Deep Learning Methods for Extracting Metaphorical Names of Flowers and Plants

Métodos de aprendizaje profundo para la extracción de nombres metafóricos de flores y plantas

Amal Haddad Haddad,[1] **Damith Premasiri,**[2] **Tharindu Ranasinghe,**[3] **Ruslan Mitkov**[4]
[1]University of Granada, Spain
[2]University of Wolverhampton, UK
[3]Aston University, UK
[4]Lancaster University, UK
amalhaddad@ugr.es

Abstract: The domain of Botany is rich with metaphorical terms. Those terms play an important role in the description and identification of flowers and plants. However, the identification of such terms in discourse is an arduous task. This leads in some cases to committing errors during translation processes and lexicographic tasks. The process is even more challenging when it comes to machine translation, both in the cases of single-word terms and multi-word terms. One of the recent concerns of Natural Language Processing (NLP) applications and Machine Translation (MT) technologies is the automatic identification of metaphor-based words in discourse through Deep Learning (DL). In this study, we seek to fill this gap through the use of thirteen popular transformer based models, as well as ChatGPT, and we show that discriminative models perform better than GPT-3.5 model with our best performer reporting 92.2349% F1 score in metaphoric flower and plant names identification task.
Keywords: Deep Learning, Transformers, Automatic Extraction of Metaphor, Metaphor-based Terms.

Resumen: El dominio de la Botánica es rico en términos metaforicos. Estos términos tienen un papel importante en la descripción e identificación de flores y plantas. Sin embargo, la identificación de este tipo de términos en el discurso es una tarea difícil. Esto puede conducir a errores en los procesos de traducción y otras tareas lexicográficas. Este proceso es aún más difícil cuando se trata de traducción automática, tanto en el caso de las unidades monoléxicas, como en el caso de las unidades multiléxicas. Uno de los desafíos a los que se enfrentan las aplicaciones del Procesamiento del Lenguaje Natural y las tecnologías de Traducción Automática es la identificación de términos basados en metáfora a través de métodos de aprendizaje profundo. En este estudio, tenemos el objetivo de rellenar este vacío a través del uso de trece modelos populares basados en transformadores, además del ChatGPT. Asimismo, demostramos que los modelos discriminativos aportan mejores resultados que los modelos de GPT-3.5. El mejor resultado alcanzó una puntuación de 92,2349% F1 en las tareas de identificación de nombres metafóricos de flores y plantas.
Palabras clave: Aprendizaje profundo, Transformadores, Extracción automática de metáfora, Términos basados en metáfora.

1 Introduction

Metaphor is a pervasive phenomenon in human language (Lakoff and Johnson, 2008). It is defined as "mapping of conceptual structure from a source to a target domain" (Ruiz de Mendoza Ibáñez, 2017). Depending on the dimension of complexity of metaphor, authors distinguish two types of metaphors: image metaphors and conceptual metaphors (Lakoff and Johnson, 2008). Image metaphors compare one single image in one domain with another image belonging to another domain, such as the image metaphor "she is as good as gold". Conceptual metaphors are more complex at the con-

ceptual and cognitive levels, and they refer to the resemblance established between a whole set of experiences, such as the metaphor in "life is a journey", which implies a whole set of elements activated within the metaphoric target domain. Metaphor-based terms, or the so-called terminological metaphors, are common in specialised languages. Their use is abundant, as they help in the conceptualisation of phenomena and their description by establishing a resemblance between images and domains. They also help in understanding abstract phenomena in terms of more concrete notions and in modelling scientific thought (Urena Gomez-Moreno and Faber, 2010). However, the identification of metaphor-based terms in discourse is an arduous task. This leads in some cases to committing errors during translation processes and lexicographic tasks. The process is even more challenging when it comes to machine translation, both in the cases of single-word terms and multi-word terms, which are represented by Multiword Expressions (MWEs). The main common error while carrying out the translation processes is that the metaphorical lexical items forming part of a term would be transferred literally into other languages without taking into consideration its metaphoric and cultural dimension or without taking into account that they form part of an MWE.

Previous studies focused on the extraction of metaphorical terms from discourse, such as Mu, Yannakoudakis, and Shutova (2019) and Razali et al. (2022); however, to the best of our knowledge, there are no programs that could automatically retrieve those terms both as single-word terms and MWEs in specialised languages. This study seeks to fill in this gap and proposes a novel method based on transformer models (Premasiri et al., 2022; Premasiri and Ranasinghe, 2022); (Ranasinghe et al., 2021) for automatic extraction of metaphor-based terms from the specialised domain of Botany and concerning the names of flowers and plants in English and Spanish. The main contributions of this study are:

1. We empirically evaluate thirteen discriminative transformer models and one generative transformer model (Chat-GPT) for the tasks of metaphoric flower and plant names identification on English and Spanish datasets.

2. We show that discriminative models perform better in the metaphoric flower and plant names identification task.

3. We release new annotated datasets for metaphoric names identification in English and Spanish.

4. We make our code freely available for further research[1].

This paper is organised as follows: in Section 2 we present previous related work. In Section 3 we describe the dataset used and its annotation process. In Section 4 we detail the experimental set-up and methodology, while in Section 5 we report our experiment's results and evaluation. Finally, we summarise the main conclusions and propose future work in Section 6.

2   Related work

The study of metaphor-based terms in discourse has been a subject of study in the last few decades. One of the main concerns in this field is the detection of metaphor-based words in discourse. With this aim, the Pragglejaz Group suggested a method for the manual identification of metaphor, called Metaphor Identification Procedure (MIP) (Group, 2007). This method has been used extensively (Nacey et al., 2019). Studies like Turney et al. (2011), Jang et al. (2015) and Coll-Florit and Climent (2019) have a similar approach. Other projects such as the VU Amsterdam Metaphor Corpus (Leong et al., 2020) offer a manually annotated corpus for all metaphorical language use. Moreover, studies like Yaneva (2016), show how the use of metaphor and figurative language in discourse is of utmost difficulty for people with Autism Spectrum Disorder (ASD); hence, studies like Yaneva (2016) and Štajner et al. (2017) endeavour to identify and disambiguate complex sentences which contain metaphor and metonymy among other features through the application of Complex Word Identification modules. The above studies were partially inspired by the FIRST Project[2] (Orăsan, Evans, and Mitkov, 2018) and the development of the Open Book tool which helps people with ASD.

Concurrently, one of the recent concerns of Natural Language Processing (NLP) appli-

---

[1]https://bit.ly/3pYAYXK
[2]http://www.iwebtech.co.uk/project-first/

cations and Machine Translation (MT) technologies is the automatic identification of metaphor-based words in discourse through Deep Learning Methods (DLM). For example, Mu, