

ChatGPT: Detection of Spanish Terms Based on False Friends

Amal Haddad Haddad
University of Granada

amalhaddad@ugr.es

Damith Premasiri
Lancaster University

d.dolamullage@lancaster.ac.uk

Abstract

One of the common errors which translators commit when transferring terms from one language into another is erroneously coining terms which are based on a false friend mistake due to the similarity between lexical units forming part of terms. In this case-study, we use ChatGPT to automatically detect terms in Spanish which may be coined based on a false friend relation. To carry out this study, we implemented two experiments with GPT and compared the results. In the first, we prompted GPT to produce a list of twenty terms in Spanish extracted from the UN discourse, which are possibly based on false friend relation, and its English equivalents and analysed the veracity of the results. In the second experiment, we used an aligned corpus to further study the capabilities of the Language Model on detecting false friends in English and Spanish Text. Some results were significant for future terminological studies.

Keywords: ChatGPT; false friends; Translation of terms

1 Introduction

English is the lingua franca in science and technology (Faber, 2010), and most primary terms are coined in English and then transferred into other languages (Steurs and Kockaert, 2011; Deppecker, 2015). The most common techniques for the transfer of primary terms are literal translation or calques (Quesada et al., 2002). One of the common errors which human translators may commit when transferring terms among languages is using false friends. False friends are defined as “lexemes in at least two different languages with identical or similar graphic/spoken forms whose semantics overlap either partially or not at all” (Szpila, 2006). When the overlap is partial, or the so-called semantic false friends (Dominguez and Nerlich, 2002),

only parts of the meaning are maintained, while other parts are different or lost, which may lead to semantic and conceptual loss. For example, the lexical units *eventually* in English and *eventualmente* in Spanish are considered partial false friends, as they share the meaning of *something to happen*, but in English, the adjective denotes something that will be realised for sure at the end of a process or activity, while the Spanish adjective denotes something that may or may not happen (Dominguez and Nerlich, 2002). In cases of no overlap, or the so-called chance false friends (Dominguez and Nerlich, 2002), or false cognates (Trishina, 2022) the meaning is totally different and leads to erroneous understanding. For example, the lexical units *sensible* in English and *sensible* in Spanish, meaning *sensitive*.

Previous studies showed that false friends are the source of miscommunication not only in common languages, but also in scientific languages (Lewis, 2020; Makayev et al., 2020). For example, Lewis gives a list of false friends between English and Croatian in the legal domain. Yetkin shows through corpus analysis the frequent use of partial false friends between English and Turkish in the diplomatic discourse and its terminology. In this context, we define a term based on false friend as a secondary term which was transferred from one language into another by committing a false friend mistake in the transfer of at least one lexical unit forming part of the term, and the false friend inequivalent meaning may be based on either a total false friend or a partial false friend.

Lewis (2020) highlights the need to carry out more research to show how frequent is the use of false friends in coining terms at interdisciplinary level, and to provide more tools to avoid this translation problem not only in common languages but also in scientific and technical discourse to prevent miscommunication.

Makayev et al. (2020) also highlight the need to investigate the influence of false friends and false cognates in the construction of domains. However, when a term is coined, it starts spreading all over different countries and when academics discover that it is erroneously coined, it is too late by then to control the spread of those terms. Even more, in cases of a false friend relations in a term, in some cases, the meaning becomes part of the language and becomes part of the official dictionaries, which may have influence at conceptual level and understanding of phenomena.

One of the difficulties, in relation to this problem, is that identifying the terms which are based on a false-friend relation is not an easy task. Firstly, it is because users read scientific discourse in one language, and no comparison is carried out by users unless it was the subject of comparative study. In second place, it is important to take into consideration that this phenomenon is not as common as other linguistic phenomena, and it usually requires high linguistic competence by the users to detect them. It also needs a detailed and cautious examination not only of the term itself, but also of the conceptual systems underlying the term, in order to detect those terms based on false friends in discourse and whether they impede or obstacle comprehension or not depending on the level of semantic equivalence between the original word in the first language and the acquired meaning in the second language. One of the approaches to detect those conceptual layers is Frame-Based Terminology (FBT) (Faber et al., 2012; Faber, 2015). There are AI models which try to detect false friends in discourse (Mitkov et al., 2007; Palmero Aprosio et al., 2020), but to the best of our knowledge, there are no automatic tools or programmes, or AI models which focus on the automatic detection of terms in specialised domains which are based on false friend relations.

To fill this gap, we propose the use of ChatGPT to identify possible terms based on false friend relations in specific specialised domains between English and Spanish, then we propose corpus analysis to verify how widely spread the use of such false friends in specialised discourse is.

This paper is organised as follows: in Section 2, we give a brief introduction about the use of ChatGPT and other Large Language Models (LLMs) to generate relevant results in the field of Translation and Terminology in relation to false friends

and its detection; in Section 3 we provide information about the methodology and the experiments carried out to prompt ChatGPT and to obtain relevant and significant results; in Section 4 we discuss the results, while in Section 5 we draw on several conclusions and highlight new trends for future research.

2 ChatGPT and LLMs

Using NLP to carry out research in Translation and Terminology is gaining increasing popularity. For example, studies like Mitkov et al. (2007) carried out an experiment to extract false friends and cognates from monolingual corpora in the linguistic combination of English-French, English-German and English-Spanish. Castro et al. (2018) use word vector representations to build a false friends classifier implementing their research in the combinations of Spanish and Portuguese, while studies like Al-Athwary (2021) used Large Language Models (LLMs) to detect false friends in less similar languages like Arabic and English. Several studies also focused on the detection of cognates (Gamallo and Garcia, 2012; Batsuren et al., 2022; Kanojia et al., 2021). Other studies used BERT for the detection of multiword expressions and term extraction (Rigouts Terryn et al., 2020). In this case study, we used ChatGPT (GPT) as a tool for the automatic detection of terms based on false friends in specific domains. To our knowledge, no previous studies focused on this aspect or proved the efficiency of GPT in carrying out those sorts of tasks.

3 Methods and experiments

In this small-scale case study, we use GPT to detect terms in Spanish which may be coined based on a false friend relation in some specialised domains. To carry out this study, first we asked GPT to produce a list of false friends between English and Spanish. We started with this query to limit the search to only a selected set of false friends to be used in the coming step. Then we asked GPT to produce a list of terms in Spanish that contain any of the words produced in the previous list. This step allowed for collecting a list of terms in the domain of study which are possibly based on a false friend relationship, but they would indeed still need further analysis to detect whether this assumption is true or false. Afterwards, we asked GPT to produce a list of terms in Spanish equivalent to the English

terms, which are possibly the result of a false friend relationship. This step gave more insightful results after prompting GPT with other queries based on previous research, which will be discussed in Section 3.1. Finally, we studied some of the terms and compared their use in Corpus, and we detected terms which are erroneously coined in the domains of study, which may lead to difficulties in communication or comprehension in specialised discourse. For the comparison of exact meaning of lexical units and to detect whether the false friend relation is correct or not, we looked out the meaning of those lexical units in the Cambridge Online Dictionary (Cambridge University Press, Accessed 2024) and the Real Academia Española (RAE) online dictionary (Real Academia Española, Accessed 2024). With respect to the resources used to verify the frequency of use of some terms or if they really exist in specialised discourse or not, we used the open parallel corpus of the United Nations (UN) (Kilgarriff et al., 2014), available on Sketch Engine, and a search in Google Scholar to detect the use of such terms in scientific articles. Regarding the second experiment (Section 3.2), we used the parallel corpus of the UN available in the UNPC webpage to further study the capabilities of the Language Model on detecting false friends in English and Spanish Text.

3.1 First Experiment

Our first prompt to GPT aimed at providing a definition of the concept of false friends to guarantee the coherence of the study. The next query was asking GPT to provide a list of false friends between English and Spanish. We obtained a list of 89 false friends with the explanation of the reason behind considering them as false friends by GPT (The whole list is provided in Appendix A). Next, GPT was asked to provide a list of terms in the domain of Climate Change in English containing one of the lexical units mentioned in the previous answer, that may be part of a term based on a false friend (Appendix B). We selected the domain of Climate Change in our prompt to GPT as the prompting of GPT is based on a term based on a false friend in this domain. Afterwards, GPT was asked to provide the translation of the provided terms (Appendix E). GPT explained after providing the list that “none of the translated terms provided are based on false friend translations. In this case, the translations are accurate and do not involve

words that are false friends between English and Spanish”.

In the next step, and in order to train GPT so that it provides a list with terms that it does consider as based indeed on false friends, we asked GPT a question based on the results of a previous case study in the domain of Climate Change, which shows, on the bases of FBT analysis (Faber et al., 2012; Faber, 2015) that the term *captura y secuestro del carbono* in Spanish is a calque from the English term *carbon capture and sequestration*, and that it is indeed a result of a false friend, as the lexical unit *secuestro* in Spanish means kidnapping and not storage as the meaning of the lexical unit *sequestration* implies in English (Haddad, 2022). Accordingly, we asked GPT the following question: “If we compared the terms *carbon capture and sequestration* and *captura y secuestro del carbono* is the term in Spanish based on a false friend?”. The answer of GPT was: “The terms *carbon capture and sequestration* and *captura y secuestro del carbono* are not based on a false friend translation. They are accurate translations from English to Spanish in the context of carbon capture and sequestration technology. *Capture* in English and *captura* in Spanish, as well as *sequestration* in English and *secuestro* in Spanish, maintain their intended meanings in this technical context without causing confusion or misunderstanding”. We provided the following information to train GPT: “But sequestration in English means *storage* while *secuestro* in Spanish means *kidnapping*”. GPT answered: “You are correct, and I apologize for the oversight. You have pointed out an important distinction between the two terms.”

After receiving this answer from GPT, we asked it to provide two other terms which it considers are based on a similar false friend relation, and it provided the terms *actual pollution* and *Library of Green Technology*. The first term, *contaminación actual*, appeared in some papers such as Herrera et al. (2012), detected in the abstract in English and its translation into Spanish, and in other documents found in the Repository of The National Autonomous University of Nicaragua. However, it is not widely used or coined as a term. With respect to the second term, *librería de tecnología verde*, it is not in use either, which means that the suggestions of GPT did not match our query.

Then we asked GPT to provide a larger list of terms, and consequently, we obtained a list of 20

terms. we explained that we need terms which exist in real discourse and that the identified lexical units are indeed clear and total false friends. GPT explained that it was a very challenging task to obtain such terms in the domain of environment, for this reason, we asked GPT to provide examples in other domains, such as UN discourse. This query led to more satisfactory results as GPT retrieved the term: *treaty ratification* and its Spanish equivalent *ratificación de tratados* among other terms that are discussed in Section 4.

