5. Boyce Avenue has played shows with *for example* Secondhand Serenade in the Hammersmith Apollo.

However, there are two main differences between exemplification and selection. First, while the former consists of two units (i.e. GE and EE, as in (3)), the latter does not have any explicit GE (as in (4) and (5)). It would be more accurate to say that in selection there is no GE at the linguistic level, though the two units do exist at the cognitive level (Barotto, 2018:26). In these cases, the EM depends on the relationship established by a structural connective or discourse adverbial positioned to its left, on which it is parasitic (a verb in (4); a preposition in (5); or a conjunction; Webber et al., 2003:566; Eggs and McElholm, 2013:8), but it does not refer back to a previous unit that has been mentioned overtly, and hence the construction is not appositional (Koktová, 1985:62). In selection, then, the EM

presents a range of options, and the example provided is just one among many possibilities; in other words, it has a generalizing effect. A second difference between exemplification and selection is that, while selection necessitates the mention of at least one instance or fulfillment case, exemplification retains grammaticality in the event that the EE be omitted (Eggs and McElholm, 2013:14). Thus, in (3) we can just say “Boyce Avenue has played shows with Secondhand Serenade in important venues”, and the omission of “for example the Hammersmith Apollo” does not make the sentence ungrammatical; the omission of the EE in (5), however, would result in the ungrammatical sentence “*Boyce Avenue has played shows with in the Hammersmith Apollo”. The third type of construction, argumentation, is transphrastic in that the GE and the EE are whole sentences and together they form a text (as in (6)-(8); Eggs and McElholm, 2013:12).

6. *For example* Boyce Avenue has played with Secondhand Serenade in the Hammersmith Apollo.
7. Boyce Avenue *for example* has played with Secondhand Serenade in the Hammersmith Apollo. (Eggs and McElholm 2013:11)
8. Many people experience considerable harm or suffering caused by other individuals, without defining themselves as victims. *For example*, Elizabeth, who is very religious, has long been a victim of domestic violence, but she still considers herself responsible. (Eggs and McElholm, 2013:85)

3. METHODOLOGY

For our purposes, we have used the *Santiago University Learner of English Corpus* (SULEC; Palacios-Martínez, 2002-2022), an ongoing project for which compilation began at the University of Santiago de Compostela in 2002 and in its latest update (November 2022) comprises samples from 1374 students (mostly Spanish students born in Galicia), with a total of 365 030 words (296 830 in the written subcorpus and 68 200 in the spoken subset) across various proficiency levels.

Spoken material was gathered via semi-structured interviews, short oral presentations, and brief story descriptions, all recorded and transcribed following standard conventions. The written material was compiled from written compositions and argumentative essays. All informants were either Baccalaureate or Higher Education students (more specifically, students enrolled in an Education, English Philology or Translation degree program). The corpus allows for analyses that incorporate factors such as participants’ proficiency level in English (pre-intermediate, intermediate, advanced), their sex (male, female), and the text type (spoken, written).

The procedure that we followed involved searching for (i) all forms identified as EMs by Meyer (1992) and Quirk et al. (1985), and (ii) the terms provided by Siepmann (2005) which may function as non-grammaticalized strategies for introducing examples, as discussed in Section 2.2 above. These include, but are not limited to, words such as *take*, *see*, *consider*, *let*, *example*, *instance*, *exemplify*, *demonstrate*, and *illustrate*, among others. The results are discussed in detail in the following section. Then, closer attention will be given to the grammaticalized EMs, which will be analyzed from both a sociolinguistic and a textual perspective.

4. THE USE OF EXAMPLES AND EXAMPLE MARKERS IN SULEC

This section explores how Spanish learners of English introduce examples according to the data in SULEC. We consider semi-fixed expressions first (Section 4.1), then fully grammaticalized EMs (Section 4.2).

4.1. Semi-fixed formulas to introduce examples

We have seen how Siepmann (2005) provides a list of non-grammaticalized strategies to introduce examples. In order to explore how Spanish EFL students’ use these, we searched the corpus for the key terms provided by Siepmann (2005). Queries here returned very few results: only four examples were found in which semi-fixed expressions were used to provide examples, and in three of these they were combined with a fully grammaticalized EM. The expressions found were: (i) the hortative expression *let’s* in combination with the verbs *say* and *imagine*, but also with the EMs *like* and *for example* (see examples (9) and (10)); an imperative containing the form *take* in combination with the EM *for example* (see (11)); and a clause built around the verbs *see* and *illustrate* plus the noun *example* (see (12)).

