

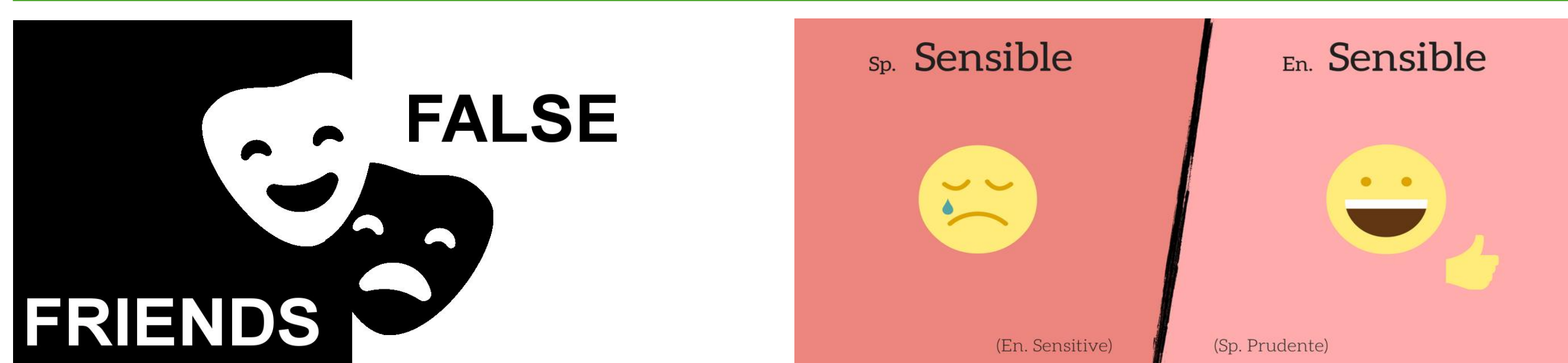
ChatGPT: Detection of Spanish Terms Based on False Friends

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.

Definition: Terms based on false friends

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.



Carbon Capture and **Sequestration**
Captura y **secuestro** de carbono

Experiment 1

ChatGPT: provide a list of terms based on false friend:

1. Defining a false friend and a term-based on false friend.
2. Give the example of previous study in a specialised domain: carbon capture and sequestration.
3. Ask GPT to produce similar examples.
4. It rendered the examples:
 - actual pollution = **polución actual**
 - Library of Green Technology = **Librería de tecnología verde**
5. We asked for a list of 20 terms based of false friends.



Amal Haddad Haddad
University of Granada
amalhaddad@ugr.es

Lexicon research group

Damith Premasiri
Lancaster University
d.dolamullage@lancaster.ac.uk

Experiment 2

GPT: automatic detection of terms based on false friends from the UN parallel corpus

Dataset: We used the dataset of 10,000 aligned English and Spanish sentences from the UN corpus (available on UNPC webpage).

Prompt: (based on trial-and-error mechanism)

1. Definition of false friends.
2. Task explanation.
3. An example.
4. Instruction on response.
5. Manual evaluation of results

Results of Experiment 1

Manual analysis of each term is carried out and terms are categorised.

Results:

- 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;
- 12 irrelevant terms.

Results of Experiment 2

GPT rendered 269 sentences which GPT considered as containing a term based on false friend relation.

Results:

- 17 terms based on false friends
 - vaccine preservatives → **preservativos de vacunas**
 - sensible to natural distares → **sensibles a los desastres naturales**
- 14 terms with transfer problems from English into Spanish.
- 240 term which are irrelevant for this study.

CONCLUSION

- Terms based on false friends are more frequent than we imagine in specialised discourse, but there are hardly any studies carried out.
- Taking into consideration the great difficulty to detect terms based on false friends, GPT is a useful tool to start with to compile a list to be studied.
- In spite of the scarcity of results, they are relevant and valuable for terminological studies.

→ Future research: using other LLMs to detect terms based on false friends from corpus.

Funding: 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.