3.2 Second experiment

On further investigation on the matter of detecting false friends in English and Spanish, we conducted a larger experiment with the UN parallel Corpus. We used nearly 10,000 aligned English and Spanish sentences to ask ChatGPT whether it can find any false friends in these texts. We created an informative prompt which had 4 main parts containing the instruction to follow. The prompt consisted of, 1. Definition of false friends; 2. Task explanation; 3. An example; 4. Instruction on response.

We decided to have the above parts in our prompt by performing trial-and-error mechanism. Final prompt can be found in Appendix C.

We used *gpt-3.5-turbo-0125* as our base model. We have listed the yielded parameter values in Table 1.

Parameter	Value
Model	<i>gpt-3.5-turbo-0125</i>
Temperature	0.1
response_format	<i>json_object</i>

Table 1: Parameters of ChatGPT experiments

Once we run the experiment, we saved all the data and conducted a manual evaluation if ChatGPT has detected false friends.

4 Results and Discussion

4.1 Results of first experiment

To analyse the obtained results of the first query, we checked all the list of terms in Spanish, whether they are considered as terms at first place, and we verified whether they exist or not. We categorised each term provided by GPT. Terms considered as based on false friend relation were tagged as (1); not a false friend relation but contains terms which are erroneously coined in Spanish due to translation errors or inaccuracy (tagged as 2), and terms or

words which are not considered relevant for this study (tagged as 3). For example, the term *bosque actual* and its equivalent *actual forest* were tagged as (1) as they show a false friend relation. The term *granja de paneles solares* as equivalent to Solar Panel Farm were tagged as (2), as GPT suggests the use of the term *planta de paneles solares* instead, explaining that it would be more accurate. The suggested term *señal de éxito* and its equivalent *exit sign*, were tagged as 3 as the term does not exist in Spanish. The list of all terms and their classification can be seen in Appendix D.

Out of 20 suggested terms by GPT, we obtained 2 terms based on a total false friend relation; 6 terms which are based on partial false friends and may be relevant for terminological studies related to the adequate coinage of terms; and 12 irrelevant terms.

4.2 Results of second experiment

In the second experiment, we provided GPT with the parallel corpus of the UN available in the UNPC webpage. GPT rendered 269 sentences which GPT considered as containing a term based on false friend relation. After analysing the results, and comparing the meaning of the lexical units and its lexical construction in comparison to its English equivalents, we classified the results as: (1) terms based on false friends; (2) terms and lexical units which are relevant to terminological and translation studies; (3) erroneous results. Under the first category, the results showed 17 sentences in the corpus of the UN with presence of false friends, based on the lexical units: *paradas*, *sensible*, *preservativo*, *comprensivo* and *secuestro*, as equivalents to *parades*, *sensible*, *preservative*, *comprehensive* and *sequestration*, and forming terms such as: *paradas mmilitares* as equivalent to *military parades* (*paradas* in Spanish means bus stops); *sensibles a los desastres naturales* as equivalent to *sensible to natural disasters* (*sensible* in Spanish means *sensitive* in English); *preservativos de vacunas* as equivalent to *vaccine preservatives* (*preservativos* in Spanish means condoms and the proper word is *conservantes*); and *marco comprensivo* as equivalent to *comprehensive approach* (*comprensivo* en Spanish means *understanding*). Under the second category, GPT detected terms which are not accurately translated or are relevant to other terminological and translation studies, but are not based on a false friend relation from English into

Spanish, such as the terms *paraíso fiscal* in Spanish and its English equivalent tax haven, which shows that the translation in Spanish is more similar to the word *heaven* than *haven*; *seguridad de la biotecnología* and its English equivalent *biosafety*, taking into consideration that the literal translation of the Spanish term would be *safety of biotechnologies*, which means that it is a possible candidate term for a future terminological research to verify the dimension of its adequacy. The results obtained were 14. The rest of the list rendered by GPT, 240 sentences, showed to be erroneous false friend and irrelevant, such as name of countries, similar cognates or proper lexical units translated from English into Spanish.

5 Conclusion

This case study uses ChatGPT to detect false friends in translation and terminological research, addressing a gap in computational linguistics and translation studies. It focuses on specialised domains such as the environmental discourse and UN terminology and provides insights into the intricacies of professional translation. The paper also sheds light on the challenges of translation in specialised fields, and demonstrates the potential use of LLMs for detecting false friends and terms based on false friends.

In spite of the scarcity of valid results obtained through the two experiments, and taking into consideration that it is difficult to detect terms based on false friends in discourse, we consider those results as useful for terminological research, and are considered evidence on the presence of transfer problems among languages in specialised discourse.

We also agree with Lewis (2020) who highlights the need for more research to show the ubiquitous presence of false friends in interdisciplinary Terminology and calls for a more systematic approach to identify them and avoid them. At European level, Lewis (2020) suggests creating a multilingual European database of false friends with corpus-based sentences to avoid falling in such errors across languages. We believe that the approach used in this study can be used as a starting point to obtain insightful preliminary results. In our view, GPT is one of the tools that can be used to detect a list terms and to create a database with possible false friends in any specialised domains; however all the results provided by GPT must be analysed manually to verify their veracity and suitability, as GPT

also provides erroneous results.

