

The Necessary Symbiosis: How ChatGPT Co-authored a New Type of Learner's Grammar to Be Displayed in a Digital Writing Assistant

Qian Li

Guangdong University of Foreign Studies (China) ✉

Sven Tarp

Aarhus University (Denmark) ✉

Antoni Nomdedeu-Rull

Universitat Rovira i Virgili (España) ✉

<https://dx.doi.org/10.5209/clac.97409>

Recibido: 30 de julio de 2024 • Aceptado: 18 de octubre de 2024

EN Abstract: This paper reports on an ongoing research project aimed at developing a new type of Spanish learner's grammar, different from those found in textbooks, grammar books and dictionaries. The new grammar, designed to be displayed in digital writing assistants, will explain problems that occur in written learner texts. The paper first describes the main features and functionalities of this grammar and how it will be presented to Spanish learners. It then discusses the development of a methodology for categorising relevant error types, using a unique combination of existing grammars, dictionaries and ChatGPT, all of it supervised by lexicographers with experience in language teaching. Based on this categorisation, the paper explains how the chatbot is prompted to write explanations of the different error types, which it does very well in fruitful interaction with the human lexicographers. The methodology is described in detail with several examples. Finally, the paper explains how the original Spanish explanations are machine translated into English and Chinese, and provides examples of the final result in each language. Throughout the paper, the complex relationship between generative AI and humans is discussed, and it is concluded that a successful result like the one achieved requires both the ability to handle the chatbot properly and the knowledge of the topic being dealt with.

Keywords: writing assistant, learner's grammar, learner corpora, language didactics, generative artificial intelligence, human-AI symbiosis.

Contents: 1. Introduction. 2. Presentation of the overall project. 3. Selecting error sub-categories. 4. Writing explanations in Spanish. 5. Producing explanations in learners' native languages. 6. Conclusions and perspectives. Acknowledgements. References. A. Corpora. B. Digital writing tools. C. Other literature.

Cómo citar: Li, Q.; Tarp, S.; Nomdedeu-Rull, A. (2024). The Necessary Symbiosis: How ChatGPT Co-authored a New Type of Learner's Grammar to Be Displayed in a Digital Writing Assistant. *Círculo de Lingüística Aplicada a la Comunicación* 100 (2024) 145-160. <https://dx.doi.org/10.5209/clac.97409>

"The formulation of a problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill. To raise new questions, new possibilities, to regard old questions from a new angle, requires creative imagination and marks real advance in science."

Einstein & Infeld (1938)

1. Introduction

This paper continues the reflections of Tarp & Nomdedeu-Rull (2024) and Huete-García & Tarp (2024) on the development of a writing assistant for Spanish learner. It discusses how lexicographers involved in the project can benefit from interacting with OpenAI's ChatGPT, or a similar generative AI chatbot, when producing grammatical explanations to be integrated into the writing tool. The title refers to this interaction as a symbiosis. In biology, a symbiosis is defined as a long-term, close relationship or interaction between two species. The relationship can be *mutualistic*, where both symbionts benefit, *commensalistic*, where one benefits and

the other is unaffected, or *parasitic*, where one benefits and the other is harmed. Furthermore, the interaction can be either *factual*, where the two symbionts can exist independently of each other, or *obligate*, where at least one is completely dependent on the other for survival.

The close interaction with a chatbot is, of course, a *factual* symbiosis, since each lexicographer is free to decide whether or not to use this method. At the same time, however, the symbiosis is considered *necessary* if the aim is to speed up the production process without compromising quality. Moreover, the relationship between the human lexicographer and this new artificial lexicographer is essentially *commensalistic*, since only the former benefits, while the latter remains unaffected, except for the spillover of a small amount of data that OpenAI can use for further training, which also adds a touch of *mutualism* to the relationship.

In the concrete project, ChatGPT is used only as an inspiration. However, it has proved to be such a powerful and productive source of inspiration that it deserves the title of co-author. This generative AI tool is prompted by lexicographers with language teaching experience to suggest the content and structure of short texts on grammar, which they then *evaluate* and, in most cases, *edit by adding, deleting or modifying* what they consider necessary on the basis of their knowledge and experience in teaching Spanish to non-native speakers. In this way, they transform the AI-generated texts into didactic explanations of grammatical problems that learners might typically encounter in the writing process. The reason for choosing this method is that it makes the whole process much faster than it would have been if they had to write them from scratch.

The teaching of a foreign language usually involves three types of basic material designed to facilitate the learning process: textbooks, grammar books and dictionaries. Each has its own role to play and deals with grammar from its own perspective, or at least is expected to do so. Textbooks and grammar books deal with the general rules, the latter in a more systematic way and the former usually by means of short explanatory notes or mini-grammars in the backmatter, both adapted to the learners' increasing proficiency level. However, neither textbooks nor grammar books explain these rules word by word, and this, according to Nomdedeu-Rull (2023: 193), is "the main problem that the learner encounters and a task that the dictionary should solve". Unfortunately, as he also notes, this is rarely the case in learners' dictionaries produced in Spain, a criticism shared by Gutiérrez-Cuadrado (1994), Barcalló-Escrivà (1999, 2001, 2010), Nomdedeu-Rull & Tarp (2018, 2024), Dam-Jensen & Tarp (2019) and Nomdedeu-Rull & Barcroft (2023), among others.

The ideal solution is undoubtedly to explain grammar in the users' respective mother tongues, especially for beginners and intermediate learners. However, grammar books made in Spain also tend to fail in this respect, because "publishers usually conceive this type of text as monolingual" (Nomdedeu-Rull, 2023: 193). In contrast, bilingual grammar books for Spanish learners have a long tradition in other countries like Italy and Denmark (see also Tarp, 2018: 368). Besides, most learners cannot be expected to read grammar books from cover to cover, but rather to use them to consult specific problems. Such books should therefore be suitable for consultation well beyond the traditional, ill-conceived indexes. In fact, we know of only one Spanish grammar book that has been designed from the outset as a reference work according to lexicographic access principles and whose dictionary-like index takes up almost a third of the book, namely the bilingual *Spansk Basisgrammatik* published by Jensen (1990).

With all this in mind, we intend to produce a fourth type of Spanish grammar that differs in several important ways from those found in textbooks, grammar books and learners' dictionaries:

- It is written in the learner's native language, which makes it bilingual if that language is not Spanish.
- It is monofunctional, meaning that it has only one lexicographic function, that of helping learners to write Spanish texts in a digital writing assistant, allowing them to gain an immediate and deeper understanding of the different types of grammatical errors they can make.
- It is easily accessible, by simply clicking on words or passages highlighted as problematic by the AI-powered language model that supports the writing assistant.
- It is neither systematic, like grammar books or mini-grammars in textbooks, nor word-specific, like the explanatory notes in textbooks or the syntactic and other grammatical data users would expect to find in dictionaries.
- It consists of a set of explanations, most of which address a type of grammatical problem common to a larger group of words or sequences, but which at the same time can be individualised to explain a specific problem encountered in a learner's text.

As can be seen, the new type of grammar does not compete with traditional grammars, except to meet learners' needs for explanations of specific grammatical problems when using a digital writing assistant. It is not a systematic presentation as in grammar books or mini-grammars in textbooks and sometimes also in the backmatter of dictionaries, nor is it limited to explaining the grammatical rules associated with individual words, as in dictionaries.

Against this backdrop, *the main aim of this paper is to report on a research project* that is currently being conducted, which focuses on developing the methodology for 1) identifying and prioritising those of the learners' grammatical writing problems that require explanations, 2) writing the Spanish explanations in close interaction with ChatGPT, and 3) producing or reproducing the explanations in the learners' respective native languages, here simplified with the two most widely spoken languages in the world, English and Chinese.

The next section briefly introduces the overall project and the role of explanations in it. Section 3 describes the methodology used to select the most relevant grammatical error categories for explanation. Section 4 then discusses how the chatbot was prompted and how the challenges posed by the interaction between human and chatbot were handled. Section 5 provides examples of how the Spanish explanations

were reproduced in the learners' native languages, while the final section presents the main conclusions and perspectives.