9. Nowadays, when there are not many jobs, and there is a lot of unemployment, is easier to find a job if you study a course (*let's say like* "FP" or "Ciclos") than an university degree. (SULEC, intermediate, written, female)¹
10. On the other hand, there are many students who finish their degree with excellent marks but they cannot face the real life properly afterwards. *For example*, *let's imagine* a person who has just finished a degree on law and actually know a lot about law but he/she cannot deal with a particular case someone proposed him/her due to the lack of practical work he/she has been doing during the five-year-degree. (SULEC, intermediate, written, male)
11. But in restaurants, museums or cinemas, being a non-smoker and having people around smoking can be very disturbing, and unhealthy too, because even with a ventilated room, the smoke and its odor pervails, *take for example* a bar where people are smoking, even when the doors and windows are widely open, once you get out, your clothes will smell smoke. (SULEC, advanced, written, female)
12. [I]f you've got a map of the street with several turning you have to specify if you take if you have to take the first turning on the left or on the right [...] okay so now we we ah going to *see an example to to illustrate* this explanation *let's see* and we've got here a map well if you want to ask for a direction in English you are here and you want to to to ask where the railway station is its here you have to ask excuse me can you tell me the way to the station? (SULEC, advanced, spoken, female)

The scarcity of non-grammaticalized forms to introduce examples is in line with Paquot's (2008:110) claim that learners of English seem to be "unaware" of these semi-fixed expressions and for this reason underuse them. This idea is underlined in our findings, given that most corpus examples combine these semi-fixed expressions with a well-established EM, suggesting that learners of English do not rely on these open formulas to any great extent. This may be due to several factors related to language acquisition and proficiency.

As noted above, grammaticalized EMs are commonly taught on language courses as part of basic vocabulary and grammar, which makes them more accessible and familiar to learners. Hence, it is easier for learners to understand and apply them in their speech and writing than those semi-fixed formulas which may have a more limited applicability and which require learners to know how to adapt them to a specific context (e.g., *take X*, *consider Y*, *let me illustrate*, *to exemplify*, etc.). Other studies have shown that at more advanced levels of proficiency, L2 students use "phrase-frames that are more varied and less predictable, indicating a development of phraseological competence as proficiency advance[s]" (Su and Ye, 2023:3 on Tan and Römer's 2022 study on the use of phrase-frames by L1 Chinese learners of English). This is not the case in our present study, as semi-fixed formulas were not found to be common at either basic or more advanced levels (that said, the four examples retrieved were produced by learners with either an intermediate or an advanced level, with no such cases in the pre-intermediate level).

4.2. Grammaticalized Example Markers

The next step was to search for all the forms which were used as EMs and then manually exclude those which did not perform an exemplifying function. Regarding *for example* and *for instance*, we searched for the nouns *example* and *instance*, respectively, instead of seeking the entire periphrastic marker. This was to identify instances in which students might have mistakenly used a preposition other than *for* with these nouns (but also to check other non-grammaticalized expressions with these two nouns, as discussed in the preceding section). As a result of this approach, examples such as (13) below were retrieved, where the Spanish preposition *por* is used alongside the English noun *example*. Similarly, we also checked the plural variants of both forms, yielding examples like (14), where the marker is pluralized. Despite containing mistakes, these examples were included as they provide insights into the students' level of English.

13. Some persons not feel the same. *Por example* my father and my sister like the programma the operación triunfo and they wach all the days in the TV. (SULEC, pre-intermediate, written, female)
14. But smokers have to respect some places, *for examples* where there children, ill people, old people or people who molest cigarettes, because all people are free to decide if they want or they don't want to smok and if they don't want to smok they don't have to suffer from smokers. (SULEC, pre-intermediate, written, female)

The searches yielded a total of 4074 sequences containing one of these terms regardless of its function. The next step was to discard those instances in which the terms served purposes other than as EMs. The distribution of exemplifying vs. other uses is presented in Figure 2.

¹ Examples are copied directly from the corpus, reproducing spelling and grammar mistakes.

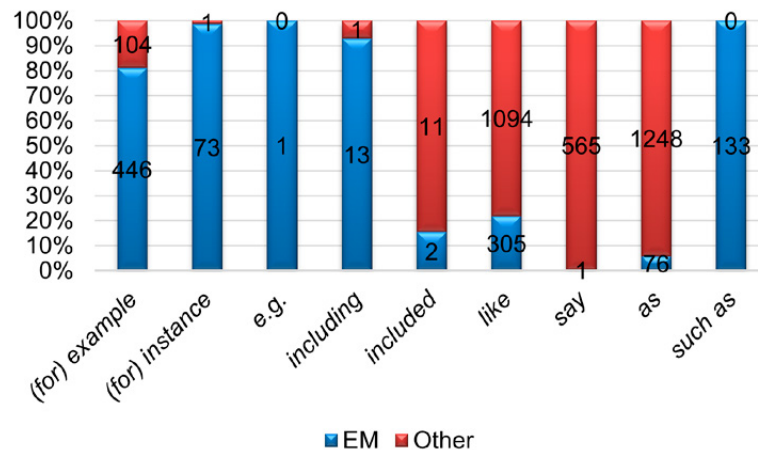


Figure 2. Frequency of EMs in SULEC in comparison to their non-exemplifying uses.

