http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

Using Generative AI to Provide High-Quality Lexicographic Assistance to Chinese Learners of English

Qian Li, *Centre for Lexicographical Studies, Guangdong University of Foreign Studies, China (lqchristina@gdufs.edu.cn) (https://orcid.org/0009-0002-8267-7762)*
and
Sven Tarp, *Centre for Lexicographical Studies, Guangdong University of Foreign Studies, China; Department of Afrikaans and Dutch, Stellenbosch University, South Africa; and Aarhus University, Denmark (st@cc.au.dk) (https://orcid.org/0000-0003-1941-9082)*

Abstract: This paper reports on a research project that aims to explore how and to what extent generative AI can be used to produce different types of explanations that can be activated in writing assistants for Chinese learners of English. It first places the project in a lexicographic context and describes the general methodology used, including the limited usefulness of a learner corpus as an empirical basis and the need to use ChatGPT as a supplement to determine the error subcategories to be explained. As a result, 26 error sub-categories are identified within the main category of subject–verb disagreement. The paper then compares two generative AI chatbots, Baidu's Ernie Bot and OpenAI's ChatGPT, and describes how the latter was found to be more efficient and therefore prompted by lexicographers with experience in second-language teaching to write long explanations for each of the error sub-categories, with several examples demonstrating both the chatbot's remarkable performance and the constant need for human supervision and intervention. At the same time, the paper argues for the integration of generative AI directly into writing assistants to produce short default explanations for errors found in learners' texts. Finally, the paper summarises the findings, including the complex relationship between human and artificial intelligence.

Keywords: AUTOMATIC ERROR CORRECTION, CHATBOTS, ERROR EXPLANATIONS, FREQUENCY CRITERIA, GENERATIVE AI, L2 LEARNING, LANGUAGE MODELS, LEARNER CORPUS, MODERN GLOSSES, WRITING ASSISTANTS

Opsomming: Die gebruik van generatiewe KI om hoëkwaliteit leksikografiese hulp aan Chinese aanleerders van Engels te bied. In hierdie artikel word verslag gelewer oor 'n navorsingsprojek wat daarop gemik is om te ondersoek hoe en tot watter mate generatiewe KI gebruik kan word om verskillende tipes verklarings te verskaf wat in skryfhulpmiddels vir Chinese aanleerders van Engels geaktiveer kan word. Die projek word eerstens in 'n leksikografiese konteks geplaas en die algemene metodologie wat gebruik word, word beskryf. Die

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

beperkte bruikbaarheid van 'n leerderkorpus as empiriese basis en die behoefte aan die gebruik van ChatGPT as 'n hulpmiddel om die foutsubkategorieë wat verklaar moet word te bepaal, word hierby ingesluit. Dit het tot gevolg dat 26 foutsubkategorieë binne die hoofkategorie van onderwerp–werkwoord-kongruensie geïdentifiseer is. Twee generatiewe KI-kletsbotte, Baidu se Ernie Bot en OpenAI se ChatGPT, word dan met mekaar vergelyk, en daar word beskryf hoe laasgenoemde meer doeltreffend bevind is. Daarom is ChatGPT deur leksikograwe met ervaring in tweedetaalonderrig versoek om lang verklarings vir elk van die foutsubkategorieë te skryf, wat verskeie voorbeelde insluit wat beide die kletsbot se merkwaardige werkverrigting en die konstante behoefte aan menslike toesig en intervensie demonstreer. Terselfdertyd word die direkte integrasie van generatiewe KI in skryfhulpmiddels bepleit om kort verstekverklarings vir foute wat in leerders se tekste gevind word, te lewer. Laastens word die bevindings, insluitend die komplekse verhouding tussen menslike en kunsmatige intelligensie, opgesom.

Sleutelwoorde: OUTOMATIESE FOUTKORRIGERING, KLETSBOTTE, FOUTVERKLARINGS, FREKWENSIEKRITERIA, GENERATIEWE KI, L2-LEER, TAALMODELLE, AANLEERDERSKORPUS, MODERNE GLOSSE, SKRYFHULPMIDDELS

1. Introduction

In his reflections on the future of lexicography, and in response to Grefenstette's (1998) famous question of whether there will be lexicographers in the year 3000, Rundell (2012: 18) optimistically predicts that there will still be lexicographers, but that they will be doing something different from what their 21st century colleagues are doing. We agree with Rundell in principle, but we would like to emphasise even more that the lexicographers of the future will not only carry out their work using different methods and techniques from those of today. The results of this work are also likely to be presented to future users in entirely new ways. As McArthur (1986) has shown, lexicography has undergone similar shape-shifting from time to time in its millennia-long evolution from clay tablet to computer, and there is no reason to doubt that it will not do so again in the future. With this in mind, Tarp and Gouws (2023) have proposed a redefinition of the discipline of lexicography to include not only dictionaries but also glosses, both the traditional ones from which dictionaries have evolved according to Hanks (2013) and Benati and Händl (2019), and the new ones that are emerging, supported by cutting-edge technologies and integrated into writing and reading aids as well as other kinds of digital software. Tarp and Gouws (2023: 439) therefore recommend that lexicographers should:

> shift their focus from dictionaries to databases containing both new and old types of lexicographical data that can serve various tools, including but not limited to digital dictionaries.

From this perspective, Tarp and Gouws distinguish between two different categories of lexicographic databases that can already be observed in practice, namely "traditional" *lemma-centred databases* and new *problem-centred databases*. The latter

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

do not focus on specific words (lemmas), but on classes of grammatical, orthographic and stylistic challenges and problems that appear in texts. Because of these characteristics, problem-centred databases cannot support dictionaries as we know them, as they only contain data (glosses) that can be visualised in various digital tools to explain problems and help solve language challenges. These glosses are not related to specific words, but to specific types of problems, usually associated with a wider group of words.

It is obvious that the preparation, organisation and usefulness of problem-centred databases are much less studied than those of lemma-centred databases. However, it is not just that the latter have been around for longer. It is also a matter of taking a broader view of lexicography and breaking new ground. An example of this is the increasing use of Generative Artificial Intelligence (AI) in the discipline, especially after the launch of OpenAI's ChatGPT in November 2022. To date, most of the academic publications on the subject have focused on the use of this new technology to perform various tasks related to dictionary making; see, for instance, Alonso-Ramos (2023), Jakubíček and Rundell (2023), Lew (2023), Phoodai and Rikk (2023), Rees and Lew (2023), Rundell (2023), De Schryver (2023), and McKean and Fitzgerald (2024). So far, Huete-García and Tarp (2024), Li, Tarp and Nomdedeu-Rull (2024), and Tarp and Nomdedeu-Rull (2024), who are all concerned with the creation of lexicographic data to be used in writing assistants, are among the few exceptions to this trend. And the same can be said of Abdullayeva and Muzaffarovna (2023), Song and Song (2023) and Wu (2024), who approach writing tools from a different disciplinary perspective.

Against this background, we have conducted a research project to explore how and to what extent generative AI can be applied to produce different types of glosses — hereafter referred to as explanations — that can be activated in writing assistants for Chinese learners of English. The hypothesis is that this technology can increase productivity, at least without compromising quality, but probably improving it as well. This hypothesis is based on some reflections made by Huete-García and Tarp (2024), who experimented with ChatGPT to develop a writing assistant for learners of Spanish. The two researchers make a distinction between oral and written communication from teacher to student. On the one hand, they note that experienced Spanish teachers can easily explain the different types of language problems and challenges to their students in class. On the other hand, however, Huete-García and Tarp (2024: 36) observe that:

> it is less straightforward to write a concise explanation that gets to the heart of the matter in a language that is easily understood by the target audience. In addition to selecting the key aspects to be covered, determining the most appropriate and pedagogical structure can be quite time-consuming.

They therefore recommend using ChatGPT for this task, but only as an inspiration, as experienced teachers or lexicographers should always have the last say. Li et al. (2024), who have further developed and tested this way of writing explanations, define it as a "necessary symbiosis" between human and artificial intelli-

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

gence. As lexicographers with experience in second-language teaching, we can easily recognise ourselves in the above description and have therefore adopted the same approach in our project.

In the next section, we will briefly explain the overall methodology used to carry out the project, including why we have based the work on an English corpus containing errors made by Chinese learners of different proficiency levels. Section 3 describes how some of the error types to be explained are determined. Section 4 reports on the main part of the project, i.e. the direct work with generative AI, where two different chatbots are used to generate explanations and their efficiency is compared. Section 5 then summarises the main findings and presents the general conclusions, together with some reflections on future work.

2. Methodology

The lexicographic glosses, i.e. the explanations that are the subject of the research project described here, cannot be planned, produced or evaluated without knowing exactly how they will be used and what specific purpose they will serve. So, the very first step to be taken is to clearly identify and define that purpose, i.e. *who* might need the explanations, *for what* they might need them, *in what situation* they might need them, and *in what technological environment* the need might arise.

Now, the explanations are intended to help Chinese beginner and intermediate learners of English who are writing English texts using an AI-based writing assistant, similar in many ways to Grammarly or ProWritingAid (see Fitria 2021, 2023), but unlike these, bilingual with explanations in Chinese, i.e. the target users' native language. It is trained to identify and highlight possible problems and suggest alternative solutions that Chinese learners may want to understand in more detail as part of their English learning process. *Helping the learners achieve this deeper understanding is the genuine purpose of the explanations*.

Thus, unlike Wiegand's (1987) classic concept of "genuine purpose", which refers to a dictionary as a whole, here it refers only to the explanations, but not to the writing assistant as such. The reason for this is that the assistant has a broader purpose related to the writing process. Apart from simple text correction, the alternative suggestions generated by the underlying language model for this specific purpose are presented in a way that supports *incidental learning*, as defined primarily in relation to reading by Krashen (1989), Shu, Anderson and Zhang (1995) and Hulstijn (2013), among others, and later adapted to writing and even lexicography by Graham (2020) and Tarp (2022), respectively. Finally, as an additional service to motivated learners, the design also allows them to move on to *intentional learning* — see Leow and Zamora (2017) — if they decide to access and read the detailed explanations that are the subject of this paper.

All this suggests that the explanations should be written as *short didactic texts in plain language, without too much technical terminology, providing the most*

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

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.

As mentioned in the previous section, writing short didactic texts with these characteristics can be time-consuming, even for experienced and knowledgeable second-language teachers, as it usually requires some prior in-depth reflection on content, style and structure. It might therefore be interesting to explore whether, how and to what extent lexicographers can benefit from generative AI in this task. For this purpose, two well-known chatbots were chosen, namely Baidu's *Ernie Bot* and OpenAI's *ChatGPT*. Both of them were instructed to write explanations of selected problems in both Chinese and English. This means that four different approaches or methods were used to test their performance for this specific purpose, after which their respective performances were compared. We are fully aware that Chinese generative AI chatbots like Ernie Bot are generally considered to be a year or two behind the most advanced Western ones, such as ChatGPT, but as the writing assistant in question is intended to correct errors made by Chinese learners of English, there may appear to be some deviation from this general "rule". In any case, generative AI is a technology that is developing almost exponentially, and much is expected to change in the next few years. For now, the initial hypothesis was that Ernie Bot would be more efficient at writing Chinese than English, and ChatGPT would be more efficient at writing English than Chinese. But regardless of whether this turns out to be true or false, the comparison of the four methods provided evidence to better determine the most advantageous way to produce the explanations using current technology, i.e. either writing them directly in Chinese, or writing them in English and then translating them into Chinese, a process that poses other challenges.

The *Chinese Learner English Corpus* (CLEC) was used to select the types of errors or problems to be explained. This corpus, compiled by Gui and Yang (2003), is currently the only tagged corpus in China that contains errors made by Chinese learners of English. It consists of a just over a million words divided into five parts of about 200,000 words each, according to the learner's proficiency level. It is a relatively small corpus for the specific task, but its size does not differ much from similar tagged corpora in other languages, as they are very time-consuming and costly to produce.

The use of some kind of learner corpus is definitely a must in order to identify typical learner errors that can be used and explained in didactic language tools. These corpora have been used in one way or another to develop numerous writing tools, as discussed by Bestgen and Granger (2011), Paquot (2012), Wanner, Verlinde and Alonso-Ramos (2013), Alonso-Ramos and García-Salido (2019), Frankenberg-García, Lew, Roberts, Rees and Sharma (2019), and Granger and Paquot (2022). The best type of corpus for this purpose is undoubtedly a tagged corpus with parallel correction of the errors detected, such as the Spanish one described by Davidson, Yamada, Fernández-Mira, Carando, Sánchez-Gutiérrez and Sagae (2020). There are two main types of tagged or parallel corpora, namely

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

those that contain errors made by real learners, and those in which these errors — also referred to as "noisy examples" — are introduced using different types of software, such as those presented by Xie, Genthial, Xie, Ng and Jurafsky (2018) and Zhao, Wang, Shen, Jia and Liu (2019). According to the former, it is now possible to "synthesize noisy examples that human evaluators" are "nearly unable to discriminate from nonsynthesized examples" (Xie et al. 2018: 626). This technique, which can easily generate parallel corpora of several million words, is clearly useful and practical for a whole range of purposes, as also pointed out by Huete-García and Tarp (2024) in relation to their Spanish writing assistant project. In this respect, the "corpus revolution in lexicography" celebrated by Hanks (2012), who himself contributed significantly to its success, has indeed entered a new phase with possibilities and perspectives that have yet to be fully explored.

Be that as it may, a corpus of synthesised "noisy examples" simply does not serve as the basis for writing explanations of the kind discussed in this paper. Learners make a large number of errors of different types, and so does the software that synthesises this "noise" and feeds it into a corpus. Many errors, especially typos and other misspellings, are quite banal and do not lend themselves to explanation. Even if these are eliminated, there will still remain a significant number of error types that will require considerable work and time to explain properly. It is therefore necessary to prioritise, which can be done on the basis of frequency starting with the most common types. To be meaningful, such a frequency determination can only be made from a corpus of human-made errors. Using the above "noise-synthesising" software for this purpose would be arbitrary and the results would not reflect the true frequency of real learners' errors.

This discussion is reminiscent of a similar discussion about frequency as a lemma selection criterion in a traditional dictionary project, best illustrated by Kilgarriff's (2013: 79) idea that "if a dictionary is to have N words in it, they should be the N words from the top of the corpus frequency list". This approach has been challenged by Trap-Jensen, Lorentzen and Sørensen (2014), among others, who argue that corpus frequency is not necessarily identical to look-up frequency. Consequently, Nomdedeu-Rull and Tarp (2024: 174) point to another empirical source, namely log files, which in some cases have recorded hundreds of millions of look-ups in online dictionaries and therefore provide a much more accurate picture of the most frequently consulted words from which the lemmas in a new dictionary project can be selected. In this case, just as in the case above, the challenge is to put the real human users at the centre of the lexicographic work and to focus on their evidence-based needs.

To this end, the *Chinese Learner English Corpus* was used together with the AntConc corpus tool. This allowed us to determine the frequency of the specific error categories relevant to the project once they had been identified. How this was done, as a symbiosis of corpus search, knowledge and generative AI, will be discussed in the next section.

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

3. Determining error categories

The *Chinese Learner English Corpus* groups all identified errors into eleven major domains (*word formation*, *verb phrase*, *noun phrase*, *pronoun*, *adjective phrase*, *adverb*, *preposition*, *conjunction*, *lexical*, *collocational* and *syntactic*). These domains are further divided into 61 general categories, all of them at a very high level of abstraction. As such, they do not lend themselves to explanation in the form of short didactic texts which, as defined above, provide "the most relevant information about the specific language problem". If the 61 categories were explained as they are, such explanations would be either far too long or far too general for learners who only want to know more about a specific problem highlighted by a writing assistant in a text they have written. So, we had to break them down into sub-categories that were more suitable for explanation.

An example of this is *agreement* (or *concord*), which the corpus records as a frequent problem for Chinese learners of English, and which is treated as three different categories under the domains of *verb phrase*, *noun phrase* and *pronoun*, respectively. Each of these categories comprises several sub-categories that are not listed separately in the corpus. Not all of them have the same frequency, but in order to work systematically it was necessary to identify them and then group them under the respective categories, just as Bestgen and Granger (2011: 239-240) did with spelling errors. For the specific purpose of this paper, we chose the overall error category SUBJECT–VERB DISAGREEMENT under the *verb phrase* domain. However, our method differed from theirs in that we decided to experiment with ChatGPT and use it as an inspiration to speed up the sub-categorisation process:

1. First, we asked it to give us a list of relevant error types, which it did surprisingly well.
2. We then clicked the regenerate button to see if it would give us more useful suggestions, which it did in most cases. We kept doing this until it started repeating itself and nothing new came up.
3. If the result was not satisfactory, we also tried modifying or rewriting the prompt to improve the output.
4. The next step was to use our own grammatical knowledge and teaching experience to add a few more error types or to split some of the ones the chatbot provided into two.
5. As proof of the pudding, we consulted the corpus to see if it actually contained errors belonging to all the proposed sub-categories. If it did not, the sub-category was ignored for the time being.
6. Finally, we refined the terms used to describe the different sub-categories, as the chatbot was not consistent in this regard.

The end result was the following list of 26 sub-categories under the overall error category SUBJECT–VERB DISAGREEMENT:

— **Nouns:** Singular noun + plural verb
— **Nouns:** Plural noun + singular verb

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

— **Compound subjects:** Compound subjects joined by "and" + singular verb
— **Compound subjects:** Compound subjects joined by "or" or "nor" with nearest subject in singular + plural verb
— **Compound subjects:** Compound subjects joined by "or" or "nor" with nearest subject in plural + singular verb
— **Proper nouns:** Title of book, film and other works + plural verb
— **Uncountable nouns:** Uncountable noun + plural verb
— **Uncountable nouns:** Uncountable noun ending in "s" + plural verb
— **Infinitives:** Single infinitive as subject + plural verb
— **Gerunds:** Single gerund as subject + plural verb
— **Gerunds:** Two or more gerunds as subject + singular verb
— **Personal pronouns:** Third-person singular personal pronoun + plural verb
— **Personal pronouns:** Personal pronoun except for third-person singular + singular verb
— **Personal pronouns:** Personal pronoun + verb "to be" inflected in wrong person in present tense
— **Personal pronouns:** Personal pronoun + verb "to be" inflected in wrong person in past tense
— **Indefinite pronouns:** Singular indefinite pronoun + plural verb
— **Indefinite pronouns:** Plural indefinite pronoun + singular verb
— **Demonstrative pronouns:** Singular demonstrative pronoun + plural verb
— **Demonstrative pronouns:** Plural demonstrative pronoun + singular verb
— **Relative pronouns:** Relative pronoun with singular referent + plural verb
— **Relative pronouns:** Relative pronoun with plural referent + singular verb
— **Formal subject "there":** There + singular verb + plural real subject
— **Formal subject "there":** There + plural verb + singular real subject
— **Adverb "here":** Here + singular verb + plural subject
— **Adverb "here":** Here + plural verb + singular subject
— **Additions to subject beginning with "as well as", "together with", "along with", etc.:** Singular subject + addition + plural verb

It is important to note that the above is not a traditional linguistic classification, but one that considers only those cases where it is possible to unambiguously explain the respective sub-categories. In this respect, it should also be noted that the list does not include all the problems we are aware of. *Collective nouns*, for instance, are not included in the list, because, depending on what the writer wants to express, they can be used with both singular (most of the time) and plural verbs, but the technology to distinguish between these two behaviours with reasonable accuracy is not yet available. However, should the language model, after being trained, start to occasionally highlight and suggest alternatives to some verbs in relation to collective nouns, an explanation can be prepared that presents the general rules for this type of agreement, without taking a position on the specific suggestion. And the same can be done in some other cases, such as *subjects made up of two or more infinitives*, which usually require a plural verb,

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

but can also be combined with a singular verb in special circumstances.

Since the described method is not induction from existing empirical data, but a combination of deduction and generative AI, it cannot be excluded that some less frequent sub-categories of the 61 general categories registered in the corpus may be overlooked. However, this should not be a major concern, as they can be easily detected and added when the prototype of the writing assistant is ready to be tested on real users. Although we have no hard facts as such to prove it, we are convinced that the chosen method significantly speeds up the identification process. Once this identification of relevant error sub-categories for our research project was done, we started prompting the two chatbots to write the corresponding explanations. This proved to be a new and unexpected challenge, which will be discussed in the next section.

4. Generating explanations

As mentioned in Section 2, our intention was to find out both whether Baidu's *Ernie Bot* or OpenAI's *ChatGPT* produce the best explanations, and whether these explanations could actually be considered useful for Chinese learners who might need to know more about specific problems identified in one of their texts. Since the ultimate goal is to provide explanations in the learners' native language, both chatbots were instructed to write explanations both directly in Chinese and in English for subsequent translation into Chinese, and the respective results were then compared. To this end, we first designed a default English prompt based on the experience of a similar research project conducted by Li et al. (2024):

> I am training a language model to be used for didactic purposes. I want you to briefly and didactically explain to a Chinese learner of English [problem type]. I need the text to have the following structure 1) an introduction to the problem, 2) an example sentence with an error, 3) the same example sentence with the error corrected and 4) a brief explanation of the incorrect and correct examples. The example sentences must be at least 12 words long.

This prompt was then adapted into Chinese and later modified whenever the respective chatbots' responses were deemed unsatisfactory, either in terms of content, style or terminology used.

4.1 Ernie Bot disappoints

It soon became clear that Ernie Bot had much more difficulty than ChatGPT in producing acceptable explanations, whether in English or Chinese. In some cases, its responses to our prompts contained only a few minor problems that were relatively easy to correct. In other cases, however, it seemed to have completely misunderstood the grammatical issue raised, even though we modified the respective prompts several times. For example, when asked to write an expla-

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

nation of the relatively simple error of using a plural verb together with a singular noun as the subject, it first gave the "incorrect" example: *The boy loves playing footballs and basketballs*. It then corrected this example to: *The boy loves playing football and basketball*, explaining it as follows:

> In the incorrect example, 'The boy loves playing footballs and basketballs', the subject 'The boy' is singular, but the verb 'loves playing' is correctly paired with a singular subject. However, the problem lies in the objects of the verb, 'footballs' and 'basketballs'. Both of these nouns are incorrectly pluralized, as the sports 'football' and 'basketball' are typically treated as singular concepts when referring to the game or sport itself, not individual balls. But more importantly for this exercise, the focus is on the subject–verb agreement. Even if the objects were correctly singular, the subject–verb agreement is not the issue here. The critical point is to emphasize that a singular subject should not be mistakenly thought to require a plural verb.

So, despite being instructed to deal with problems of subject–verb agreement, Ernie Bot first gives two examples of correct agreement in the singular (i.e. no incorrect example of this problem), then claims that "the problem lies in the objects of the verb" because "the subject–verb agreement is not the issue here", and finally states that the "critical point" is precisely this agreement! Apart from the two unhelpful example sentences, such nonsensical explanations would take too much time and effort to reformulate and make useful for the specific purpose. We therefore decided to ditch Ernie Bot for the time being and continue with ChatGPT alone.

This disappointing conclusion could be due to a number of reasons, firstly, as mentioned in Section 2, that Ernie Bot still lags behind ChatGPT from a technological or programming point of view. But it could also be that this large language model has not been trained on a sufficient amount of English grammar texts compared to its Western counterpart, and that the result would be different if the purpose was to explain Chinese grammar. To test this last hypothesis, we asked both chatbots to write Chinese explanations of some grammatical issues in Chinese texts, similar to those in English. However, the respective responses show that ChatGPT is also qualitatively a step ahead of Ernie Bot when it comes to explaining Chinese grammar, so the hypothesis turned out to be false, or at least premature.

4.2 ChatGPT passes the test

ChatGPT was tested with its version 4o. Once we learned how to prompt it in the most appropriate way, the chatbot's responses were generally of a quality that could be easily used for our purposes, although varying degrees of editing were required. For instance, in some cases, both in English and Chinese, the explanations generated contained some "noise" with comments and even whole sen-

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

tences that only served to obscure the message without adding anything new and important to the specific issue being addressed:

— This ensures that the sentence is grammatically accurate and clearly conveys the intended meaning.
— This agreement ensures clarity and accuracy in the sentence.
— This ensures that the subject and verb agree, making the sentence clearer and grammatically correct.
— This can lead to confusion and grammatical errors.
— This common mistake can lead to grammatical errors.

The above are just a few examples of excessive verbosity that does not fit the genre and purpose. The last sentence is even nonsensical, since a mistake can not "lead to" an error, but is by definition an error. However, such unnecessary verbiage, which makes the explanations too long and therefore less readable and didactic, could easily be deleted in a few seconds.

In other cases, the example sentences also contained unnecessary wording that could distract learners from the real issue. An example of this is when ChatGPT was asked to explain the incorrect use of a singular verb with a plural noun as subject:

— The dogs runs in the park every morning and play with each other happily.

This sentence includes two verbs (*runs* and *play*) that refer to the same subject (*dogs*), but only the first verb has the incorrect singular form, which requires explanation. The second part of the sentence is therefore superfluous and likely to confuse learners. Therefore, to present a more didactic example, this part of the sentence was removed leaving only the following short incorrect sentence to be corrected:

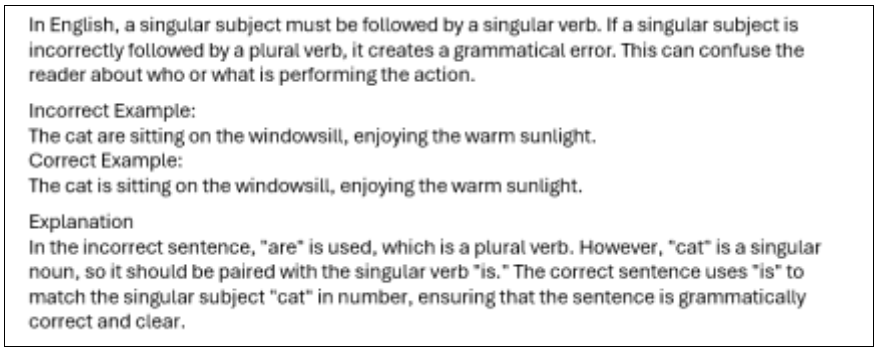
— The dogs runs in the park every morning.

Finally, ChatGPT was sometimes sloppy in its formulations, for example when it correctly stated in various explanations that a verb should match its subject, but did not make it clear that this match was in terms of number. Again, this problem could easily be solved by simply adding "in number" to avoid any misunderstanding. And a similar method could be used to correct other minor but distracting inaccuracies in the wording.

All in all, despite its shortcomings, ChatGPT has passed the test, thanks to its fruitful symbiosis with experienced and open-minded lexicographers who patiently instruct and correct it, always having the last word. After the initial, somewhat frustrating process of learning how to handle it, the time needed to write the desired explanations was significantly reduced, and in most cases considered to be much shorter than if these explanations had to be conceived from scratch without the inspiration of this generative AI tool. However, this does not exclude that in a very small number of cases a partial or complete reworking of

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

the explanations generated by the chatbot could be beneficial in order to improve their quality, but even this appeared to be relatively straightforward, as the chatbot's original explanation was an inspiration in both a positive and negative sense. Figures 1 and 2 show examples of an original ChatGPT-generated and a modified explanation of noun subject–verb agreement, aimed primarily at learners at the beginner level.

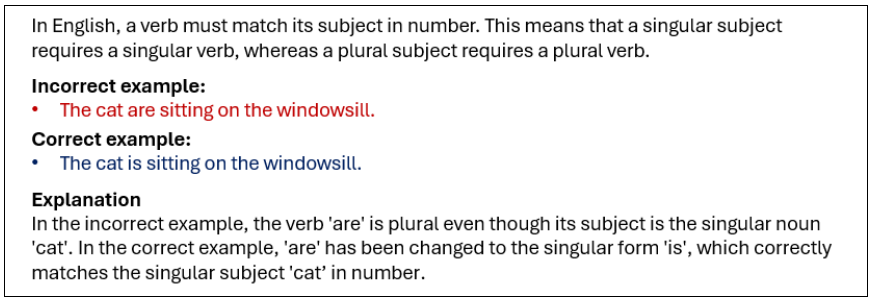

In English, a singular subject must be followed by a singular verb. If a singular subject is incorrectly followed by a plural verb, it creates a grammatical error. This can confuse the reader about who or what is performing the action.

Incorrect Example:
The cat are sitting on the windowsill, enjoying the warm sunlight.
Correct Example:
The cat is sitting on the windowsill, enjoying the warm sunlight.

Explanation
In the incorrect sentence, "are" is used, which is a plural verb. However, "cat" is a singular noun, so it should be paired with the singular verb "is." The correct sentence uses "is" to match the singular subject "cat" in number, ensuring that the sentence is grammatically correct and clear.

Figure 1: ChatGPT's explanation of noun subject–verb agreement

Apart from the obligatory layout improvements, there are three major differences between the modified explanation in Figure 2 and the original one in Figure 1. First, the introduction has been completely rewritten and now includes a short presentation of the general rule for subject–verb agreement. Second, the example sentences have been shortened to put more focus on the specific problem, by deleting the superfluous present participial phrase *enjoying the warm sunlight*. And third, the argumentation structure in the final explanation has been adjusted so that the reference to the incorrect example deals only with the error, while the reference to the correct example focuses on how it has been changed to achieve a grammatically correct sentence.


In English, a verb must match its subject in number. This means that a singular subject requires a singular verb, whereas a plural subject requires a plural verb.

Incorrect example:
- The cat are sitting on the windowsill.

Correct example:
- The cat is sitting on the windowsill.