2. Presentation of the overall project

Recent years have seen a growing number of digital tools designed to help writers produce texts, either in their mother tongue or a foreign language. These writing assistants have a wide variety of purposes, functionalities, designs and target audiences; see for example *ArgRewrite*, *DeepL Write*, *Ginger*, *Grammarly*, *HARTA*, *LanguageTool*, *LEAD*, *ProWritingAid*, *QuillBot*, *Tekstretter* and *Wordtune*, just to mention a few. Apart from these and similar tools, ChatGPT itself can also be used to correct and even to some extent comment on written texts; see Faiz et al. (2023), Kim et al. (2023), Li et al. (2023), Schmidt-Fajlik (2023), Song & Song (2023) and Wu (2024), among many others.

The earliest writing aids were based on advanced and often statistical programming, but more and more are now AI-powered. The majority still seem to be monolingual English tools, with other languages lagging behind, although more and more are joining the club. The writing tools released so far, at least those known to the authors of this paper, are overwhelmingly monolingual, even if their target audience is non-native speakers. Despite being the third most spoken language in the world, Spanish is woefully underrepresented in this area. Apart from the one presented here, and especially the newly launched Spanish version of *DeepL Write*, the only other project seems to be *HARTA*, which focuses on the appropriate use of collocations in academic texts; see Guzzi & Alonso-Ramos (2023).

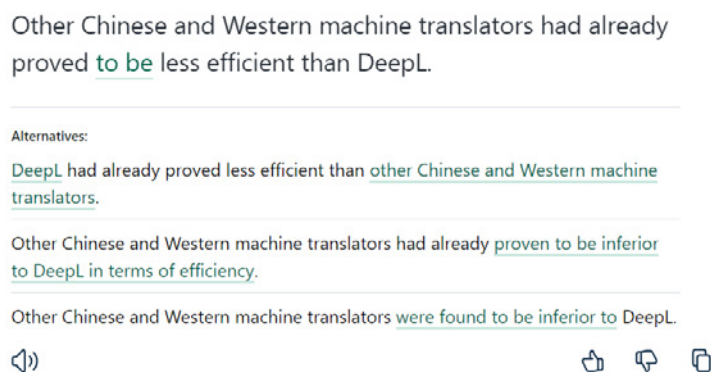
The scope, quality and usability of existing writing assistants vary considerably. Among the absolute top ten, there are a few – like *DeepL Write* and *Grammarly* – of very high quality in terms of helping their users with spelling, word choice, inflection, syntax, style and almost total rephrasing of the original text when necessary. From this perspective, technological development has reached a point of no return, where benefits and risks are mixed in the same pot. On the one hand, there are better opportunities than ever before to write fast, correctly, and linguistically varied in the desired genre and style. On the other hand, there are also great risks lurking on the horizon if the new technology is not used wisely. It is therefore important to distinguish between *writing performance* and *writing skills*:

“If future generations allow themselves to be seduced and become slaves of the new tools, without making a sustained effort to acquire the art of writing, they could end up losing track of what they themselves write.” (Tarp, 2023: 114)

This is perhaps the biggest challenge in this brave new world that would make even Aldous Huxley turn in his grave (cf. Huxley, 1932). If learners adapt to high-tech writing tools without adequately developing their own skills, especially in a second language, they may end up linguistically handicapped, possibly with many negative side effects. The new AI technology is undoubtedly here to stay. Whatever countermeasures are taken here and now, it cannot be suppressed or rolled back in the longer term, and it is likely that ever more advanced writing tools will soon become part of our everyday lives and educational reality. The challenge, then, is to make the best of it. Accordingly, the vision for the Spanish writing assistant under construction is to create a tool that focuses not only on performance but also on didactics. This assistant should not only help learners to produce correct texts in Cervantes' mother tongue, but also support the long-term development of adequate writing skills in this language.

A brief comparison with *DeepL Write* makes this point clear. This tool can correct learners' texts and make really excellent alternative suggestions, also in Spanish, although not yet as good as the English version. But it only highlights possible problems in the texts, without explaining what these problems are and whether they are actually errors or just statistically less frequent solutions. In addition, it often gives alternative suggestions that express something different or even the opposite of what was originally written, an example of which is shown in Figure 1. This implies that a relatively high proficiency level is required to understand the problem highlighted and to judge which of the alternative suggestions can be used in the concrete context. In this respect, *DeepL Write* completely lacks a didactic dimension that can contribute significantly to improved writing skills.

Figure 1. Screenshot with alternative suggestions from *DeepL Write*



The Spanish writing assistant under construction has now been fed with so-called synthetic data from a lexicographic database and then trained on a parallel Spanish corpus with errors and their corrections in order to learn how to distinguish between right and wrong, as explained by Tarp (2023) and Huete-García & Tarp (2024). This has taken longer than expected, as the original plan was modified several times, mainly due to technological improvements, including the programming of a new, more efficient language model and the introduction of ChatGPT, which mainly affected the lexicographers' tasks. The model is now able to identify a large number of different types of errors and problems in written texts and to suggest alternative solutions. The project has entered a phase where these automatically generated suggestions need to be explained in the learners' native language, so that they can better understand them and gain a deeper insight into the problems involved.

Hence, the new writing assistant will not only highlight possible problems and propose alternatives, as DeepL Write does. It will also provide ultra-short default explanations of the problem in the learner's native language, in many cases with the option to obtain additional and more detailed information by clicking the MORE button, as can be seen in Figures 2 and 3.

Figure 2. Short default explanation

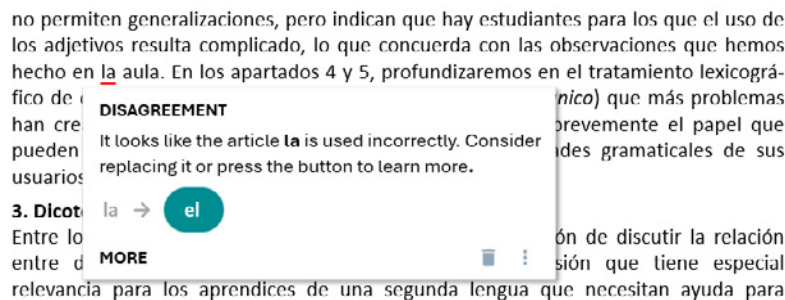
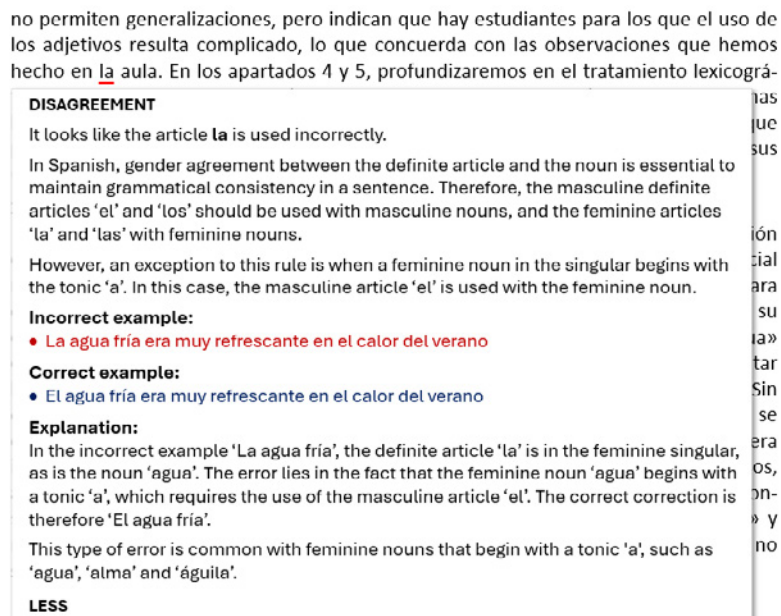


Figure 3. Supplementary explanation



The supplementary explanation shown in Figure 3 was generated by ChatGPT and adapted by the lexicographers. The didactic philosophy behind this design is to open the doors to both incidental and intentional learning, as redefined from a lexicographic perspective by Tarp (2022a). The combination of highlighting, alternative suggestion and ultra-short explanation allows the learner to maintain the flow of writing (or reading, when proofreading), i.e. to remain focused on the text, thus creating the conditions for picking up some grammar and vocabulary incidentally. In contrast, the link to the longer explanation invites the writer to engage in intentional learning, i.e. to shift the focus from the content and form of his or her own text to an explanatory text about a particular grammatical rule, phenomenon or problem. Learners should then be encouraged to take up the invitation, a didactic challenge that requires further research. For now, the plan is to produce the short and long definitions first in Spanish and then reproduce them in the learners' respective mother tongues, as will be discussed in more detail in sections 4 and 5. Before this can be done, however, it is necessary to develop a methodology that will allow the identification of the types of problems that require more detailed explanation. This is done in the next section.

3. Selecting error sub-categories

The basic requirements for grammatical errors that are suitable for explanation in the writing assistant are, on the one hand, that they have a certain minimum frequency in learners' texts and, on the other hand, that they are neither too general nor too specific. If they are too general, the explanations will also be too general and will not explain the concrete problem convincingly to the learners, and if they are too specific, there will be too many of them to be dealt with in a rational way. It is therefore important to develop a set of methods for selecting the relevant error categories. One possibility would be to wait for the writing assistant to be released, see what errors its users make, and then start writing and adding explanations. But this would be more or less like publishing an online dictionary whose articles are only created when people start looking up words. Hence, this method will only be relevant as a supplement once the tool has been introduced to its target audience.

Another, more obvious solution would be to use a tagged corpus composed of texts written by Spanish learners, in which the errors are categorised and their frequency indicated. However, since the creation of a corpus with these characteristics is very time-consuming, the most appropriate option would be to use a corpus that already exists. In this respect, there are several Spanish corpora – such as CAES, CORESPI, CELEN, CATE and CEDEL2 – that contain texts written by non-native learners of Spanish from Italy, Japan, China and other language communities. Some of these corpora were never really completed, probably because the research projects behind them ended when their funding ran out, and most of them do not aim to categorise the different types of errors that learners make. As such, they are not useful for the current project, although they may serve other didactic purposes, as shown by Valverde (2023), among others. From this perspective, these corpora cannot be criticised for their limitations, as they were not designed for the very specific purpose dealt with in this paper.

Among the corpora that do categorise learners' errors, the *Corpus of Written Spanish of L2 and Heritage Speakers* (COWS-L2H) is the only one that provides open data on the most frequent errors, previously categorised by type of error. As reported by Hernández Muñoz et al. (forthcoming), this corpus is divided into 7 main error categories, but there is no information on further sub-categories, at least at the time of writing; see also Yamada et al. (2020). As such, the categorisation is currently much less detailed than in the *International Corpus of Learner English* (ICLE), used by Bestgen & Granger (2011) to sub-categorise spelling errors, which contains 8 main domains and 56 error categories, or in the *Chinese Learner English Corpus* (Gui & Yang 2003), which structures errors into 11 domains and 61 categories and is used as a starting point by Li & Tarp (2024) to further sub-categorise grammatical errors to be explained in a writing assistant project similar to the one described above. This suggests that the current project cannot copy the methodology applied in either of these two projects, but has to develop its own, although this does not mean that the COWS-L2H is entirely useless in this regard.

The COWS-L2H corpus is a collection of texts written by L2 and heritage speakers of Spanish which, according to the team behind it, “includes annotations of grammatical errors based on a previously established taxonomy”, with a total of “9463 errors related to the categories annotated in the nominal syntagm”. The categories with the highest frequency are “the improper presence of a subject personal pronoun (3051 cases, 32%), followed by gender agreement errors (2235 cases, 24%) and the absence of an article (1695 cases, 18%)” (Hernández Muñoz et al., forthcoming). For the purposes of this paper, the category “gender agreement errors”, which accounts for a quarter of all errors registered, was chosen as the basis for further sub-categorisation, using a unique combination of two Spanish reference grammars, a dictionary and a style guide, together with ChatGPT, all supervised by experienced lexicographers and language teachers.

The Royal Spanish Academy states in its *Nueva gramática de la lengua española* (New grammar of the Spanish language) that “gender agreement is not optional in Spanish” (RAE, 2009: 82), i.e. it is mandatory. In his long chapter on agreement in the *Gramática descriptiva de la lengua española* (Descriptive grammar of the Spanish language), edited by Bosque & Demonte (1999), José Antonio Martínez explains that in Spanish agreement is:

“a relationship between at least two words, established by the repetition in each of them of one of the morphemes of gender, number or person, and which basically serves to relate and identify lexically and syntactically the agreed words, among them, article and adjective with noun, pronoun with noun, and verb with noun or pronoun” (Martínez, 1999: 2697).

In other words, there are two main types of agreement: 1) *gender* (masculine or feminine) and *number* (singular or plural) agreement, when it concerns the relation of article and adjective to noun, and pronoun to noun; 2) *person* agreement, when it concerns the relation of verb to noun or pronoun.

The *Curricular Plan* of the Cervantes Institute (2006) was consulted for further inspiration, but this plan, which refers to the *Common European Framework of Reference for Languages* (Council of Europe, 2001), does not categorise error types. In contrast, the *Style Guide* of the public broadcaster RTVE, an authoritative source for language policy, discusses some error types in its section on “Gender and number disagreement”. Most of them are exclusively number-related, not gender-related, and as such, although relevant to the overall project, they are not relevant to the issue addressed in this paper, which is gender agreement errors. But at the end of the section, it is stated that “the gender of the cardinal numbers (*un, una*) must always agree with the nouns they accompany”. It is therefore correct to write “*sesenta y una mujeres*” (sixty-one women), but incorrect to write “*sesenta y un mujeres*”, a useful error sub-category for the present purpose.

For its part, the *Nueva gramática de la lengua española* (RAE 2009) does explain gender, number and person agreement, but it does so by means of a special case-by-case treatment that does not shed light on the different categories of agreement:

- gender agreement
- agreement with *alguien* (someone) and *nadie* (nobody)
- agreement with definiteness
- agreement with compound groups
- agreement with *yo* (I) and *tú* (you)
- agreement with adjective preceded by *lo* (how)
- agreement with the quantifier *medio* (half)
- agreement between reflexive pronoun and its antecedent
- agreement within partitive and pseudo-partitive constructions
- agreement with the pattern *algo/nada* (something/nothing) + *de* + *adjective*
- agreement between article and noun
- agreement between relative and antecedent
- agreement between subject and attribute
- agreement between subject and verb
- partial agreement

In the section on gender agreement, it explains different cases of this phenomenon in Spanish, but does not group them into categories. In one subsection, for example, it notes that “the gender characteristics of the noun extend to the nominal group it forms”. Thus, the adjective *pequeña* in *La mesa del comedor era pequeña* (the dining table was small) agrees with the noun *mesa* and, by extension, with the entire nominal group *La mesa del comedor*. In another subsection, it states that there is no gender disagreement in sequences such as *el alma dormida* (the sleeping soul), because the use of the article form *el* before feminine nouns beginning with the tonic *a* “is due to morphophonological reasons”. And so forth.

The information on agreement offered by the *Diccionario panhispánico de dudas* (RAE & ASALE, 2024) is easier to consult than the grammar from the Royal Spanish Academy. This “Pan-Hispanic Dictionary of Doubts” begins by defining agreement as “the obligatory coincidence of certain grammatical accidents (gender, number and person) between different variable elements of the sentence”, then gives general rules and explains both verbal and nominal agreement. It lists 15 types of nominal agreement, some of which are useful for the current project, but without categorising them. The list includes, among others:

- Single determiner for multiple nouns
- Coordination of nominal syntagms with elided nucleus
- Coordinations by gender splitting
- Adjective postposed to several nouns in copulative coordination
- Adjective postposed to several nouns in disjunctive coordination
- Adjective or participle anteposed to several nouns
- Several singular coordinated adjectives modifying a plural noun
- Whole numbers or decimals anteposed to a noun
- Multiple coordinated ordinals modifying the same noun
- Cardinal postposed to a feminine noun
- Partitive constructions

There seems to be a general consensus that gender agreement in Spanish should be classified within nominal agreement, which is established between a noun (or pronoun) and its adjuncts (determiners and adjectives) in terms of gender (feminine or masculine) and number (singular or plural). With regards to gender, which is the focus here, the adjuncts of a feminine singular noun must agree in the feminine singular: *la* (article determiner) *mesa* (noun) *redonda* (adjective). Similarly, the adjuncts of a masculine plural noun must agree in the masculine plural: *unos* (indefinite determiner) *profesores* (noun) *maravillosos* (adjective). There is also nominal agreement between the pronoun and its antecedent or consequent: *Le dije a Juan que regresara pronto / A Juan le dije que regresara pronto* (both: I told Juan to come back soon), and between the subject and the participle of the passive voice: *El libro fue escrito por Juan* (the book was written by Juan).

However, none of the above references alludes to the frequency of the different types of gender agreement errors, so ChatGPT was asked for help. All communication with the chatbot was in Spanish, but the prompts and its responses are reproduced here in English.

When asked the question *What are the most frequent gender agreement errors in Spanish?*, ChatGPT first divides its answer into three parts: 1) Introduction to the problem, 2) Categorisation and sub-categorisation with examples of typical errors, and 3) Summary. However, it also explains number agreement, which was not asked for in the first prompt. So the prompt was changed to *What are the most frequent gender and gender-only (non-number) agreement errors in Spanish?* The chatbot again divides its answer into three sections, this time with only the required data. As a result, it provides the most common types of gender agreement errors grouped by sub-category, including examples of typical errors:

- Noun and Adjective (*La perros negros*) (the black dogs)

- Noun and Article (*La problema es serio*) (the problem is serious)
- Noun and Pronoun (*Ella es mi amigo*) (she is my friend)
- Noun and Participle (*Las puertas abierto*) (the open doors)
- Adjectives that (do not) change with gender (*Una chica grande / Un hombre trabajadora*) (a big girl / a hard-working man)

Some of the examples provided by the chatbot are incorrect, such as *La perros negros*, where there is agreement between noun and adjective, but not between article and noun, although it should be the other way around. After consulting all the different types of sources mentioned above, we can see that gender agreement is associated with article, noun, pronoun, adjective and participle (and, within the category “adjective”, also with pronouns, demonstratives, some possessives and quantifiers). All in all, from the combined contributions of the chatbot, two grammars, the style guide, the dictionary consulted and, last but not least, the lexicographers who always have the last word, it is possible to establish the different sub-categories of gender agreement errors in Spanish:

1. **Noun and adjective**
 - 1.1. Disagreement between noun and postposed adjective.
 - 1.2. Disagreement between singular noun and postposed non-variable adjective
 - 1.3. Disagreement between singular noun and anteposed adjective
 - 1.4. Disagreement between same-gender coordinated nouns and adjective
 - 1.5. Disagreement between different-gender coordinated nouns and adjective
 - 1.6. Disagreement between singular coordinate adjective and plural noun
2. **Noun and determiner**
 - 2.1. Disagreement between article and noun
 - 2.2. Disagreement between masculine singular article and feminine noun
 - 2.3. Disagreement between coordinated nouns and article
 - 2.4. Disagreement between noun and demonstrative determiner
 - 2.5. Disagreement between noun and definite determiner
 - 2.6. Disagreement between noun and indefinite determiner
 - 2.7. Disagreement between noun and possessive determiner
3. **Noun and pronoun**
 - 3.1. Disagreement between noun and atonic personal pronoun
 - 3.2. Disagreement between antecedent and possessive pronoun
 - 3.3. Disagreement between antecedent and relative pronoun
4. **Noun and participle**
 - 4.1. Disagreement between noun and participle
5. **Pronoun and adjective**
 - 5.1. Disagreement between personal pronoun and adjective
 - 5.2. Disagreement with adjective after pronouns *alguien* and *nadie*
6. **Adverb and adjective/adverb**
 - 6.1. Disagreement with adjective after quantifier adverb *medio*
 - 6.2. Disagreement with adjective/adverb after indefinite adverbs *algo* and *nada*
7. **Number and noun**
 - 7.1. Disagreement with noun after cardinal or decimal
 - 7.2. Disagreement with cardinal ending in *un* or *una* before noun
 - 7.3. Disagreement with partitive constructions such as ‘*uno de los* + noun’
8. **Adjective and subject in *que*-clause**
 - 8.1. Disagreement in constructions with ‘*lo* + adjective + *que*-clause’

The above list may not look like a traditional linguistic one. This is because the starting point is not grammatical rules as such, but only those error types that can and should be explained separately with the specific purpose of being displayed in a digital writing assistant. In addition to the categories mentioned, there are numerous pairs of nouns that have the same form but different meanings depending on their gender, as determined by the article that precedes them. Examples of this phenomenon are: *el capital* and *la capital* (money / city), and *el editorial* and *la editorial* (editorial article / publisher). Since the explanation of these double-gender nouns also involves semantics, it cannot be written as a single standard text, but has to be individualised, at least to a some extent. It will therefore not be dealt with in this paper.

It has not been possible to check the specific frequency of the different sub-categories in the COWS-L2H corpus, as it does not allow queries of this type. In any case, the lexicographers’ combined sixty years of experience in teaching beginner and intermediate Spanish to non-native speakers clearly indicates that all these error types are highly relevant for this group of learners. Thus, although the above list may not include all possible sub-categories, it is undoubtedly a good starting point for the further development of the language model, until it can be tested by its target users and their errors used to fill in any gaps.

4. Writing explanations in Spanish

This section discusses how and to what extent lexicographers can interact with and benefit from generative AI to produce the most appropriate Spanish explanations to be used in the writing assistant under construction.

Such explanations cannot be written without knowing their specific purpose. Thus, before starting the interaction with the chatbot, it is necessary – as also required by Li & Tarp (2024) for a similar project – to identify and clearly define 1) *who* might need the explanations, 2) *for what* they might need them, 3) *in what situation* they might need them, and 4) *in what technological environment* the need might arise.

The answer to these questions is that the explanations are mainly intended to help 1) *beginner and intermediate Spanish learners*, 2) *who may want to better understand the different errors they make*, 3) *when writing Spanish texts*, 4) *using an AI-based writing assistant* that corrects their texts and highlights these errors. This implies that the explanations should be written as:

“short didactic texts in plain language, without too much technical terminology, providing the most relevant information about the specific language problem, and structured in a way that makes it easy for the reader to get an overview and grasp the essence of the problem.” (Li & Tarp, 2024: 400-401)

Writing short didactic texts with these characteristics usually requires in-depth prior reflection not only on the relevant content, but also on the most learner-friendly style and structure. This can be very time-consuming – even for experienced language teachers and lexicographers – and it is therefore worth exploring whether generative AI can help them to shorten the process without compromising on quality, thus avoiding the need to write explanations from scratch.

The following is a discussion of the interaction between ChatGPT and lexicographers. As we will see, it was initially very difficult and almost impossible to prompt it in a way that would produce results with the characteristics described above. These were moments of great tension and frustration, when the lexicographers tended to blame the chatbot for its misguided and useless responses, rather than looking at themselves and their ability to write a more appropriate prompt that would bring out the best in the chatbot. It was the kind of situation that would cause many to give up, but after a few awful days, things started to improve and the chatbot’s responses got better and better, thanks to the lexicographers’ accumulated experience and growing understanding of how to deal with their new interlocutor. That was the first big lesson to be learnt.

To illustrate the initial problems that arose, we will discuss here the lessons learned from experimenting with “Gender disagreement between noun and adjective”, a super-category that will most likely not be explained separately in the final writing assistant presented to users, as it is further divided into six other sub-categories, each of which will receive its own explanation. The chatbot was prompted to explain this type of disagreement in a clear and straightforward way, as if it were explaining it to a layman, and to give an incorrect and correct example of the problem:

“Explain to a non-native learner of Spanish, who is also a layman in grammatical terminology, the error in the gender agreement in Spanish between noun and adjective. I only need one incorrect example and the same example corrected.”

(Prompt 1)

The chatbot provided a relatively long response to the prompt, which is structured into seven sections, where it:

1. introduces the problem and presents the topic of gender agreement in Spanish in a very short and understandable way;
2. explains what gender agreement is, focusing on the agreement requested in the prompt;
3. shows an incorrect example followed by the correct one;
4. gives a simple explanation of the two examples;
5. provides a summary;
6. offers a practical exercise to reinforce what has been learned, and
7. ends with a motivational call: “Keep practising and you’ll soon master agreement in Spanish!”

This response generated by ChatGPT has a number of problems. First of all, the exercise does not meet the requirements to the explanation in this context, as it would take up the learner’s time, nor does the moralising call. In addition, the explanation suffers from inconsistencies at the level of content. For example, it defines a noun as “a word that names things, such as ‘house’ or ‘book’”, and an adjective as “a word that describes things, such as ‘nice’ or ‘big’”, which is not correct because a noun designates entities of a different nature, not just things, whereas an adjective not only describes things but also qualifies or determines the noun, i.e. it *modifies* the noun. Moreover, the response also includes excessive wording that does not help the learner to understand the error better, such as in the sentence “Let’s see how this works with a noun and an adjective in the singular”, or in “Let’s look at a sentence with an error”, the latter being completely redundant because it is already preceded by “Incorrect example” in bold. Finally, it provides too much data.