Figure 2 clearly shows that the majority of the terms considered are used either exclusively (e.g. and *such as*) or predominantly (*(for) example*, *(for) instance* and *including*) as EMs. However, certain terms tend to serve non-exemplifying functions more frequently. This applies to *included* (which appears only twice as an EM compared to 11 non-exemplifying occurrences), *like* (305 examples as an EM vs. 1094 in other roles), *say* (just one EM use and 565 with other functions) and *as* (76 examples as an EM and 1248 with other functions). These patterns may be explained by considering the overall frequency, functional versatility and syntactic position of each form.

Like, *say* and *as* are three highly frequent forms in the English language with a wide variety of functions. Indeed, *like* is well known for its versatility across various grammatical categories (according to the OED, it can function as a verb, noun, adjective, adverb, preposition or conjunction; see Peters, 2004:323; Rodríguez-Abrueñas, 2017, 2019a). Regarding *say* and *as*, they rank within the top 100 and 20 most common words in English, respectively, according to the OED. It is therefore unsurprising that their non-exemplifying uses outnumber their exemplifying ones, since they are functionally loaded. As for *included*, although it shares both formal and functional similarities with *including*, these two terms exhibit a division of labor: while *including* precedes the example it introduces, *included* follows it, this final position potentially leading to ambiguity in interpretation. Moreover, *included* appears more frequently as a verbal form (simple past and past participle; see (15) below) compared to *including* (present participle; see (16)). Consequently, *included* has not been widely adopted as an EM (see Rodríguez-Abrueñas, 2019b).

15. As a good nonsmoker, I am *included* in the first group. (SULEC, intermediate, written, female)

16. I am not against homosexuality, but they haven't the right of *including* an innocent creature (creature creo que no é a palabra, casi que no) in nothing. (SULEC, intermediate, written, male)

Once non-exemplifying uses were discarded, the data from SULEC comprised 1050 examples containing a grammaticalized EM. Figure 3 below sets out the frequency of all EMs found in the corpus.

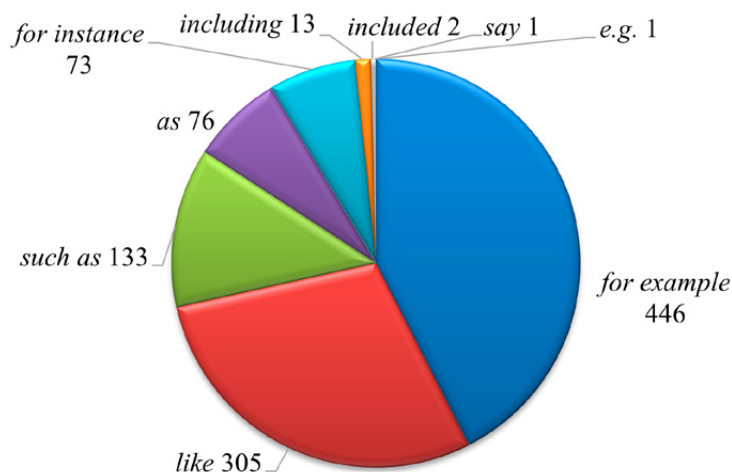


Figure 3. Distribution of EMs in SULEC.

As can be seen, *for example* (446 tokens) and *like* (305) are by far the most common EMs in the data. In turn, *including* (13), *included* (2), *say* (1) and *e.g.* (1) are barely attested. One final consideration regarding the EMs found in the corpus is that in some instances students use the form *how* instead of *like* or *such as* to introduce the EE, as in (17), where *how* comes before the EE and *for example* after it. This incorrect use of EMs appears to be common at the most basic levels of proficiency, and can be explained in terms of the similarities in meaning between *like*, *such as* and *how*, at least in their translation into Spanish (they may all be translated as 'como'). Such examples have not been included in this study given that finding all forms which are mistakenly used as an EM would be unfeasibly time consuming.

17. I am agree exist places where we can't smoke *how* hospitals, pharmacia, or banks *for example* because you should (molestar) persons stay there. (SULEC, pre-intermediate, written, male)

In what follows, we will address the corpus data in Figure 3 from different perspectives, starting with a sociolinguistic analysis (Section 4.2.1) and continuing with the analysis of different textual features (Section 4.2.2).