Explanation
In the incorrect example, the verb 'are' is plural even though its subject is the singular noun 'cat'. In the correct example, 'are' has been changed to the singular form 'is', which correctly matches the singular subject 'cat' in number.

Figure 2: Modified explanation of noun subject–verb agreement

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

4.3 Writing Chinese explanations

After the first round of testing, it was apparent that ChatGPT performed just as well in Chinese as it did in English when asked to explain typical learner errors in English grammar. Since it has probably not been trained on the same amount of Chinese as English texts dealing with English grammar, this is most likely due to some sort of internal translation process, as the explanations in Chinese, and even the example sentences provided, were almost identical to those in English. In any case, once this became clear, we decided to continue with prompts and explanations in Chinese only. Thus, the examples provided in this section were all originally written in Chinese and then translated for the benefit of English-speaking readers.

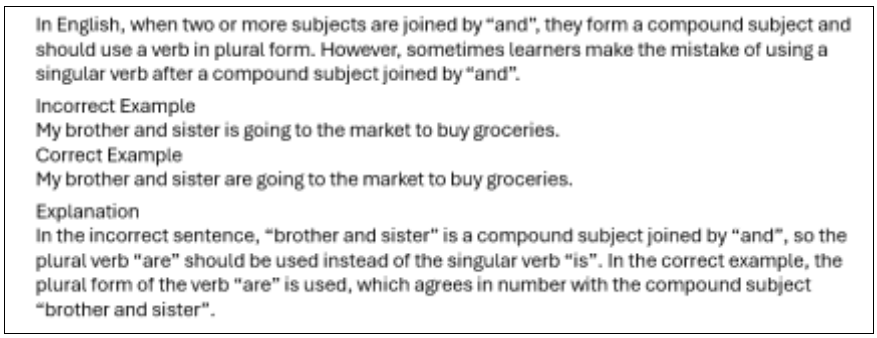
Figure 3 reproduces ChatGPT's response in Chinese after being prompted to explain why it is a mistake to pair singular verbs with compound subjects, and Figure 4 then provides the translation of this explanation.


在英语中，当两个或多个主语通过“and”连接时会构成复合主语，应该使用复数动词。然而，有时学习者会错误地在用“and”连接的复合主语后面使用单数动词。

错误例句
My brother and sister is going to the market to buy groceries.
正确例句
My brother and sister are going to the market to buy groceries.

解释
在错误例句中，“brother and sister”是通过“and”连接的复合主语，因此应该使用复数动词“are”，而不是单数动词“is”。正确例句中，使用了动词的复数形式“are”，这与复合主语“brother and sister”在数量上保持一致。

Figure 3: Chinese explanation of compound subject–verb agreement


In English, when two or more subjects are joined by “and”, they form a compound subject and should use a verb in plural form. However, sometimes learners make the mistake of using a singular verb after a compound subject joined by “and”.

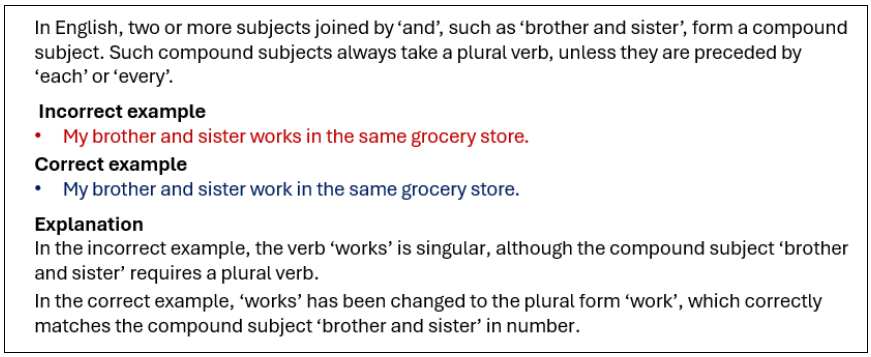
Incorrect Example
My brother and sister is going to the market to buy groceries.
Correct Example
My brother and sister are going to the market to buy groceries.

Explanation
In the incorrect sentence, “brother and sister” is a compound subject joined by “and”, so the plural verb “are” should be used instead of the singular verb “is”. In the correct example, the plural form of the verb “are” is used, which agrees in number with the compound subject “brother and sister”.

Figure 4: Translation of ChatGPT's explanation in Figure 3

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

As can be seen in Figure 4, there are some good points in ChatGPT's response, especially the initial explanation of what is meant by *compound subject* and what this type of subject requires of the paired verb in terms of number. However, given its intended audience and genre-specific purpose, there are several things that could and should be improved, as a comparison with the edited explanation in Figure 5 will clearly show. Firstly, there is some distracting "noise" that should be removed, such as the superfluous sentence beginning "However …" in the introduction and the redundant repetition of the definition of a compound subject in the final explanation. Secondly, the compound subject *brother and sister* used in the example sentences has been inserted into the definition of this grammatical category to illustrate what it refers to. Thirdly, the only two words (*each* and *every*) that override the general rule when they precede a compound subject are briefly mentioned. Fourthly, the argumentation structure in the final explanation has been made more logical and straightforward, similar to the refinement of the explanation in Figure 2. As the cherry on the cake, the two example sentences have also been modified, although this was not strictly necessary. The intention behind this move was simply to use a verb other than *to be*, as it turned out to be quite easy to come up with an alternative inspired by the words *groceries* and *market* in the original examples. Figure 5 shows the result of this careful editing, always keeping in mind the specific purpose and anticipated target users.


In English, two or more subjects joined by 'and', such as 'brother and sister', form a compound subject. Such compound subjects always take a plural verb, unless they are preceded by 'each' or 'every'.

Incorrect example

- My brother and sister works in the same grocery store.

Correct example

- My brother and sister work in the same grocery store.

Explanation

In the incorrect example, the verb 'works' is singular, although the compound subject 'brother and sister' requires a plural verb.

In the correct example, 'works' has been changed to the plural form 'work', which correctly matches the compound subject 'brother and sister' in number.

Figure 5: Modified version of ChatGPT's explanation in Figure 3

The revision and editing of the other test explanations followed the same general pattern as the one discussed above in relation to Figures 3, 4 and 5.

— A quick read through of the AI-generated explanations.
— Remove some background noise from the text so as not to distract the learner's attention from the main point.
— Change some words and phrases to improve readability.

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

— Select relevant data, such as subjects and verb forms, from the example sentences and add them to the introduction whenever it makes this section easier to read and helps to clarify the grammar problem being explained, especially if the use of some technical terminology is unavoidable.
— Refine the argumentation structure in the explanation of the incorrect and correct example sentences to make this section more logical, straightforward and concise.
— Improve the layout of the whole explanation to make it as easy as possible for the reader to quickly gain the necessary overview and grasp the essence of the problem being addressed.

As for the example sentences generated by the chatbot, although a few of them were shortened by deleting irrelevant wording to maximise the focus on the specific issue, in only one case was it considered beneficial to rephrase them, as in the explanation in Figure 5. It obviously took some thought, discussion and practice to develop and become familiar with the method described, but once it was internalised the whole editing process became quite straightforward and could be completed in much less time than it would have taken to write the explanations from scratch without inspiration from the chatbot. Although it is for others to judge, the end result can be considered both satisfactory and of the required quality.