To avoid lengthy explanations, ChatGPT was therefore instructed to be more concise, providing only a short introduction to the problem posed, an incorrect and correct example, and a brief explanation of the examples:

“Explain briefly and pedagogically to a non-native learner of Spanish the gender agreement error between a noun and an adjective in Spanish. I need you to introduce the problem, give an example with an error and its correction and briefly explain the example. The example sentence must be at least 10 words long.”

(Prompt 2)

In this case, the chatbot significantly reduces the volume of data and simplifies the structure, as shown in Figure 4.

Figure 4. First response to Prompt 2

Introduction: In Spanish, adjectives (words that describe) must match in gender (masculine or feminine) with the nouns they describe. If they do not match, there is a gender agreement error.

Example with error and its correction:

- Incorrect example: La casa bonito tiene una puerta roja.
- Correct example: La casa bonita tiene una puerta roja.

Brief explanation of the example: In the incorrect example, “casa” is a feminine word, but “bonito” is a masculine word. To make them match, we must use “bonita”, which is the feminine form of the adjective. Thus, in the correct example, “casa” and “bonita” are both feminine words, and the sentence has gender agreement.

However, some of the problems identified above are reproduced in this new response, albeit to a lesser extent. There is an inconsistency in the content when indicating that adjectives describe, as adjectives also qualify or determine the noun. And when explaining the noun-adjective disagreement in the incorrect example, the chatbot uses the expression “to make them match”, whereas the correct expression would be “to make them match in gender” or “to make them agree in gender”, as it is always a matter of gender agreement.

On the other hand, when interacting with ChatGPT, the option to get a different response by clicking on its regeneration button was generally used to see if the new text would be more useful in terms of content or structure. In general, differences between the two responses were observed in all cases, giving alternatives to the same problem statement. This is positive because it provides more elements to inspire the lexicographers. At the same time, however, it shows that the chatbot modifies its answers for whatever reason, which is surprising because it implies that there is no guarantee that the same prompt will produce the same results, as can be seen in Figure 5.

Figure 5. Second response to Prompt 2

Introduction: In Spanish, adjectives (words that describe) must match in gender (masculine or feminine) with the nouns (words that name things or people) they describe. This is an important aspect of correct sentences.

Example with error and correction:

- Incorrect example: La casa bonito tiene un jardín muy hermoso.
- Correct example: La casa bonita tiene un jardín muy hermoso.

Brief explanation of the example: In the incorrect example, “casa” is a feminine word, but “bonito” is a masculine word. To make them match, “bonito” must be changed to its feminine form, “bonita”. In the correct example, “casa” and “bonita” are both feminine words, so the sentence is correct.

In the second response to the same prompt, the chatbot claims that nouns are “words that name things or people”. Similarly, the short explanation in this second response says “To make them match, ‘bonito’ must be changed to its feminine form, ‘bonita’”, which is an almost identical repetition of the first response: “To make them match, we must use ‘bonita’, which is the feminine form of the adjective”. In both cases, the expression is still “to make them match”, not “to make them agree in gender”.

Throughout the process described, ChatGPT gave several examples of a striking gap between word and action, between theory and practice, as also observed by Tarp & Nomdedeu-Rull (2024: 317), who concluded that “the chatbot is unable to think and reason like a human”. In most cases, ChatGPT seems to know the grammatical rules well, but occasionally it fails to apply them, even when corrected several times by the lexicographers. This happens, for instance, when, in a third response to Prompt 2, it introduces the noun-adjective agreement rule well, but then claims that the sentence “La casa bonita tiene un jardín muy hermoso” (The big house has a very nice garden) is an incorrect example, which it is not. To make matters worse, the short explanation it gives for this alleged error is completely confusing.

However, it is not difficult for skilled lexicographers to spot and correct the error, proving that even chatbot-generated explanations with some errors can be used to advantage in the production process. In this respect, one of the biggest initial challenges was stylistic, i.e. to obtain a text “without too much technical terminology”, as Li & Tarp (2024) require above, so that the learners do not need to consult an external source

to understand the terms used. However, once the style level had been established and the lexicographers had gained some experience in implementing it, the editing process became much smoother and easier.

Figure 6 shows what a final explanation might look like after passing through the hands of lexicographers and their careful editing. As can be seen, the overall structure of the chatbot's suggestions in Figures 4 and 5 has been retained, but with some design improvements, including the different colours given to the incorrect and correct examples. In terms of content, it has taken into account the critical observations above, and consequently uses the expression "*agree in gender*" instead of just "*match*". It also offers a more correct, but still easy to read, definition of the two word classes NOUN and ADJECTIVE, exemplified with Spanish words that appear in the two examples, so that learners can easily make the connection. The final explanation of these examples is also strongly inspired by those in Figures 4 and 5, with only a few corrections and adaptations.

Figure 6. Adapted explanation based on Figures 4 and 5

In Spanish, adjectives, such as 'bonito' and 'hermoso', must agree in gender (masculine or feminine) with the noun they modify, for example, 'casa' or 'jardín'. This is an important aspect of correct sentences.

Incorrect example:

- La casa bonito tiene un jardín muy hermoso.

Correct example:

- La casa bonita tiene un jardín muy hermoso.

Explanation:

In the incorrect example, 'casa' is a feminine word, whereas 'bonito' is a masculine word. To make them agree in gender, 'bonito' must be changed to its feminine form, 'bonita'.

In the correct example, 'casa' and 'bonita' are both feminine words. They therefore agree in gender and the sentence is correct.

The tests carried out to obtain Prompt 2, to analyse the answers generated by ChatGPT and finally to edit the text reinforce the idea that the human hand is crucial in this process. The chatbot does not always provide relevant data, sometimes gets things wrong and other times misapplies the theory it seems to "know". But overall, it is an important source of inspiration for finding the most didactic way, both in terms of content and structure, to give qualified feedback to learners who make mistakes when writing Spanish texts. The process of interacting with the chatbot served not only to obtain data on the issues raised more easily and quickly, but also to inspire what the explanations should look like when they are eventually linked to the language model supporting the writing assistant under construction, such as the example shown in Figure 3 in Section 2.

5. Producing explanations in learners' native languages

As mentioned in Section 1, the fully developed writing assistant will be bilingual with the explanations written in the learners' native languages. The plan is to produce them first in Spanish, using the methodology developed above, and then automatically translate them with DeepL into a number of languages, including English and Chinese, followed by linguistic revision and based on the experience from another research project; see Tarp (2022b). The English explanations shown in Figures 3-6 were all produced in this way, as were the test explanations for the 25 sub-categories of gender disagreement identified in Section 3.

When the original Spanish explanations were concise and written in a relatively straightforward language, as in Figures 3-6, this method proved to be very fast, efficient and reliable, taking on average 2-4 minutes per translation. When they were more complex, the time taken was longer, mainly because it also involved some re-editing of the style and structure of the Spanish source text as part of the initial learning process and teething troubles. In this way, the English translation was used to further improve the Spanish explanations and gain valuable experience, which of course added time to the process. As such, the problem was not with DeepL, which generally did a good job with only minor problems that could easily be corrected, but with the text material used to feed it in the first place. Apart from that, and although the empirical base is far from sufficient to make a final judgement, it does suggest that the arrow is pointing in the right direction. It can be predicted that other European languages like Italian, German and Danish will show a similar picture. However, this may not be the case for translation into Chinese.

Chinese is the language with the most native speakers in the world. The number of Spanish learners in China is considerable, as is the number of Chinese students studying Spanish abroad. There is even an on-line journal, *SinoELE*, dedicated to the teaching of Spanish as a foreign language to Chinese native speakers. Thus, it makes sense to prepare the writing assistant for this target group as well. However, this is not without its challenges.

The roots and structure of the Chinese language are very different from Indo-European languages like Spanish and English. This suggests that machine translation from Spanish into Chinese may not work as smoothly as it does from Spanish into English, and that Chinese learners of Spanish may need additional information to understand Spanish grammar and the mistakes they make when using the writing assistant. To get a first idea of how to approach this, it was therefore decided to conduct a small test. The Spanish

explanations of the 25 sub-categories of gender disagreement identified in Section 3 were translated into Chinese using the same method as for English, and the result was then analysed to 1) check the quality of the translations and the time taken to produce them, and 2) assess whether other techniques are also needed to make the explanations as useful as possible for Chinese learners.