4.2.1. Sociolinguistic analysis

Let us start by considering learners' level of English. Overall, the use of EMs is rather balanced across the three levels represented in the corpus (i.e. preintermediate, intermediate and advanced), albeit with a slight upward trend as learners' command of English improves. This is reflected in their normalized frequencies (NF): NF 28.00 at the pre-intermediate level, NF 28.96 at the intermediate level and NF 29.22 at the advanced level. However, a more fine-grained analysis shows notable differences when markers are considered individually, as illustrated in Figure 4.²

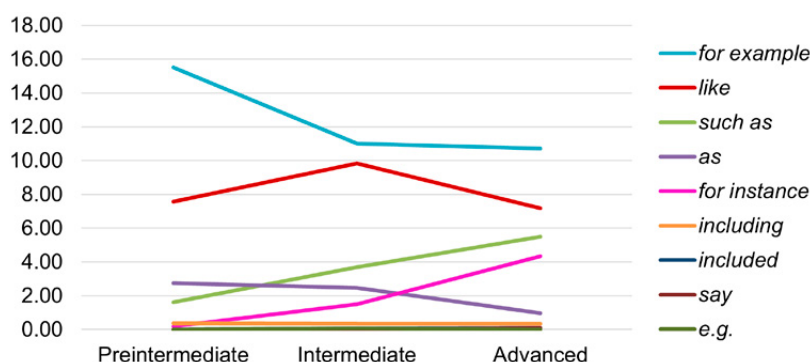


Figure 4. EMs in SULEC according to learners' proficiency level (NFs).

As we can see here, *for example* and *like* are the preferred markers regardless of students' level of English. In fact, at the pre-intermediate level the use of these two forms alone accounts for 82% of all EMs. More advanced levels show the use of a wider variety of markers. Whereas the use of *like* does not exhibit any definite tendency, *for example* clearly recedes in frequency as students' command of English improves, thus showing an opposing trend to that of *such as* and *for instance*, whose use is very limited in the pre-intermediate level but increases substantially at more advanced levels. The use of *as* is consistently infrequent, with tendency to decrease in use as proficiency increases. The remaining EMs are extremely rare at all three levels.

The distribution of EMs according to learners' command of English may be expected if we take into account how learners' dictionaries classify these words. In both the *Oxford Learner's Dictionaries* (henceforth OXD) and the *Cambridge Advanced Learner's Dictionary* (henceforth CALD) *for example* is cited as the first EM to be learned, with both sources classifying it as typical of an A1 level. Moreover, Spanish has an exact equivalent (*por ejemplo*), which will most probably have a direct impact on the use of this EM in our data.

Also common in the most basic levels but slightly more advanced are *such as* (A1 in OLD, A2 in CALD) and *like* (A2 in OLD, B1 in CALD). Finally, *for instance* (B1 in CALD) and *say* (B2 in OLD) are ascribed to more advanced levels. No reference to *e.g.* is made in any of these dictionaries. By and large, this form is not very common in English, especially since it is a rather formal marker whose use is often restricted to parenthetical references (cf. *The Chicago Manual of Style*, 1982:383; see Rodríguez-Abruñeiras, 2017, 2019a). We might, then, expect its use to be marginal in the discourse of young learners of English. Note, however, that it can be very popular in other

² Subsets of the corpus data vary in terms of word count, and thus normalized frequencies (NF) per 10000 words are used in the remainder of this paper to make direct comparisons possible.

formal texts (for instance, e.g. ranks first in Na's 2023 study of exemplification in research articles). Likewise, the exemplifying function of *as* is not mentioned in either OXD or CALD, which is not surprising given that this form is not usually listed as an EM, as we have already pointed out in Section 2.2 above.

The second variable we considered was learners' sex. As illustrated in Figure 5, there is a comparable use of EMs among male and female informants, although women generally exhibit a higher frequency of use than men. Notably, there are two exceptions: *like* and *such as* are both more prevalent in male production. These two EMs typically introduce short and integrated EEs, i.e. EEs which are not separated from the GE by means of a pause (see Rodríguez-Abruñeiras, 2017:94, 2019a:111), as in examples (18) and (19), where the EEs are *Germany* and *cancer*, respectively.

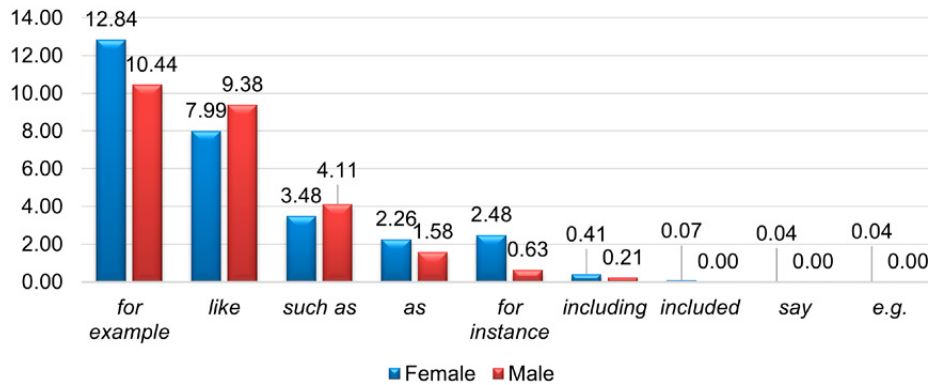


Figure 5. EMs in SULEC according to learners' sex (NFs).

18. In my opinion, we should follow the example of some countries *such as* Germany where, although smoking is not banned, there are separate areas for smokers and non-smokers. (SULEC, intermediate, written, male)
19. Smoking is a harmful habit wich has become natural, although it provokes many severe diseases *like* cancer. (SULEC, intermediate, written, male)

4.2.2. Textual analysis

When it comes to the use of EMs in the written vs. spoken datasets, these were found overall to be more common in the former (904 tokens, NF 30.46) than the latter (146 tokens, NF 21.41). However, the frequency of each marker varies according to the spoken vs. written criterion, as shown in Figure 6.

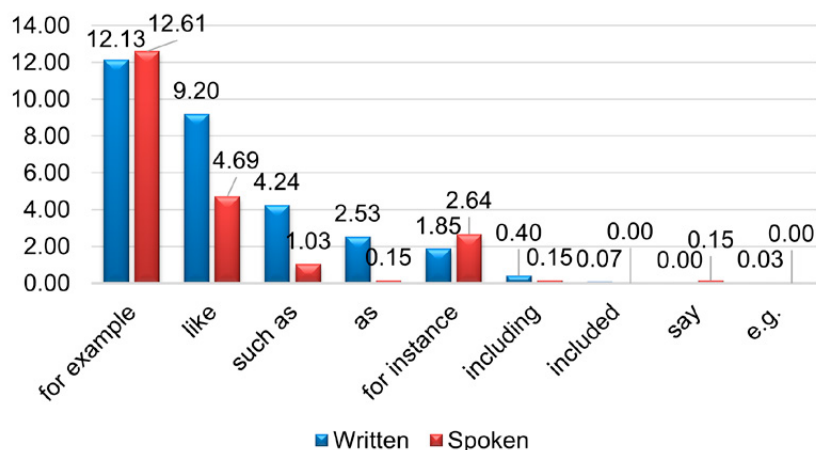


Figure 6. Frequency of EMs in the spoken vs. written data (NFs).

As can be seen in Figure 6 above, most EMs follow the general trend we have noted. EMs are more frequent in the written form than in spoken production, with three exceptions: *for example*, *for instance* and *say*. Given that *say* is attested only once, no satisfactory conclusions regarding its use can be drawn. As for *for example* and *for instance*, these are EMs which can appear in all three types of sequences distinguished by Eggs and

McElholm (2013): exemplification, argumentation, and selection. The versatility of these two forms may have to do with their general prevalence in speech, although before speculating on their behavior we need to explore the data in more detail.

There is one EM in Figure 6 which deserves further attention: *like*. This form ranks second in both spoken and written data, which is somewhat unexpected given its stigmatized status in some late 20th-century sources (Freeman, 1990:252; Burchfield, 1998:459; cf. Bernstein, 1971; Kilpatrick, 1984; see also Rodríguez-Abruñeiras, 2017:93–94; 2019a:99–101). The controversy surrounding *like* stems largely from its potential ambiguity and informality. Functionally, *like* can serve to introduce an example or to express similarity, which can obscure its intended meaning. For instance, the phrase “a girl *like* you” could either be an example (i.e., someone such as you) or a comparison (similar to you). In addition, *like* is often associated with informal or colloquial speech, leading to its discouragement in more formal contexts (Carter et al., 2011). The fact that *like* appears so frequently in the written subset of SULEC (even more so than in the spoken subset) suggests that the students who contributed to the corpus may not have been aware of either its controversial usage or its informal tone.

We return now to the use of *for example* and *for instance* in exemplification, argumentation and selection, which, as we pointed out, merits closer attention. Figure 7 presents the overall distribution of these two forms across the three kinds of functions, taking text type into account.

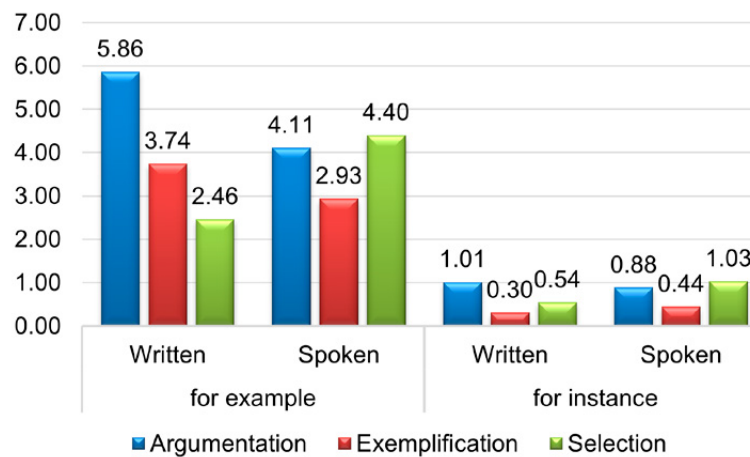


Figure 7. *For example* and *for instance* in SULEC according to text type and type of sequence (NFs).