4.4 Explanations at work

The final destination of the explanations discussed in the previous sections is their integration into a bilingual English writing assistant, similar to the bilingual Spanish one presented by Li et al. (2024), but different from theirs in that it is aimed exclusively at Chinese learners of English. The writing assistant will be supported by an AI-powered language model that has been trained to detect errors in written English. The idea is that beginner and intermediate learners can either paste their English texts into the tool or use it to write them, as in Chinese apps of the 1–7 Zuoye type, albeit with different functionality. The writing assistant will then highlight possible problems in the text. If the learner does not know why a particular word has been highlighted, he or she can simply click on the word in question to display a pop-up window with an alternative suggestion followed by a short explanation (see Figure 6). As can be seen, the problem in this case is noun subject–verb disagreement. Accordingly, the short Chinese text reads:

> The verb 'are' is plural, but must be singular to agree in number with the subject 'impact'.

This combination of highlighting, alternative suggestion and ultra-short explanation enhances the possibility of incidental learning as defined from a lexicographic perspective by Tarp (2022).

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

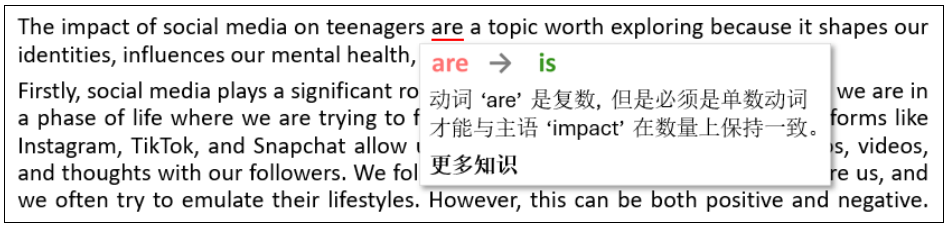


Figure 6: Pop-up window with alternative suggestion and short explanation

If the learner is satisfied with the information provided, a simple click on the green *is* will insert that verb form into the text instead of the highlighted *are*. If, on the other hand, the learner is a beginner who wants to know more about this fundamental grammatical issue in English, a click on the Chinese characters 更多知识 (LEARN MORE) in the bottom left corner of the pop-up window will open another window with a long supplementary explanation (see Figure 7).

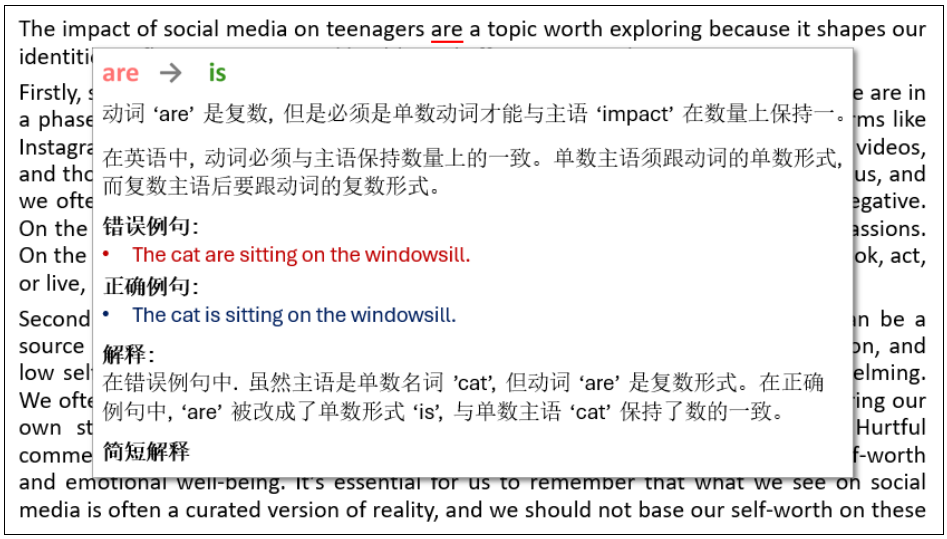


Figure 7: Pop-up window with supplementary explanation

The first line of the explanation in Figure 7 is a repetition of the short explanation from the pop-up window in Figure 6. The following text is the Chinese version of the English explanation shown in Figure 2, and the characters 简短解释 (LESS) in the bottom left corner indicate how to close the pop-up window.

The whole construction is based on the dialectical relationship between the Hegelian concepts of the *individual*, the *particular* and the *universal*. The under-

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

lined error "are" in the learner's text represents the individual, the short explanation the particular and the long explanation the universal. In this way the particular, i.e. the short explanation, acts as a mediator or bridge between the individual and the universal, allowing the learner to relate the long explanation to the error he or she has made and vice versa. This seemingly simple construction, but with complex underlying relationships, allows for intentional learning of English grammar and is primarily aimed at the motivated student who is eager to study, learn and make progress in English language acquisition.

4.5 In the borderland between the possible and the impossible

It should be emphasised that the short explanation in Figure 6 does not use phrases such as *It seems that*, *There may be* and *It looks like*, which have been employed in tools like Grammarly and ProWritingAid, at least until recently, and also by Li et al. (2024) in their proposal for a future writing assistant. The reason for using these and similar phrases is to avoid misinforming the user, as the underlying language models have not been entirely reliable in identifying possible errors and suggesting alternatives. This is about to change. The short explanation proposed in Figure 6 not only states directly that there is an error, but also explains the nature of the problem in very few words.

This new approach is driven not only by improved language models, but also by the integration of generative AI into the writing tool to support its functionality. To explore these new technological advances in the current borderland between the possible and the impossible, ChatGPT was asked to explain why it is a mistake to use the verb form *are*, highlighted as an error by the language model, in the sentence partially covered in Figures 6 and 7. Its response is shown in Figure 8, and as can be seen, it was perfectly able to identify both the subject (*impact*) and its number (*singular*) and compare it with the number (*plural*) of the verb (*are*), and on this basis explain the grammatical problem at hand.

To check whether this was more than a lone wolf, ChatGPT was then asked to explain errors belonging to the 26 sub-categories of subject–verb disagreement listed in Section 3. In all cases, without exception, it managed to identify both the subject and the verb, as well as their respective numbers, and explain the nature of the problem. Although 26 tests are not statistically sufficient to draw a definitive conclusion, the result (26 hits out of 26 possible) strongly suggests that generative AI has an important role to play in improving the quality of writing assistants.

However, as useful as they are, ChatGPT's responses cannot be uploaded directly into small pop-up windows, such as the one in Figure 6, which are intended to give the writer a very brief explanation of why one word or phrase should replace another. They are simply too long for that purpose, and sometimes they are not clear enough or use terminology that learners cannot be expected to know. Nevertheless, the information they provide about a problem and its nature can be of great value in preparing such short default explana-

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

tions. The best solution would therefore be to prepare an explanation template for each problem category, into which the relevant words, in this case *are* and *impact*, can be inserted, the first one provided by the language model (a technology that already exists) and the second one by the chatbot. How to extract this last information from the chatbot's responses and insert it into the explanations is a technical question. So, the challenge now lies with the computer experts, but lexicographers should also have their say and be prepared for this brave new world. The discussion above is part of that preparation.