In a nutshell, the following conclusions can be also highlighted:

- GPT does not always provide a correct answer, and not all the provided terms fulfil the pre-established criteria. Prompting is very important, as well as the manual analysis of each term.
- GPT is useful to start a hypothesis and to obtain preliminary results to build upon them in more detailed research in the field of translation and terminology.
- GPT is a useful aid during the process of coining new terms, as once prompted, it provides insightful information about the adequacy of use, whether the suggested term is a false friend, as well as providing suggestions for a more accurate term.

Finally, we recommend a supervised and monitored use of GPT in terminology-related tasks, for the coinage of new terms and for the creation of new models that aim at creating new terms for low resources languages.

In future research, we aim at applying this methodology on other pairs of languages which are less similar, such as English and Arabic, as well as carrying out more experiments with other LLMs.

Acknowledgments

This research was carried out as part of the projects PID2020-118369GB-I00 and A-HUM-600-UGR20, funded by the Spanish Ministry of Science and Innovation and the Regional Government of Andalusia.

References

- Anwar AH Al-Athwary. 2021. False friends and lexical borrowing: A linguistic analysis of false friends between english and arabic. *Journal of Language and Linguistic Studies*, 17(S1):368–383.
- Khuyagbaatar Batsuren, Gábor Bella, and Fausto Giunchiglia. 2022. A large and evolving cognate database. *Language Resources and Evaluation*, pages 1–25.
- Cambridge University Press. Accessed 2024. Cambridge dictionary. <https://dictionary.cambridge.org/>.

- Santiago Castro, Jairo Bonanata, and Aiala Rosá. 2018. A high coverage method for automatic false friends detection for spanish and portuguese. In *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*, pages 29–36.
- Loïc Depecker. 2015. How to build terminology science. 2015). *Handbook of Terminology*, 1:34–44.
- Pedro J Chamizo Dominguez and Brigitte Nerlich. 2002. False friends: their origin and semantics in some selected languages. *Journal of pragmatics*, 34(12):1833–1849.
- Pamela Faber. 2010. English as an academic lingua franca. *Revista Alicantina de Estudios Ingleses*, 23:19–32.
- Pamela Faber. 2015. Frames as a framework for terminology. *Handbook of terminology*, 1(14):14–33.
- Pamela B Faber et al. 2012. *A cognitive linguistics view of terminology and specialized language*. De Gruyter Mouton Berlin, Boston.
- Pablo Gamallo and Marcos Garcia. 2012. Extraction of bilingual cognates from wikipedia. In *International Conference on Computational Processing of the Portuguese Language*, pages 63–72. Springer.
- Amal Haddad Haddad. 2022. False friend in scientific and technical language: a corpus-based study in climate change discourse. In *Investigaciones de nuevo cuño en la academia*, pages 297–312. Thomson Reuters Aranzadi.
- Alan Humberto Herrera, Doctor Francisco Picado Pavón, Maestra Thelma Salvatierra, and Maestro Emilio Peña. 2012. Universidad nacional autónoma de nicaragua centro para la investigación en recursos acuáticos de nicaragua.
- Diptesh Kanojia, Kevin Patel, Pushpak Bhattacharyya, Malhar Kulkarni, and Gholamreza Haffari. 2021. Utilizing wordnets for cognate detection among indian languages. *arXiv preprint arXiv:2112.15124*.
- Adam Kilgarriff, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. The sketch engine: ten years on. *Lexicography*, 1(1):7–36.
- Kristian Lewis. 2020. False friends in terminology: Croatian lost in translation. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje*, 46(1):141–164.
- Khanif Makayev, Svetlana Takhtarova, Mery Gulkyan, Ilvira Kuznetsova, and Tatiana Kazymova. 2020. The urgency of considering “false friends of a translator” for construction specialists. In *IOP Conference Series: Materials Science and Engineering*, volume 890, page 012205. IOP Publishing.
- Ruslan Mitkov, Viktor Pekar, Dimitar Blagoev, and Andrea Mulloni. 2007. Methods for extracting and classifying pairs of cognates and false friends. *Machine translation*, 21:29–53.
- Alessio Palmero Aprosio, Stefano Menini, and Sara Tonelli. 2020. Adaptive complex word identification through false friend detection. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, pages 192–200.
- MGd Quesada, Fuertes Olivera, and Silvia Montero. 2002. Propuesta de estructura definicional terminográfica en ontoterm®. *Terminology*, 8(1):57–90.
- Real Academia Española. Accessed 2024. *Diccionario de la Real Academia Española*. Real Academia Española, Madrid.
- Ayla Rigouts Terry, Véronique Hoste, and Els Lefever. 2020. In no uncertain terms: a dataset for monolingual and multilingual automatic term extraction from comparable corpora. *Language Resources and Evaluation*, 54(2):385–418.
- Frieda Steurs and Hendrik Kockaert. 2011. Language planning and domain dynamics: challenges in term creation. In *Annual meeting of the Societas Linguistica Europaea (SLE): book of abstracts*, pages 275–276. Universidad de la Rioja.
- Grzegorz Szpila. 2006. False friends in dictionaries. bilingual false cognates lexicography in poland. *International Journal of Lexicography*, 19(1):73–97.
- Valentina Trishina. 2022. False cognates by spanish and russian speaking efl students from the perspective of cross-linguistic similarity in second language acquisition.
- Nihal Yetkin. 2011. Partial false friends in english-turkish translations: diplomatic texts. *Hacettepe Üniversitesi Edebiyat Fakültesi Dergisi*, 28(1).