With both EMs, argumentation emerges as the predominant function in the written data (*for example* NF 5.86, *for instance* NF 1.01), whereas selection prevails in the spoken material (*for example* NF 4.40, *for instance* NF 1.03). Interestingly, exemplification (in the narrower sense of the term as used by Eggs and McElholm, 2013) is the least common of the three sequences under analysis here for both of the EMs.

The use of examples in the three sequences under scrutiny (i.e. argumentation, selection and exemplification) reflects different textual and pragmatic needs. Argumentation probably prevails in written texts because of the intrinsic characteristics written discourse: in written texts, the writer and reader do not share time and space, which is why the information introduced by the EMs is deemed as essential in transmitting a clear and unequivocal message.

In instances such as (20) the example provided is necessary to persuade the reader of the more general thesis given in the previous sentence (i.e. in the GE). In turn, EMs in the spoken material are frequently used with a variety of pragmatic implications. For instance, in (21) *for example* is used after the verb *like* with a generalizing effect to indicate that the examples mentioned are not exhaustive.

20. Students who belong to theoretical degrees usually study the past. *For example*, history is based on the past of a civilization, a country; philosophy is based on the theories of Kant, the Platonism. (SULEC, intermediate, written, female)
21. Yes and I like eh very much *for example* eh eh to walk eh along Santiago and and listening to these people playing the guitar or eh or other instrument. (SULEC, intermediate, spoken, female)

Let us consider one final example. In the long extract provided in (22), *for example* is used four times: the first two occurrences correspond to exemplification, the latter ones to selection. In the two selective uses, *for example* follows the EE and is parasitic on the preposition *with* and the verb *ask*, respectively. In both cases, it also has a generalizing effect: it is added in order to indicate that the examples provided (i.e. *Brad Pitt* and *because [why] they choose these films*) are just two possibilities from many examples which could have been chosen as representative of the implicit GEs.

22. [T]here are lot of celebrity in the world in general *for example* Julia Roberts Antonio Banderas mm my favourite celebrity is Brad Pitt and Julia Roberts and he is my favourite celebrity because he is a in my opinion a handsome man and i like a lot of films er my favourite film is mm in my opinion in was made a er there is er one film er Brad Pitt and Julia Roberts is sorry are together when there is a film er was in the cinema i wasn't because my favourite celebrities have made a film together and Spanish celebrities mm *for example* Antonio Banderas i like or mm i don't remember Antonio Banderas is the best and i would like to be a celebrity because i would like to know the lights and mm i would like can sorry to speak sorry can speak with them and because if i speak with Brad Pitt *for example* i would ask er because they choose these films *for example* and in my opinion er their lives sorry are difficult because there are pression on although it is their lives are also beautiful because they have money and they can travel a lot mm. (SULEC, advanced, spoken, female)

Examples like (21) and (22) above shed light on the pragmatic load that EMs may carry in any chunk of discourse. In this regard, the use of examples has recently been associated with functions such as hedging and attenuation, drawing on the idea that examples imply a set of similar items to represent a general unit and that any of those items could have been given as an illustration of that general concept (i.e. the examples are, in principle, interchangeable). In other words, the examples chosen are arbitrary selections from a pool of possible options (Manzotti, 1998:121, Mihatsch, 2010:108; Barotto, 2018:26; cf. Lo Baido, 2018). In other examples from our corpus, the EM serves as a filler alongside other hesitation marks to allow time for thought, as in (23).

23. And there they are er I don't know er *for example* they have a children one year and three years later they had other children. (SULEC, intermediate, spoken, female)

5. CONCLUSIONS

The analysis of the use of examples and EMs in SULEC reveals interesting aspects of learners' linguistic behavior and proficiency in English. The examination of semi-fixed formulas to introduce examples underlines a preference among Spanish learners of English for grammaticalized EMs over semi-fixed expressions (RQ1). This preference for well-established and more predictable formulas reflects their practicality, versatility, and perceived suitability for various communicative purposes. By mastering such markers, learners can enhance their ability to convey examples accurately and effectively in both spoken and written communication.

Moreover, the analysis of grammaticalized EMs has shown the prevalence of *for example* and *like* in learners' discourse, whereas markers such as *including*, *included*, *e.g.* and *say* are hardly attested at all. However, the frequency and distribution of these markers vary depending on learners' proficiency levels (RQ2), with more advanced learners exhibiting a wider range of EM usage. Thus, whereas students with a lower level of proficiency resort almost exclusively to *for example* and *like*, the frequency of these two EMs declines as students' level of English increases, thus leaving room for other EMs. Differences have also been found regarding sex (RQ3), with female learners generally using all EMs more frequently than males, with the exception of *like* and *such as*, both more frequent in male discourse.