The result was better than expected, although not of the same high standard as the Spanish-English translation. Most of the translated explanations presented only minor problems, which were easily spotted and corrected by an experienced lexicographer and language teacher. These explanations took an average of five minutes to translate and revise.

Among the “minor” errors, which were easy to correct but would have confused future users, there were a few that changed the meaning of the explanation. For example, when translating the explanation of sub-category 3.3 into Chinese, DeepL was unable to provide a consistent translation of the Spanish term *pronombre relativo* (relative pronoun), which appears five times in this explanation, and which was translated three times with the incorrect Chinese term 相对代词, but only twice with the correct one, 关系代词.

Another challenge was the insufficiency displayed by DeepL in its translation of the Spanish verb *preceder* (to precede), the core meaning of which it seems to understand without grasping its combinatorial properties. For example, in the explanation of sub-category 2.4, which deals with gender and number disagreement between a noun and a demonstrative pronoun, DeepL translated the sequence *el sustantivo al que precede* (the noun it precedes) with 代词前面的名词, which means the exact opposite, namely “the noun before the pronoun”. If left uncorrected, mistakes like this would obviously have consequences for Chinese learners of Spanish, who would have difficulty understanding the gist of the explanation and would probably learn little from it.

An even more serious problem occurred in about a fifth of the test translations, where DeepL missed one or two whole sentences in a single explanation, making it completely incomprehensible. This sin of omission was easy to spot, but it took much more time to find an acceptable solution. Nevertheless, with a good dose of dedication, linguistic flair and creativity, it proved perfectly possible to correct the translation errors, even the serious ones, without having to resort to other techniques, such as asking ChatGPT – or its Chinese counterpart, Ernie Bot – to generate new Chinese explanations from scratch for the respective subcategories of gender and number disagreement.

All this shows, just as in the case of ChatGPT, that it is risky to use DeepL Translate without knowledge of the topic treated, combined with a high proficiency level in the languages involved. But if these requirements are met, DeepL can certainly be used to great advantage for professional purposes, especially when the intention is to subsequently publish the translated explanations and make them available to the specific target audience.

Figure 7 shows the original Spanish explanation of agreement errors between an adjective – placed between the definite article *lo* and a *que*-clause – and the subject of that clause (sub-category 8.1). Figures 8 and 9 then show its respective English and Spanish translations.

Figure 7. Spanish explanation of sub-category 8.1

Un adjetivo como ‘listos’, pospuesto a ‘lo’ y seguido por una oración subordinada introducida por ‘que’, debe concordar en género y número con el sujeto de esta oración, p. ej. ‘niños’.

El artículo definido ‘lo’ es neutro y no cambia de género ni de número en construcciones como ‘lo + adjetivo + que...’, aunque el adjetivo que lo siga sea femenino o plural.

Ejemplos incorrectos:

- No sabes lo listo que son estos niños.
- El examen que esta alumna hizo el otro día demuestra la lista que es.

Ejemplos correctos:

- No sabes lo listos que son estos niños.
- El examen que esta alumna hizo el otro día demuestra lo lista que es.

Explicación:

En el primer ejemplo incorrecto, el adjetivo ‘listo’ es masculino singular, aunque el sustantivo ‘niños’ es masculino plural. Para que haya concordancia de género y número en este caso, el adjetivo también debe ser masculino plural, es decir, ‘listos’.

En el segundo ejemplo incorrecto, el artículo definido ‘la’ tiene forma femenina, pero esto es un error, aunque el adjetivo ‘lista’ sea femenino. Para que la oración sea correcta, debe usarse el artículo neutro ‘lo’, que no cambia de forma en construcciones como ‘lo lista que es’.

Figure 8. English explanation of sub-category 8.1

An adjective such as 'listos' – placed after 'lo' and followed by a *que*-clause – must agree in gender and number with the subject of that clause, e.g. 'niños'.

The definite article 'lo' is neuter and does not change gender or number in constructions like 'lo + adjective + que...', even if the adjective following it is feminine or plural.

Incorrect examples:

- No sabes lo listo que son estos niños.
- El examen que esta alumna hizo el otro día demuestra la lista que es.

Correct examples:

- No sabes lo listos que son estos niños.
- El examen que esta alumna hizo el otro día demuestra lo lista que es.

Explanation:

In the first incorrect example, the adjective 'listo' is masculine singular, although the noun 'niños' is masculine plural. To match the gender and number of the subject in this case, the adjective must also be masculine plural, i.e. 'listos'.

In the second incorrect example, the definite article 'la' is feminine, but this is an error even though the adjective 'lista' is feminine. For the sentence to be correct, the neuter article 'lo' must be used, as 'lo' does not change the form in constructions such as 'lo lista que es'.

Figure 9. Chinese explanation of sub-category 8.1

在'lo'后面的形容词，如 'listos'，后面是由 'que' 引导的从句，其性别和数必须与该从句的主语一致，如 'niños'。

定冠词 'lo' 是中性的，在 'lo + 形容词 + que...' 等结构中即使后面的形容词是阴性或复数，也不会改变性别或数。

错误示例:

- No sabes lo listo que son estos niños.
- El examen que esta alumna hizo el otro día demuestra la lista que es.

正确示例:

- No sabes lo listos que son estos niños.
- El examen que esta alumna hizo el otro día demuestra lo lista que es.

解释:

在第一个错误例句中，形容词 'listo' 是阳性单数，而名词 'niños' 是阳性复数。在这种情况下，为了使性别和数量一致，形容词也必须是阳性复数，即 'listos'。

在第二个错误例句中，定冠词 'la' 的形式是阴性的，但这是错误，尽管形容词 'lista' 是阴性的。要使句子正确，必须使用中性冠词 'lo'，'lo' 在 'lo lista que es' 等结构中不改变形式。

As can be seen, the overall structure originally proposed by ChatGPT in Figure 4 has been retained in all three languages: introduction, example sentences and explanation (in other cases a note has been added to provide additional information, e.g. a list of possessive pronouns relevant to the explanation). The three figures are a good example of the difference in approach between traditional grammars and the one we intend to develop in this project, as explained in Section 1. The former usually focuses on the grammatical rule, which in this case would be the use of the neuter definite article "lo" to emphasise the meaning of an adjective followed by a *que*-clause, as in the sentence: *No sabes lo listos que son estos niños* (You don't know how smart these kids are). By contrast, in the three explanations the focus is on agreement, as they are only displayed when a learner has either used an incorrect form of an adjective or has misinterpreted the agreement as being between the adjective and the article 'lo'. This means that the explanations are not motivated by grammatical rules as such, but by errors that learners often make in written texts.

Another thing that can be seen in the three figures is that it has been deemed necessary to use certain grammatical terms, such as *adjective*, *subject* and *definite article*, but that these terms have been explained as simply as possible, mostly by referring to words that are also used in the example sentences. Learners may not always fully understand what these grammatical categories represent, but they will be able to relate both the introduction and the main explanation to specific words that appear in a concrete context, i.e. to the error types being explained. This is the main idea behind this way of introducing the unavoidable terminology. In addition, as can be seen in Figure 3, the explanation that eventually appears in the writing assistant will be directly linked to the error that the learner has made, through a sentence such as "It looks like the article *la* is used incorrectly".

Furthermore, as the three figures also show, in some explanations it was considered more practical to offer a set of two correct and two incorrect example sentences, in this case representing two different sub-problems, i.e. the gender and number disagreement between a noun and an adjective preceded by the non-variable definite article *lo* on the one hand, and the incorrect spelling of the latter on the other. They could have been explained separately but are explained here together, both to avoid writing and translating too many explanations and to give the learners a bigger picture of the phenomenon. Whether this is a good idea

or not will become clear when it is implemented and the writing assistant is ready to be tested by the target users, who, as always, will have the final say.

Finally, in none of the 25 translated test explanations it was considered necessary to provide additional information for the benefit of the Chinese learners, which of course does not exclude that it might be required for other problem types. But it does prove, contrary to our predictions, that the methodology developed above is largely applicable also to a language as different from Spanish as the one spoken by more than a billion people in China.

6. Conclusions and perspectives

The aim of the research project discussed in this paper is to explore how ChatGPT can be used to write a new type of Spanish grammar in the form of a set of explanations that can help non-native learners understand the mistakes they make when using a digital writing assistant. The above discussion shows that this is possible to the extent that the chatbot can rightly be considered a co-author of the grammar.