Appendixes

A List of false friends between English and Spanish provided by ChatGPT

- Actual (Spanish) - Current, not "actual"
- Asistir (Spanish) - Attend, not "Assist"
- Carpeta (Spanish) - Folder, not "Carpet"
- Embarazada (Spanish) - Pregnant, not "Embarrassed"
- Ropa (Spanish) - Clothing, not "Rope"
- Sopa (Spanish) - Soup, not "Soap"
- Largo (Spanish) - Long, not "Large"
- Realizar (Spanish) - Accomplish, not "Realize"
- Éxito (Spanish) - Success, not "Exit"
- Librería (Spanish) - Bookstore, not "Library"
- Aparentemente (Spanish) - Apparently, not "Apparently"
- Constipado (Spanish) - Cold, not "Constipated"
- Sensible (Spanish) - Sensitive, not "Sensible"
- Pretender (Spanish) - Intend, not "Pretend"
- Fábrica (Spanish) - Factory, not "Fabric"
- Actualmente (Spanish) - Currently, not "Actually"
- Tuna (Spanish) - Cactus fruit, not "Tuna fish"
- Atender (Spanish) - Assist, not "Attend"
- Constipación (Spanish) - Cold, not "Constipation"
- Molestar (Spanish) - Bother, not "Molest"
- Colegio (Spanish) - School, not "College"
- Pretensión (Spanish) - Intention, not "Pretension"
- Mayor (Spanish) - Older, not "Mayor"
- Sano (Spanish) - Healthy, not "Sane"
- Ropa (Spanish) - Clothing, not "Roof"
- Abogado (Spanish) - Lawyer, not "Advocate"
- Molestar (Spanish) - Bother, not "Molest"
- Tren (Spanish) - Train, not "Trend"
- Asegurar (Spanish) - Ensure, not "Assure"
- Largo (Spanish) - Long, not "Largo"
- Mismo (Spanish) - Same, not "Mismo"
- Parentes (Spanish) - Relatives, not "Parents"
- Mayoría (Spanish) - Majority, not "Mayor"
- Suceder (Spanish) - Happen, not "Succeed"
- Actuar (Spanish) - Perform, not "Act"
- Apellido (Spanish) - Last name, not "Appellation"
- Ropa (Spanish) - Clothes, not "Rope"
- Ropa (Spanish) - Wardrobe, not "Rope"
- Pie (Spanish) - Foot, not "Pie"
- Falta (Spanish) - Lack, not "Fault"
- Gato (Spanish) - Cat, not "Gate"
- Mente (Spanish) - Mind, not "Mint"
- Cuento (Spanish) - Story, not "Count"
- Sensible (Spanish) - Sensitive, not "Sensible"
- Resumir (Spanish) - Summarize, not "Resume"
- Tarjeta (Spanish) - Card, not "Tarot"
- Exquisito (Spanish) - Delicious, not "Exquisite"
- Recordar (Spanish) - Remember, not "Record"
- Resto (Spanish) - Rest, not "Restaurant"
- Suceso (Spanish) - Event, not "Success"
- El embarazo (Spanish) - Pregnancy, not "Embarrassment"
- La carpeta (Spanish) - Folder, not "Carpet"
- La fábrica (Spanish) - Factory, not "Fabric"
- El éxito (Spanish) - Success, not "Exit"

- El colegio (Spanish) - School, not "College"
 - El abogado (Spanish) - Lawyer, not "Advocate"
 - El tren (Spanish) - Train, not "Trend"
 - La mayoría (Spanish) - Majority, not "Mayor"
 - El gato (Spanish) - Cat, not "Gate"
 - El acento (Spanish) - Accent, not "Asset"
 - La actualidad (Spanish) - Current events, not "Actuality"
 - El apodo (Spanish) - Nickname, not "Apology"
 - El edificio (Spanish) - Building, not "Editing"
 - El alquiler (Spanish) - Rent, not "Alkali"
 - La arena (Spanish) - Sand, not "Arena (sports venue)"
 - El barrio (Spanish) - Neighborhood, not "Bar (place to drink)"
 - El desayuno (Spanish) - Breakfast, not "Desert"
 - La lengua (Spanish) - Tongue, not "Language"
 - El libro (Spanish) - Book, not "Library"
 - El ruido (Spanish) - Noise, not "Rude"
 - El cuadro (Spanish) - Painting or picture, not "Quad"
 - El codo (Spanish) - Elbow, not "Code"
 - La carta (Spanish) - Letter (correspondence), not "Card"
 - La fecha (Spanish) - Date, not "Feta (cheese)"
 - La red (Spanish) - Network, not "Red (color)"
 - El rincón (Spanish) - Corner or nook, not "Rinse"
 - El suceso (Spanish) - Event, not "Success"
 - El traje (Spanish) - Suit, not "Trace"
 - El vaso (Spanish) - Glass, not "Vase"
 - El alquiler (Spanish) - Rent, not "Alkali"
 - La arena (Spanish) - Sand, not "Arena (sports venue)"
 - El barrio (Spanish) - Neighborhood, not "Bar (place to drink)"
 - El desayuno (Spanish) - Breakfast, not "Desert"
 - La lengua (Spanish) - Tongue, not "Language"
 - El libro (Spanish) - Book, not "Library"
 - El ruido (Spanish) - Noise, not "Rude"
 - El cuadro (Spanish) - Painting or picture, not "Quad"
 - El codo (Spanish) - Elbow, not "Code"
 - La carta (Spanish) - Letter (correspondence), not "Card"
- B List of terms in the domain of climate change containing a false friend from the previous list**
- "Carbon footprint reduction" - Contains "footprint," which in Spanish is "huella."
 - "Environmental impact assessment" - Contains "impact," which in Spanish is "impacto."
 - "Climate policy development" - Contains "policy," which in Spanish is "política."
 - "Renewable energy sources" - Contains "sources," which in Spanish is "fuentes."
 - "Climate data analysis" - Contains "data," which in Spanish is "datos."
 - "Emissions reduction goals" - Contains "emissions," which in Spanish is "emisiones."
 - "Sustainable development practices" - Contains "development," which in Spanish is "desarrollo."
 - "Biodiversity conservation efforts" - Contains "conservation," which in Spanish is "conservación."
 - "Climate adaptation strategies" - Contains "strategies," which in Spanish is "estrategias."