The textual analysis has revealed differences in EM usage between written and spoken language (RQ4), with EMs being more common in written texts. A closer look at the EMs *for example* and *for instance* in these two text types, and taking the type of sequence in which they appear into account, illustrates that argumentation is the predominant sequence in the written dataset, whereas selection prevails in the spoken material (RQ5). It has been argued that this distribution arises from the inherent characteristics of the two text types under analysis. Thus, in written texts authors need to make their words as clear as possible in order to transmit a clear and unequivocal message. In this regard, arguments in the form of examples are a perfect means for a writer to provide support for their argument or position. In turn, EMs may have richer pragmatic implications in the oral data, where they commonly perform a hedging function with a generalizing effect to indicate that the example provided is randomly selected from a set of elements from which other representatives could have been chosen.

Overall, the findings of this study contribute to our understanding of how Spanish EFL students use examples and EMs in their language production. By examining these linguistic phenomena in a learner corpus, we have gained valuable insights into how the use of such devices interacts with learners' proficiency levels, linguistic preferences, and communicative strategies, all of which is of direct application in language teaching and learning practices.

ACKNOWLEDGMENTS

This research was funded by the Consellería de Cultura, Educación, Formación Profesional e Universidades of the Regional Government of Galicia, Spain (grant number ED431B 2023/03).

REFERENCES

- Ash, E., Xu, Y., Pool, R., Schulenberg, K., Mikkilineni, S.D., & Baraka, T. (2023). Exemplification effects on policy support: Exemplar familiarity, narrative vividness, and perceptions of maternal health disparities. *Health Communication, 38*. <https://doi.org/10.1080/10410236.2023.2200907>
- Barotto, A. (2018). The hedging function of exemplification: Evidence from Japanese. *Journal of Pragmatics, 123*, 24-37. <https://doi.org/10.1016/j.pragma.2017.09.007>
- Bernstein, T.M. (1971). *Miss Thistlebottom's Hobgoblins. The careful writer's guide to the taboos, bugbears, and outmoded rules of English usage*. Farrar, Straus, and Giroux.
- Bigsby, E., Bigman, C.A., & Martinez Gonzalez, A. (2019). Exemplification theory: A review and meta-analysis of exemplar messages. *Annals of the International Communication Association, 43*(4), 273-296. <https://doi.org/10.1080/23808985.2019.1681903>
- Burchfield, R.W. (1998). *The new Fowler's modern English usage*. Oxford University Press.
- Cambridge advanced learner's dictionary*. (2008, 3rd ed.). Cambridge University Press.
- Carter, R., McCarthy, M., Mark, G., & O'Keeffe, A. (2011). *English grammar today*. Cambridge: Cambridge University Press.
- Eggs, E., & McElholm, D. (2013). *Exemplifications, selections and argumentations: The use of example markers in English and German*. Peter Lang. <https://doi.org/10.3726/978-3-653-03742-5>
- Freeman, M.S. (1990). *The wordwatcher's guide to good writing and grammar*. Writer's Digest Books.
- Garner, J.R. (2016). A phrase-frame approach to investigating phraseology in learner writing across proficiency levels. *International Journal of Learner Corpus Research, 2*, 31-67. <https://doi.org/10.1075/ijlcr.2.1.02gar>
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A.P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 145-160). Oxford University Press. <https://doi.org/10.1093/oso/9780198294252.003.007>
- Halliday, M.A.K., & Matthiessen, C.M.I.M. (2014). *An introduction to functional grammar* (4th ed.). Edward Arnold. <https://doi.org/10.4324/9780203783771>
- Hinkel, E. (2004). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781410609427>
- Krämer, B., & Peter, C. (2020). Exemplification effects: A meta-analysis. *Human Communication Research, 00*, 1-30. <https://doi.org/10.1093/hcr/hqz024>
- Kilpatrick, J.J. (1984). *The writer's art*. Andrews McMeel Publishing LLC.
- Koktová, E. (1985). Apposition as a pragmatic phenomenon in a functional description. *University of East Anglia Papers in Linguistics, 23*, 39-79.
- Lee, Y., & Mlynář, J. (2023). 'For example' formulations and the interactional work of exemplification. *Human Studies, 46*, 607-633. <https://doi.org/10.1007/s10746-023-09665-7>
- Lischinsky, A. (2008). Examples as persuasive argument in popular management literature. *Discourse & Communication, 2*(3), 243. 269. <https://doi.org/10.1177/1750481308091907>
- Lo Baido, M.C. (2018). Mitigation via exemplification in present-day Italian: A corpus-based study. A. García Ramón & M.A. Soler Bonafont (Eds.), *ELUA: Estudios de atenuación en el discurso*, Anexo IV (pp. 11-32). <https://doi.org/10.14198/ELUA2018.Anexo4.02>
- Mann, S. (2024). Understanding via exemplification in XAI: How explaining image classification benefits from exemplars. *AI & SOCIETY*. <https://doi.org/10.1007/s00146-023-01837-4>
- Manzotti, E. (1998). L'esempio. Natura, definizioni, problemi. *Cuadernos de Filología Italiana, 5*, 99-123.
- Meyer, C.F. (1992). *Apposition in contemporary English*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511597824>