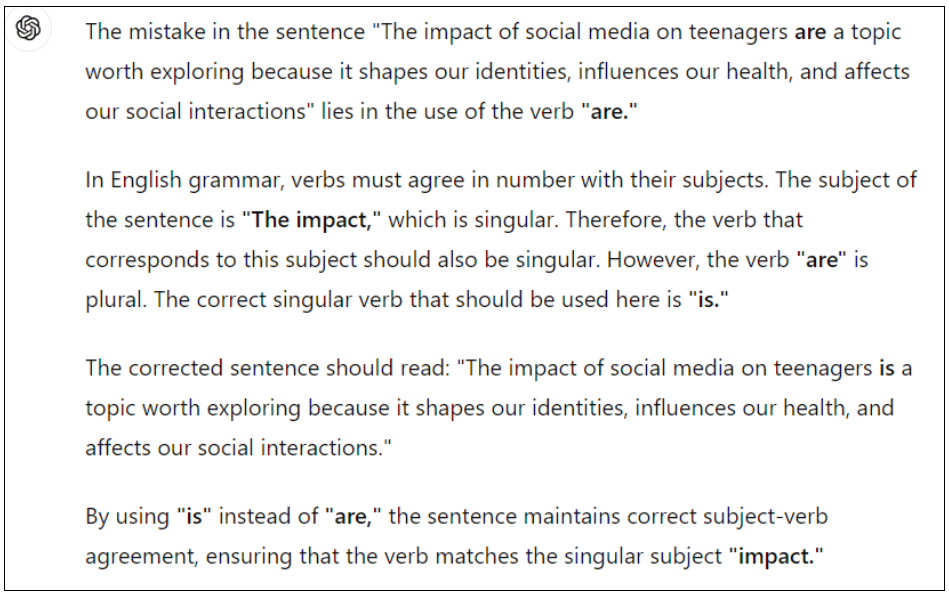


Figure 8: ChatGPT explains a subject–verb agreement error

5. Conclusions

The main objective of the research project discussed in this paper was "to explore how and to what extent generative AI can be applied to produce different types of explanations that can be activated in writing assistants for Chinese learners of English", based on the hypothesis that this technology can increase productivity without compromising quality.

The paper answered all three questions. It showed *how* generative AI could be used to perform three interrelated tasks in this perspective: (1) determining which error sub-categories should be explained and producing both (2) long and (3) ultra-short explanations of these sub-categories, developing and testing a methodology for each of them.

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

The paper also showed that the use of this technology could (1) significantly speed up the determination of error categories, but probably not to the same quality as if they were based on the much slower method of tagging and then analysing a learner corpus; (2) produce long explanations of the errors much faster, and at least to the same quality as if they had to be written from scratch by human lexicographers; and (3) contribute to the development of much more informative short explanations, and therefore to a radically different quality than those found in writing assistants to date.

Thus, the paper proved that the hypothesis that this technology could increase productivity without compromising quality was correct for long explanations, but not entirely for the prior detection of the error sub-categories to be explained, where quality is likely to be lower, while it was not possible to compare the productivity of short explanations, as those discussed did not yet exist to our knowledge.

Finally, the paper also reported on a comparative test of two different generative AI chatbots, Baidu's Ernie Bot and OpenAI's ChatGPT, both of which were asked to generate explanations in both Chinese and English. Our initial hypothesis that Ernie Bot would be more efficient at writing Chinese than English, and that ChatGPT would be more efficient at writing English than Chinese, proved to be wrong, or at least premature, as ChatGPT performed significantly better than Ernie Bot, with no qualitative difference between its English and Chinese explanations.

As an added bonus, the research project confirmed Tarp and Nomdedeu-Rull's (2024) conclusion that humans should always have "the last word", as generative AI chatbots are not entirely reliable. Like untamed dogs, they need to be kept on a short leash, despite being man's best friend.

One important thing to bear in mind when working with these tools is that they cannot do anything on their own. Beyond their intrinsic technical limitations, their actual performance depends entirely on their interaction with humans, and in particular on the prompts they receive from the latter. This implies that they can only perform at their best when properly handled by a human, in this case a lexicographer. This raises the question of what is required of the lexicographer in order to interact with and prompt the chatbot in an optimal way. Using generative AI chatbots is something that must be learned. Writing good prompts is not easy. Like any learning process, it takes time and a lot of practice. But apart from learning how to handle the chatbots, i.e. acquiring usage skills, it is also important to have a good grasp of the specific domain of knowledge that is the subject of the interaction. Without knowledge of English grammar and the errors that Chinese learners typically make when writing in English, it would be impossible to interact meaningfully with the chatbots and prompt them to improve the explanations they generate. User skills and domain knowledge are the two keys to success in working with this new technology.

Having said that, it is worth remembering the old English saying that dates back to at least the early 17th century: *The proof of the pudding is in the eating*. In

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

this case, the *pudding* is the explanations, the *proof* is testing them, the *eating* is using the writing assistant in which they are integrated, and those who *eat* it are its future target users. If they are not happy and satisfied, the lexicographers and programmers will have to go back to the drawing board. Meeting user needs is the ultimate quality criterion.

Acknowledgments

Special thanks are due to the Center for Lexicographical Studies at Guangdong University of Foreign Studies, for appointing Sven Tarp as Yunshan Chair Professor, thus making the collaboration between the authors of this article possible.

This study is supported by the Research Funding to Qian Li from Center for Linguistics and Applied Linguistics of Guangdong University of Foreign Studies for the project "Multi-modal Language Teaching for Students of Different Disciplines".

References

Abdullayeva, M. and M.Z. Muzaffarovna. 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.

Alonso-Ramos, M. 2023. El papel de ChatGPT como lexicógrafo. Garriga-Escribano, C., S. Iglesia-Martín, J.A. Moreno-Villanueva and A. Nomdedeu-Rull (Eds.). 2023. *Lligams: Textos dedicats a Maria Bargalló Escribà*: 15-27. Tarragona: Publicacions URV.

Alonso-Ramos, M. and M. García-Salido. 2019. Testing the Use of a Collocation Retrieval Tool Without Prior Training by Learners of Spanish. *International Journal of Lexicography* 32(4): 480-497.

Benati, C. and C. Händl (Eds.). 2019. *From Glosses to Dictionaries: The Beginnings of Lexicography*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Bestgen, Y. and S. Granger. 2011. Categorizing Spelling Errors to Assess L2 Writing. *International Journal of Continuing Engineering Education and Life-Long Learning* 21(2/3): 235-252.

Davidson, S., A. Yamada, P. Fernández-Mira, A. Carando, C.H. Sánchez-Gutiérrez and K. Sagae. 2020. Developing NLP Tools with a New Corpus of Learner Spanish. Calzolari, N. et al. (Eds.). 2020. *Proceedings of the Twelfth Language Resources and Evaluation Conference, May 11–16, 2020, Marseille, France*: 7238-7243. Marseille: European Language Resources Association.

De Schryver, G.-M. 2023. Generative AI and Lexicography: The Current State of the Art Using ChatGPT. *International Journal of Lexicography* 36(4): 355-387.

Fitria, T.N. 2021. Grammarly as AI-powered English Writing Assistant: Students' Alternative for Writing English. *Metathesis. Journal of English Language, Literature and Teaching* 5(1): 65-78.

Fitria, T.N. 2023. ProWritingAid as AI-Powered Writing Tools: The Performance in Checking Grammar and Spelling of Students' Writing. *Polingua. Scientific Journal of Linguistics, Literature and Language Education* 12(2): 65-75.

Frankenberg-García, A., R. Lew, J.C. Roberts, G.P. Rees and N. Sharma. 2019. Developing a Writing Assistant to Help EAP Writers with Collocations in Real Time. *ReCALL* 31(1): 23-39.

Graham, S. 2020. The Sciences of Reading and Writing Must Become More Fully Integrated. *Reading Research Quarterly* 55(S1): 535-544.

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

Granger, S. and M. Paquot. 2022. *The Louvain English for Academic Purposes Dictionary. User Manual*. Louvain: Centre for English Corpus Linguistics, Université Catholique de Louvain.