- "Natural disaster resilience" - Contains "disaster," which in Spanish is "desastre."
- "Climate change adaptation" - Incorporates "adaptation," which in Spanish is "adaptación."
- "Eco-friendly practices" - Includes "eco-friendly," with "eco" being a similar term in both languages.
- "Sustainable development goals" - Contains "sustainable," which in Spanish is "sostenible."
- "Carbon emissions reduction efforts" - Incorporates "reduction," similar to "reducción" in Spanish.
- "Environmental impact assessment" - Contains "impact," which in Spanish is "impacto."
- "Climate action plan implementation" - Incorporates "implementation," a similar term in both languages.
- "Renewable energy transition strategy" - Includes "transition," which in Spanish is "transición."
- "Climate resilience assessment" - Contains "resilience," similar to "resiliencia" in Spanish.
- "Biodiversity conservation initiatives" - Incorporates "conservation," which in Spanish is "conservación."
- "Sustainable energy solutions" - Contains "sustainable," which in Spanish is "sostenible."

C ChatGPT prompts and parameters

C.1 Prompt

A "false friend" is a linguistic term referring to words in different languages that look or sound similar but have different meanings. These similarities can often lead to confusion or misunderstandings for language learners or even native speakers who encounter them in unfamiliar contexts. For example, the English word "embarrassed" and the Spanish word "embarazada" sound similar, but "embarazada" means "pregnant" in Spanish, not "embarrassed." So, "embarazada" is a false friend for English speakers trying to communicate in Spanish.

You will be given sentences as pairs, the English sentence and the Spanish translation of it. Your task is to detect false-friends based terms in the given Spanish sentence and then return a JSON object saying if there is a false-friend term Yes/No and the false-friend Spanish term if you find any. Make sure you do not detect any country names, state names or any names.

For example : "captura y secuestro del carbono" is a term based on false friend translated from "carbon capture and sequestration", because the lexical unit "secuestro" is a false friend of "sequestration"

Example JSON object : {false_friend_term_found : 'yes', false_friend_term = ['term1', 'term2']}

Do not provide any other explanation. Just return JSON object with the results. English : {sentence goes here}

D List of all terms provided by GPT and their classification

- The term *bosque actual*, appears in the UN parallel corpus as the translation of existing forest. This use shows that it is the result of a false friend translation error. (1)
- The term *librería de estudios ecológicos* could not be found in online resources. (3)
- The term *señal de éxito de emergencia* could not be found in online resources. (3)
- The term *reserva natural* is correctly used in Spanish. (3)
- The suggested term *pesticidas de arma* does not exist in Spanish. (3)
- The term *granja de paneles solares* is a metaphor-based term, and it is a direct calque from English, not a result of a false friend; this term is used in academic research. Google Scholar search shows different articles using this term. However, GPT suggests the use of the term *planta de paneles solares* instead, explaining that it would be more accurate. (2)
- The term *gestión de recursos hídricos* is widely used in Spanish. The Real Academia Española dictionary (DRAE (Real Academia Española, Accessed 2024) defines *recurso* in one of its meanings as "Conjunto de elementos

disponibles para resolver una necesidad o llevar a cabo una empresa” [a set of available elements needed to cover a necessity or to launch a company], giving as example “Recursos naturales, hidráulicos, forestales, económicos, humanos”. Those examples coincide with the use of the term in English, hence, we do not consider this term as based on false friend relation. (3)