- Michaelis, L.A. (2017). Meanings of constructions. *Oxford online research encyclopedia of linguistics*. <https://doi.org/10.1093/acrefore/9780199384655.013.309>
- Mihatsch, W. (2010). The diachrony of rounders and adaptors: Approximation and unidirectional change. In G. Kaltenböck, W. Mihatsch & S. Schneider (Eds.), *New approaches to hedging* (pp. 93-122). Bingley. https://doi.org/10.1163/9789004253247_007
- Na, Y.H. (2023). Exploration of the act of exemplification in research article introductions. *Modern English Education*, 24, 284-304. <https://doi.org/10.18095/meeso.2023.24.1.284>
- Oxford Learner's Dictionaries*. Available at <https://www.oxfordlearnersdictionaries.com/>
- Oxford English Dictionary* (OED). Available at <http://www.oed.com/>
- Palacios Martínez, I. (dir., SPERTUS Research Group). 2002-2022. *The Santiago University corpus of learner English* (SULEC). Santiago: University of Santiago de Compostela. Available at <https://sulec.cesga.es/>.
- Paquot, M. (2007). Towards a productively-oriented academic word list. In J. Waliński, K. Kredens & S. Goźdź-Roszkowski (Eds.), *Lodz studies in LANGUAGE, Vol. 13: Corpora and ICT in language studies* (pp. 127-140). PALC 2005. Peter Lang.
- Paquot, M. (2008). Exemplification in learner writing: A cross-linguistic perspective. In F. Meunier & S. Granger (Eds.), *Phraseology in Language Learning and Teaching* (pp. 101-119). John Benjamins. <https://doi.org/10.1075/z.138.09paq>
- Peters, P. (2004). *The Cambridge guide to English usage*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511487040>
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. Longman.
- Rodríguez-Abrunheiras, P. (2017). Exemplification then and now: A historical overview of English exemplifying markers. *Miscelanea. A Journal of English and American Studies*, 55, 87-107. https://doi.org/10.26754/ojs_misc/mj.20176816
- Rodríguez-Abrunheiras, P. (2019a) *When apposition meets exemplification: A semantic and syntactic description of two related categories*. Edicions UIB.
- Rodríguez-Abrunheiras, P. (2019b). On exemplifying markers in present-day British and American English: Formal and functional implications. *Brno Studies in English*, 45(2), 155-173. <https://doi.org/10.5817/BSE2019-2-8>
- Rodríguez-Abrunheiras, P. (2021). The history of *for example* and *for instance* as markers of exemplification, selection and argumentation (1600-1999). *Atlantis*, 43(1):133-153. <https://doi.org/10.28914/Atlantis-2021-43.1.08>
- Sellnow-Richmond, D.D., Lukacovic, M.N., & Sellnow-Richmond, S.A. (2024). Exemplification in news narratives: Stigmatizing and securitizing effects. *Frontiers in Communication*, 9. <https://doi.org/10.3389/fcomm.2024.1316677>
- Siepmann, D. (2005). *Discourse markers across languages: A contrastive study of second-level discourse markers in native and non-native text with implications for general and pedagogic lexicography*. Taylor & Francis Routledge.
- Su, H., & Ye, J. (2023). Local grammar approach to investigating advanced Chinese EFL learners' development of communicative competence in academic writing: The case of 'exemplification'. *Corpus-based Studies across Humanities*, 1, 1-25. <https://doi.org/10.1515/csh-2023-0022>.
- Su, H., & Zhang, L. (2020). Local grammars and discourse acts in academic writing: A case study of exemplification in linguistics research articles. *Journal of English for Academic Purposes*, 43, 1-11. <https://doi.org/10.1016/j.jeap.2019.100805>
- Su, H., Zhang, Y., & Chau, M.H. (2022). Exemplification in Chinese English-major MA students' and expert writers' academic writing: A local grammar investigation. *Journal of English for Academic Purposes*, 58, 1-12. <https://doi.org/10.1016/j.jeap.2022.101120>
- Su, H., Zhang, Y., & Lu, X. (2021). Applying local grammars to the diachronic investigation of discourse acts in academic writing: The case of exemplification in linguistics research articles. *English for Specific Purposes*, 63, 120-33. <https://doi.org/10.1016/j.esp.2021.05.002>
- Tan, Y., & Römer, U. (2022). Using phrase-frames to trace the language development of L1 Chinese learners of English. *System*, 108, 102844. <https://doi.org/10.1016/j.system.2022.102844>
- The Chicago manual of style*. (1982, 3rd edition). The University of Chicago Press.
- Triki, N. (2021). Exemplification in research articles: Structural, semantic and metadiscursive properties across disciplines. *Journal of English for Academic Purposes*, 54, 101039. <https://doi.org/10.1016/j.jeap.2021.101039>
- Webber, B., Stone, M., Joshi, A., & Knott, A. (2003). Anaphora and discourse structure. *Computational Linguistics*, 29(4), 545-587. <https://doi.org/10.1162/089120103322753347>