Grefenstette, G. 1998. The Future of Linguistics and Lexicographers: Will There Be Lexicographers in the Year 3000? Fontenelle, T., P. Hiligsmann, A. Michiels, A. Moulin and S. Theissen (Eds.). 1998. *Proceedings of the Eighth EURALEX in Liège, Belgium*: 25-41. Liège: English and Dutch Departments, University of Liége.

Gui, S.C. and H.Z. Yang. 2003. *Chinese Learner English Corpus*. Shanghai: Foreign Language Education Press.

Hanks, P. 2012. The Corpus Revolution in Lexicography. *International Journal of Lexicography* 25(4): 398-436.

Hanks, P. 2013. Lexicography from Earliest Times to the Present. Allan, K. (Ed.). 2013. *The Oxford Handbook of the History of Linguistics*: 503-536. Oxford: Oxford University Press.

Huete-García, Á. and S. Tarp. 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.

Hulstijn, J.H. 2013. Incidental Learning in Second Language Acquisition. Chapelle, C.A. (Ed.). 2013. *The Encyclopedia of Applied Linguistics:* 2632-2637. New York: Wiley-Blackwell.

Jakubíček, M. and M. Rundell. 2023. The End of Lexicography? Can ChatGPT Outperform Current Tools for Post-editing Lexicography? Medved', M., M. Měchura, C. Tiberius, I. Kosem, J. Kallas, M. Jakubíček and S. Krek (Eds.). 2023. *Electronic Lexicography in the 21st Century (eLex 2023): Invisible Lexicography. Proceedings of the eLex 2023 Conference, Brno, 27–29 June 2023*: 518-533. Brno: Lexical Computing CZ s.r.o.

Kilgarriff, A. 2013. Using Corpora as Data Sources for Dictionaries. Jackson, H. (Ed.). 2013. *The Bloomsbury Companion to Lexicography*: 77-96. London: Bloomsbury.

Krashen, S. 1989. We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *Modern Language Journal* 73: 440-464.

Leow, R.P. and C.C. Zamora. 2017. Intentional and Incidental L2 Learning. Loewen, S. and M. Sato (Eds.). 2017. *The Routledge Handbook of Instructed Second Language Acquisition*: 33-49. New York: Routledge.

Lew, R. 2023. ChatGPT as a COBUILD Lexicographer. *Humanities and Social Sciences Communications* 10:704.

Li, Q., S. Tarp and A. Nomdedeu-Rull. 2024. The Necessary Symbiosis: How ChatGPT Co-authored a New Type of Learner's Grammar. *Círculo de lingüística aplicada a la comunicación* 100. (To appear)

McArthur, T. 1986. *Worlds of Reference. Lexicography, Learning and Language from the Clay Tablet to the Computer*. Cambridge: Cambridge University Press.

McKean, E. and W. Fitzgerald. 2024. The ROI of AI in lexicography. *Lexicography* 11(1): 7-27.

Nomdedeu-Rull, A. and S. Tarp. 2024. *Introducción a la Lexicografía en Español: Funciones y Aplicaciones.* London: Routledge.

Paquot, M. 2012. The LEAD Dictionary-cum-writing Aid: An Integrated Dictionary and Corpus Tool. Granger, S. and M. Paquot (Eds.). 2012. *Electronic Lexicography*: 163-185. Oxford: Oxford University Press.

Phoodai, C. and R. Rikk. 2023. Exploring the Capabilities of ChatGPT for Lexicographical Purposes: A Comparison with Oxford Advanced Learner's Dictionary within the Microstructural Framework. Medved', M., M. Měchura, C. Tiberius, I. Kosem, J. Kallas, M. Jakubíček and S. Krek (Eds.). 2023. *Electronic Lexicography in the 21st Century (eLex 2023): Invisible Lexicography. Proceedings of the eLex 2023 Conference Brno, 27–29 June 2023*: 345-375. Brno: Lexical Computing CZ s.r.o.

http://lexikos.journals.ac.za; https://doi.org/10.5788/34-1-1944 (Article)

Rees, G.P. and R. Lew. 2023. The Effectiveness of OpenAI GPT-Generated Definitions Versus Definitions from an English Learners' Dictionary in a Lexically Orientated Reading Task. *International Journal of Lexicography* 37(1): 50-74.

Rundell, M. 2012. The Road to Automated Lexicography: An Editor's Viewpoint. Granger, S. and M. Paquot (Eds.). 2012. *Electronic Lexicography:* 15-30. Oxford: Oxford University Press.

Rundell, M. 2023. Automating the Creation of Dictionaries: Are we Nearly There? *Asialex 2023 Proceedings. Lexicography, Artificial Intelligence, and Dictionary Users, 22–24 June 2023, Seoul, Korea:* 9-17. Seoul: Yonsei University.

Shu, H., R.C. Anderson and H. Zhang. 1995. Incidental Learning of Word Meanings While Reading: A Chinese and American Cross-cultural Study. *Reading Research Quarterly* 30(1): 76-95.

Song, C. and Y. Song. 2023. Enhancing Academic Writing Skills and Motivation: Assessing the Efficacy of ChatGPT in AI-assisted Language Learning for EFL Students. *Frontiers in Psychology* 14: 1-14.

Tarp, S. 2022. A Lexicographical Perspective to Intentional and Incidental Learning: Approaching an Old Question from a New Angle. *Lexikos* 32(2): 203-222.

Tarp, S. and R.H. Gouws. 2023. A Necessary Redefinition of Lexicography in the Digital Age: Glossography, Dictionography and the Implications for the Future. *Lexikos* 33(1): 425-447.

Tarp, S. and A. Nomdedeu-Rull. 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.

Trap-Jensen, L., H. Lorentzen and N.H. Sørensen. 2014. An Odd Couple — Corpus Frequency and Look-up Frequency: What Relationship? *Slovenščina 2.0: Empirical, Applied and Interdisciplinary Research* 2(2): 94-113.

Wanner, L., S. Verlinde and M. Alonso-Ramos. 2013. Writing Assistants and Automatic Lexical Error Correction: Word Combinatorics. Kosem, I., J. Kallas, P. Gantar, S. Krek, M. Langemets and M. Tuulik (Eds.). 2013. *Electronic Lexicography in the 21st Century: Thinking Outside the Paper. Proceedings of the eLex 2013 Conference, 17–19 October 2013, Tallinn, Estonia*: 472-487. Ljubljana/Tallinn: Institute for Applied Slovene Studies/ Eesti Keele Instituut.

Wiegand, H.E. 1987. Zur handlungstheoretischen Grundlegung der Wörterbenutzungsforschung. *Lexicographica* 3: 178-227.

Wu, L. 2024. AI-based Writing Tools: Empowering Students to Achieve Writing Success. *Advances in Educational Technology and Psychology* 8(2): 40-44.

Xie, Z., G. Genthial, S. Xie, A. Ng and D. Jurafsky. 2018. Noising and Denoising Natural Language: Diverse Backtranslation for Grammar Correction. Walker, M., J. Heng and A. Stent (Eds.). 2018. *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL–HLT 2018, New Orleans, Louisiana, USA, June 1–6, 2018, Vol. 1 (Long Papers)*: 619-628. New Orleans: Association for Computational Linguistics.

Zhao, W., L. Wang, K. Shen, R. Jia and J. Liu. 2019. Improving Grammatical Error Correction via Pre-Training a Copy-Augmented Architecture with Unlabeled Data. Burstein, J., C. Doran and T. Solorio (Eds.). 2019. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL–HLT 2019, Minneapolis, MN, USA, June 2–7, 2019. Vol. 1 (Long and Short Papers):* 156-165. Minneapolis: Association for Computational Linguistics.