- The term *Departamento de estudios ambientales* is widely used in Spanish and refers to the same concept. We do not consider this pair as a false friend relation. (3)
- In the case of the term *transición de energía limpia*, which is widely used in scientific articles, it is not accurate according to GPT due to the preposition *de* which in its opinion, should be *hacia* in its place to achieve a more accurate meaning. In this case, we consider the false friend relation as valid and insightful due to the meaning attached to the preposition. (2)
- According to GPT, the use of the term *gestión de pescas sostenibles* is not accurate as *pesca* is the fishing activity, while *pesquería* is a more accurate term to be used instead of fishing, which refer to the “area of water where fish are caught so they can be sold” as defined in the Cambridge Online Dictionary (Cambridge University Press, Accessed 2024). In spite of understanding the reason of suggesting this more accurate term, the lexical units *pesca* and *pesquerías* are not false friends of fishing, however, this suggestion provided by GPT is considered insightful for terminological case-studies. (2)
- The term *control global de la contaminación* is suggested by GPT as the lexical unit *control* in Spanish may have more meanings in other contexts. For this reason, we do not consider this suggestion as valid as a false friend example, neither relevant for our study. (3)
- GPT considers the term *ratificación de tratados* as less false friend based term, as the lexical unit *ratificación* is a false friend of *ratification*. GPT explains that *treaty ratification* refers to the formal approval or acceptance of an international treaty by a country’s government. In Spanish, *ratification* can be a false

friend, as it is often translated as *ratificación*. However, *ratificación* primarily refers to the confirmation of an action or decision, which can lead to potential misunderstandings when discussing the formal acceptance of treaties. We checked the corresponding meaning in the Cambridge Dictionary and in the RAE Dictionary, and we have noticed that the answer provided by GPT was correct. (1)

- The term *conservación de especies en peligro de extinción* is not considered a false friend-based term as the meaning of *conservación* in Spanish according to the DRAE is equivalent to *conservation*, unlike what GPT suggests; however, according to GPT, *preservation* would be more accurate than *conservation* for the exact connotation of the meaning, and we consider this suggestion as relevant, as *perservar* is defined in DRAE as: “Proteger, resguardar anticipadamente a una persona, animal o cosa, de algún daño o peligro.” (Real Academia Española, Accessed 2024) (2)
- The term *infraestructura de energía limpia* as equivalent to Clean Energy Infrastructure is not steamed to be a false friend as *infraestructura* and *infraestructura* are not false friends. (3)
- The term *evaluación de impacto ambiental* is suggested by GPT as the lexical unit *impacto* has two meanings in Spanish. For this reason, we do not consider this term as based on a false friend, neither insightful for our case study. (3)
- The term *prácticas de agricultura sostenible* is suggested by GPT, as the lexical unit *sostenible* in Spanish may have more meanings in other contexts. For this reason, we do not consider this suggestion as valid as a false friend example, neither relevant for our study. (3)
- The term *metas de conservación ambiental* is not considered a false friend-based term as the meaning of *conservación* in Spanish according to the DRAE is equivalent to *conservation*, unlike what GPT suggests; however, according to GPT, *preservation*, would be more accurate than *conservation* for the exact connotation of the meaning, and we consider this

suggestion as relevant, as *perservar* is defined in RAE as: “Proteger, resguardar anticipadamente a una persona, animal o cosa, de algún daño o peligro.” (Real Academia Española, Accessed 2024). (2)

- The term *desarrollo de energía renovable* is not considered a false friend-based term, as *desarrollo* and *development* are not false friends and both involve the same meaning. (3)
- The term *conservación de la biodiversidad marina* is not considered a false friend-based term as the meaning of *conservación* in Spanish according to the RAE is equivalent to conservation, unlike what GPT suggests; however, according to GPT, *conservación* would be more accurate than *preservación* for the exact connotation of the meaning, and we consider this suggestion as relevant, as *perservar* is defined in DRAE as: “Proteger, resguardar anticipadamente a una persona, animal o cosa, de algún daño o peligro.” (Real Academia Española, Accessed 2024). In the case of the lexical unit *biodiversidad*, we do not agree with the suggestion of GPT as the meaning is equivalent in English and Spanish, and the Spanish term includes the connotation of variety of species according to the DRAE. (2)
- The term *evaluación de la sostenibilidad ambiental* is not considered a false friend-based term, as *sostenibilidad* and *sustainable* are not false friends and both involve the same meaning. (3)

E Terms in English and their Equivalents in Spanish provided by GPT and the explanation of GPT why they are based on false friend relation

English Term	Spanish Term	GPT Explanation
Actual Forest	Bosque actual	"Actual" in Spanish means "current" or "present," not "actual" as in "real." So, "actual forest" in English might be mistakenly translated to "bosque actual" in Spanish, which means "current forest" rather than "real forest."
Library of Ecological Studies	Librería de estudios ecológicos	As mentioned before, "library" in English might be mistranslated as "librería" in Spanish, which is commonly associated with a bookstore. So, "library of ecological studies" might be inaccurately translated as "librería de estudios ecológicos," while a more appropriate translation would be "biblioteca de estudios ecológicos."
Emergency Exit Sign	Señal de éxito de emergencia	"Exit" in English might be mistakenly translated to "éxito" in Spanish, which means "success." So, "emergency exit sign" could be mistranslated as "señal de éxito de emergencia," while it should be "señal de salida de emergencia."
Natural Reserve	Reserva natural	"Reserve" in English might be translated as "reserva" in Spanish, but it's essential to be aware that "reserva" in Spanish is often associated with booking or reserving something in advance. So, "natural reserve" might be inaccurately translated as "reserva natural," while "reserva natural" can also refer to a protected natural area.
Harmful Pesticides	Pesticidas de arma	"Harmful" in English might be translated as "dañino" in Spanish. However, "harm" can be mistranslated as "arma" in Spanish, which means "weapon." So, "harmful pesticides" could be mistakenly translated as "pesticidas de arma," while it should be "pesticidas dañinos."
Solar Panel Farm	Granja de paneles solares	"Farm" in English might be translated as "granja" in Spanish, which usually refers to a place for agriculture or livestock. So, "solar panel farm" could be inaccurately translated as "granja de paneles solares," while "planta de paneles solares" would be a more appropriate translation.
Water Resource Management	Gestión de recursos hídricos	"Resource" in English might be mistakenly translated as "recurso" in Spanish. However, "recurso" in Spanish can also mean a recourse or a means to solve a problem. So, "water resource management" might be mistranslated as "gestión de recursos hídricos," while it should be "gestión de la fuente hídrica."
Environmental Studies Department	Departamento de estudios ambientales	"Department" in English might be translated as "departamento" in Spanish. However, in this context, "departamento" may imply a residential unit. So, "environmental studies department" could be inaccurately translated as "departamento de estudios