In Section 3, a methodology for identifying grammatical error categories in written Spanish was developed, in which ChatGPT played an important role. In section 4, it became the protagonist, contributing decisively to the writing of explanations of these error categories for non-native learners and, consequently, to the development of a methodology for carrying out this particular activity. In Section 5, it gave way to DeepL Translate, which was used to translate the explanations into the target users' native languages, in this case English and Chinese, which it did brilliantly in the former and acceptably with a few major but correctable problems in the latter. Even so, this high-tech tool can by no means be considered a co-author of the new type of grammar, but at best a helpful co-translator, which was also kept under strict human supervision throughout the process.

The new grammar with the features described in Section 1 now has to be completed. This involves identifying new sub-categories of relevant grammatical errors, writing the corresponding explanations and translating them into the target users' native languages, using the methodology developed above. New challenges are likely to arise in this process – such as how to easily explain modal errors, i.e. the use of either the indicative or the subjunctive form of the verb – which the human experts and the chatbot will have to solve in symbiosis, building on the experience gained. For now, the next three main areas to be addressed are “number agreement errors”, “person agreement errors” and “improper presence of subject personal pronouns”, so that grammatical categories representing well over half of the errors and problems registered in the COWS-L2H corpus will be covered.

Moreover, the discussion in this paper also shows that lexicographers, or whoever else uses generative AI for a similar purpose, need a good knowledge of Spanish grammar on the one hand, and a lot of patience, motivation and dedication on the other, as it takes time to learn how to interact appropriately with the chatbot and get it to generate explanations with the required characteristics. Especially in the beginning, the chatbot's responses can be quite challenging and confusing, and it is easy to get frustrated. It is important to understand that this is entirely the fault of the lexicographers, who do not yet know how to instruct it in the right way. But if they persist and stay focused, they will soon get faster and better results.

Writing good prompts on a specific topic is something that has to be learned. It is not enough to have done it on another topic, as was the case with the lexicographers in this project, because the nature of the topic clearly seems to influence how the prompts should be designed. This is one of the reasons why learned knowledge of the subject matter, together with the ability to master the chatbot, is essential to achieve good results.

As stated in Section 2, the new AI-technology is undoubtedly here to stay, but this does not mean that everyone is equally positive about the impact of AI on learners' *writing skills*, whether in the form of chatbots or writing assistants. For example, Abdullayeva & Muzaffarova (2023), who have studied both ChatGPT and writing tools like Grammarly, and who themselves are generally positive about the benefits that the use of such tools can have on students' *writing performance*, also summarise some of their colleagues' concerns about the long-term consequences:

“However much we praise this AI based writing tool, not everyone is positive about the influence of it on writing overall. There are also concerns about the impact of AI on student writing skills. Some worry that the use of AI will discourage students from learning how to write well. If students are relying on AI to correct their mistakes, they may not learn how to identify and correct those mistakes on their own. Additionally, there is a risk that students will become overly reliant on AI and fail to develop critical thinking skills and creativity.” (Abdullayeva & Muzaffarova 2023: 66)

Such long-term concerns should definitely be taken seriously, as the hypothetical risk of losing a generation cannot be entirely ruled out if well-considered countermeasures are not taken, as discussed, for example, in a thought-provoking paper by Liu et al. (2023). In this respect, it is worth recalling the wise words of the down-to-earth Sancho Panza in Cervantes' *Don Quixote*:

“You will see it when you fry the eggs” (*al freír de los huevos lo verá*)

In our case, the eggs to be fried are the explanations, and those who will be frying and later tasting them are the future users when they start testing the writing assistant and its output. Ultimately, however, it is not just about taste, i.e. the immediate writing performance, but above all about the nutritional value of the eggs in terms of developing students' writing skills, critical thinking and creativity. At the moment, therefore, it is most important that the explanations are correct, well structured, easy to read and readily accessible, a task that generative AI can certainly help with. As Francis Bacon (1561-1626) famously said:

“By far the best proof is experience”.

Acknowledgments

Thanks are due to the Center for Lexicographical Studies, the Center for Linguistics and Applied Linguistics and the Research Project on Postgraduate Education and Reform (24GWYJSCX-19) at Guangdong University of Foreign Studies (China) for supporting this project.

Thanks are also due to the Faculty of Arts at Universitat Rovira i Virgili (Spain), Aarhus University Research Foundation (Denmark) and Ordbogen A/S (Denmark) for their logistical, financial and technical support, respectively.

References

A. Corpora

- CAES. Corpus de aprendices de español como lengua extranjera: <https://galvan.usc.es/caes/search>
 CATE. The Corpus of Taiwanese Learners of Spanish: <https://corpora.fild.ncku.edu.tw/>
 CEDEL2. Corpus Escrito del Español como L2: <http://cedel2.learnercorpora.com/>
 CELEN. Corpus de ELE en Japón: <https://sites.google.com/view/celen>
 CORESPI. CORpus del ESPAñol de los Italianos: <https://corespiyorite.altervista.org/>
 COWS-L2H. Corpus of Written Spanish of L2 and Heritage Speakers: <https://github.com/ucdaviscl/cowsl2h>
 ICLE: International Corpus of Learner English: <https://www.i6doc.com/en/book/?GCOI=28001105280390>

B. Digital writing tools

- ArgRewrite: <https://argrewrite.cs.pitt.edu>
 ChatGPT 4o: <https://chat.openai.com>
 DeepL Translate: <https://www.deepl.com/translate>
 DeepL Write: <https://www.deepl.com/write>
 Ginger: <https://www.gingersoftware.com>
 Grammarly: <https://www.grammarly.com>
 HARTA: <http://www.dicesp.com:8083/>
 LanguageTool: <https://languagetool.org>
 LEAD: <https://leaddico.uclouvain.be>
 ProWritingAid: <https://prowritingaid.com>
 QuillBot: <https://quillbot.com>
 Tekstretter: <https://www.grammatip.com/tekstretter>
 Wordtune: <https://www.wordtune.com>

C. Other literature

- Abdullayeva, Madina & Muzaffarovna, Musayeva Zilola. 2023. The Impact of Chat GPT on Student's Writing Skills: an Exploration of AI-assisted Writing Tools. *International Conference of Education, Research and Innovation*, 1(4): 61-66. <https://zenodo.org/records/7876800>
- Bargalló-Escrivà, María. 1999. La información gramatical en los diccionarios didácticos monolingües del español. Vila-Rubio, María Nieves; Calero-Fernández, María Ángeles; Mateu, Rosa M.; Casanovas-Catalá, Monserrat & Orduña-López, José Luis. (Eds.), *Así son los diccionarios*, 15-42. Lleida: Edicions de la Universitat de Lleida.
- Bargalló-Escrivà, María. 2001. La información gramatical en los diccionarios especializados. Bargalló Escrivà, María; Forgas-Berdet, Esther; Garriga-Escribano, Cecilio; Rubio, Ana & Schnitzer, Johannes. (Eds.). *Las lenguas de especialidad y su didáctica*, 81-90. Tarragona: Universitat Rovira i Virgili.
- Bargalló-Escrivà, María. 2010. La información gramatical explícita en los diccionarios monolingües del español. Bernal, Elisenda; Torner, Sergi & DeCesaris, Janet (Eds.), *Estudis de Lexicografia 2003-2005*, 17-37. Barcelona: IULA-UPF.
- Bestgen, Yves & Granger, Sylviane. 2011. Categorizing Spelling Errors to Assess L2 Writing. *International Journal of Continuing Engineering Education and Life-Long Learning*, 21(2/3): 235-252. <http://dx.doi.org/10.1504/IJCEELL.2011.040201>
- Bosque, Ignacio & Violeta Demonte. (Eds.), 1999. *Gramática descriptiva de la lengua española*. Madrid: Espasa. <https://www.espanolavanzado.com/recursos/2771-gramatica-descriptiva-de-la-lengua-espanola>
- Cervantes, Miguel de Saavedra. 1605. *El ingenioso hidalgo don Quijote de la Mancha. Cáp. XXXVII*. Madrid: por luan de la Cuesta. <https://www.cervantesvirtual.com/obra-visor/el-ingenioso-hidalgo-don-quixote-de-la-mancha-9/html/>
- Cervantes Institute. 2006. *Plan curricular del Instituto Cervantes. Niveles de referencia para el español*. Madrid: Instituto Cervantes – Biblioteca nueva. http://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/
- Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Strasbourg: Council of Europe.
- Dam Jensen, Helle & Tarp, Sven. 2019. El tratamiento de adjetivos en gramáticas y diccionarios de español. *Revista Internacional de Lenguas Extranjeras*, 10: 13-47. <https://doi.org/10.17345/rile10.13-47>
- Einstein, Albert & Infeld, Leopold. 1938. *The Evolution of Physics*. New York: Simon & Schuster.