		ambientales," while "facultad de estudios ambientales" or "departamento de estudios del medio ambiente" might be more appropriate.
Clean Energy Transition	Transición de energía limpia	"Transition" in English might be translated as "transición" in Spanish. However, "transición" can also be used in the sense of transitioning from one gender to another. So, "clean energy transition" might be mistakenly translated as "transición de energía limpia," while "transición hacia la energía limpia" would be a more suitable translation.
Sustainable Fisheries Management	Gestión de pescas sostenibles	"Fisheries" in English might be translated as "pescas" in Spanish. However, "pescas" can refer to fishing activities. So, "sustainable fisheries management" could be inaccurately translated as "gestión de pescas sostenibles," while "gestión sostenible de pesquerías" would be a more accurate translation.
Global Pollution Control	Control global de la contaminación	"Control" in English might be translated as "control" in Spanish. However, "control" in Spanish can sometimes refer to a remote control, such as for a television. So, "global pollution control" could be mistakenly translated as "control global de la contaminación," while "control y regulación de la contaminación a nivel mundial" would provide a more accurate translation.
Treaty ratification	Ratificación de tratados	Treaty ratification & ratificación de tratados & "treaty ratification" refers to the formal approval or acceptance of an international treaty by a country's government. In Spanish, "ratification" can be a false friend, as it is often translated as "ratificación." However, "ratificación" primarily refers to the confirmation of an action or decision, which can lead to potential misunderstandings when discussing the formal acceptance of treaties.
Endangered Species Conservation	Conservación de especies en peligro de extinción	"Conservation" in English might be translated as "conservación" in Spanish. However, "conservación" can sometimes be associated with preserving something in its current state, such as preserving food. So, "endangered species conservation" could be mistakenly translated as "conservación de especies en peligro de extinción," while "protección de especies en peligro de extinción" would be a more accurate translation.
Clean Energy Infrastructure	Infraestructura de energía limpia	"Infrastructure" in English might be translated as "infraestructura" in Spanish. However, "infraestructura" can sometimes be more broadly associated with physical structures and facilities. So, "clean energy infrastructure" could be inaccurately translated as "infraestructura de energía limpia," while "desarrollo de energías limpias" or "red de energía limpia" would provide a more accurate translation.

Environmental Impact Assessment	Evaluación de impacto ambiental	"Impact" in English might be translated as "impacto" in Spanish. However, "impacto" can also refer to a collision or crash. So, "environmental impact assessment" could be mistakenly translated as "evaluación de impacto ambiental," while "evaluación de los efectos ambientales" might be a more suitable translation.
Sustainable Agriculture Practices	Prácticas de agricultura sostenible	"Sustainable" in English might be translated as "sostenible" in Spanish. However, "sostenible" can sometimes be associated with maintaining or propping something up. So, "sustainable agriculture practices" could be inaccurately translated as "prácticas de agricultura sostenible," while "prácticas agrícolas sustentables" might be a more appropriate translation.
Environmental Conservation Goals	Metas de conservación ambiental	"Conservation" in English might be translated as "conservación" in Spanish. However, as mentioned before, "conservación" can sometimes imply preservation rather than broader conservation efforts. So, "environmental conservation goals" could be mistakenly translated as "metas de conservación ambiental," while "objetivos de conservación del medio ambiente" would be a more accurate translation.
Renewable Energy Development	Desarrollo de energía renovable	"Development" in English might be translated as "desarrollo" in Spanish. However, "desarrollo" can also refer to the growth or expansion of something. So, "renewable energy development" could be inaccurately translated as "desarrollo de energía renovable," while "fomento de energías renovables" might provide a more appropriate translation.
Marine Biodiversity Conservation	Conservación de la biodiversidad marina	"Biodiversity" in English might be translated as "biodiversidad" in Spanish. However, "biodiversidad" can sometimes be more narrowly associated with a wide variety of species without considering their variation. So, "marine biodiversity conservation" could be mistakenly translated as "conservación de la biodiversidad marina," while "preservación de la variabilidad marina" might be a more accurate translation.
Environmental Sustainability Assessment	Evaluación de la sostenibilidad ambiental	"Sustainability" in English might be translated as "sostenibilidad" in Spanish. However, "sostenibilidad" can sometimes be more closely linked to maintaining or propping something up. So, "environmental sustainability assessment" could be inaccurately translated as "evaluación de la sostenibilidad ambiental," while "evaluación de la viabilidad ambiental" might be a more suitable translation.