- Faiz, Rabia; Bilal, Hafiz Ahmad; Asghar, Ijaz & Safdar, Aisha. 2023. Optimizing ChatGPT as a Writing Aid for EFL Learners: Balancing Assistance and Skill Development in Writing Proficiency. *Linguistic Forum – A Journal of Linguistics*, 5(3): 24-37. <https://doi.org/10.53057/linfo/2023.5.3.3>
- Gui, Shi Chun & Yang, Hui Zhong. 2003. *Chinese Learner English Corpus*. Shanghai: Foreign Language Education Press.
- Gutiérrez-Cuadrado, Juan. 1994. Gramática y diccionario. *Actas del Congreso de la lengua española*, 637-656. Madrid: Instituto Cervantes.
- Guzzi, Eleonora & Alonso-Ramos, Margarita. 2023. Descripción y usabilidad de HARTA, una herramienta de ayuda para la redacción de textos académicos en español. *TEISEL. Tecnologías para la investigación en segundas lenguas*, 2: 1-22. <https://doi.org/10.1344/teisel.v2.42173>
- Hernández-Muñoz, Natividad; Tomé-Cornejo, Carmela; Belío-Apaolaza, Helena Sofía; Recio-Diego, Álvaro & Davidson, Sam. Anotación de errores gramaticales en el corpus de aprendices de español COWS-L2H. (forthcoming)
- Huete-García, Ángel & Tarp, Sven. 2024. Training an AI-based Writing Assistant for Spanish Learners: the Usefulness of Chatbots and the Indispensability of Human-assisted Intelligence. *Lexikos*, 34: 21-40. <https://doi.org/10.5788/34-1-1862>
- Huxley, Aldous. 1932. *Brave New World*. London: Chatto & Windus.
- Jensen, Kjær. 1990. *Spansk Basisgrammatik*. Aarhus: Akademisk Forlag.
- Kim, Sunyoung; Shim, Joobo & Shim, Jaechang. 2023. A Study on the Utilization of OpenAI ChatGPT as a Second Language Learning Tool. *Journal of Multimedia Information System*, 10(1): 79-88. <https://doi.org/10.33851/JMIS.2023.10.1.79>
- Li, Qian & Tarp, Sven. 2024. Using Generative AI to Provide High-Quality Lexicographic Assistance to Chinese Learners of English. *Lexikos*, 34: 397-418. <https://doi.org/10.5788/34-1-1944>
- Li, Jing; Ren, Xiaohui; Jiang, Xinliang & Chen, Chiu-Hung. 2023. Exploring the Use of ChatGPT in Chinese Language Classrooms. *International Journal of Chinese Language Teaching*, 4(3): 36-55. <https://doi.org/10.46451/ijclt.20230303>
- Liu, Ming; Ren, Yiling; Nyagoga, Lucy Michael; Stonier, Francis; Wu, Zhongming & Yu, Liang. 2023. Future of Education in the Era of Generative Artificial Intelligence: Consensus among Chinese Scholars on Applications of ChatGPT in Schools. *Future in Educational Research*, 1(1): 72-101. <https://doi.org/10.1002/fer3.10>
- Martínez, José Antonio. 1999. La concordancia. Bosque, Ignacio & Demonte, Violeta. (Eds.), *Gramática descriptiva de la lengua española*, 2695-2786. Madrid: Espasa.
- Nomdedeu-Rull, Rull. 2023. Sin noticias aún de datos gramaticales útiles en diccionarios para aprendices. Garriga-Escribano, Cecilio; Iglesia-Martín, Sandra; Moreno-Villanueva, José Antonio & Nomdedeu-Rull, Antoni. (Eds.), *Lligams: Texts dedicats a Maria Bargalló Escrivà*, 191-201. Tarragona: Publicacions URV. <https://llibres.urv.cat/index.php/purv/catalog/download/544/565/1245-1?inline=1>
- Nomdedeu-Rull, Antoni & Barcroft, Joe. 2023. Los diccionarios de ELE. Torner, Sergi; Battaner, Pat & Renau, Irene. (Eds.), *Lexicografía hispánica / The Routledge Handbook of Spanish Lexicography*, 448-466. London and New York: Routledge.
- Nomdedeu-Rull, Antoni & Tarp, Sven. 2018. Hacia un modelo de diccionario en línea para aprendices de español como LE/L2. *Journal of Spanish Language Teaching*, 5(1): 50-65. <https://doi.org/10.1080/23247797.2018.1459279>
- Nomdedeu-Rull, Antoni & Tarp, Sven. 2024. *Introducción a la Lexicografía en Español: Funciones y Aplicaciones*. London: Routledge.
- RAE. 2009. *Nueva gramática de la lengua española*. Madrid: Espasa. <https://www.rae.es/gram%C3%A1tica/>
- RAE & ASALE. 2024. *Diccionario panhispánico de dudas*. <https://www.rae.es/dpd/> [consultado 15 de julio, 2024]
- RTVE. n.d. Errores de concordancia de género y número. *Manual de Estilo de RTVE. Directrices para los profesionales*. <https://manualdeestilo.rtve.es/el-lenguaje/6-5-los-verbos/6-5-6-errores-de-concordancia-de-genero-y-numero/>
- Schmidt-Fajlik, Ronald. 2023. ChatGPT as a Grammar Checker for Japanese English Language Learners: A Comparison with Grammarly and ProWritingAid. *AsiaCALL Online Journal*, 14(1): 105-119. <https://doi.org/10.54855/acoj.231417>
- SinoELE. *Revista de enseñanza de español lengua extranjera a sinohablantes*. <https://www.sinoele.org>
- Song, Cuiping & Song, Yanping. 2023. Academic Writing Skills and Motivation: Assessing the Efficacy of ChatGPT in AI-assisted Language Learning for EFL Students. *Frontiers in Psychology*, 14: 1-14. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Tarp, Sven. 2018. El desafío de los datos sintácticos en los diccionarios para aprendices extranjeros del español. *Revista de Filología de la Universidad de La Laguna*, 36: 367-391. <http://riull.ull.es/xmlui/handle/915/8624>
- Tarp, Sven. 2022a. A Lexicographical Perspective to Intentional and Incidental Learning: Approaching an Old Question from a New Angle. *Lexikos*, 32(2): 203-222. <https://doi.org/10.5788/32-2-1703>
- Tarp, Sven. 2022b. Turning Bilingual Lexicography Upside Down: Improving Quality and Productivity with New Methods and Technology. *Lexikos*, 32: 66-87. <https://doi.org/10.5788/32-1-1686>
- Tarp, Sven. 2023. Eppur si muove: Lexicography Is Becoming Intelligent! *Lexikos*, 33(2): 107-131. <https://doi.org/10.5788/33-2-1841>
- Tarp, Sven & Nomdedeu-Rull, Antoni. 2024. Who Has the Last Word? Lessons from Using ChatGPT to Develop an AI-based Spanish Writing Assistant. *Círculo de lingüística aplicada a la comunicación*, 97: 309-321. <https://doi.org/10.5209/clac.91985>

- Valverde, Pilar. 2023. El corpus de aprendices japoneses CELEN y su aplicación a la docencia y la investigación en ELE. *TEISEL. Tecnologías para la investigación en segundas lenguas*, 3: 1-31. <http://doi.org/10.1344/teisel.v3.42898>
- Wu, Leiping. 2024. AI-based Writing Tools: Empowering Students to Achieve Writing Success. *Advances in Educational Technology and Psychology*, 8(2): 40-44. <http://dx.doi.org/10.23977/aetp.2024.080206>
- Yamada, Aron; Davidson, Sam; Carando, Agustina; Sánchez-Gutiérrez, Claudia H.; Sagae, Kenji & Fernández-Mira, Paloma. 2020. COWS-L2H: A corpus of Spanish learner writing. *Research in Corpus Linguistics*, 8(1): 17-32. <https://doi.org/10.32714/ricl.08.01.02>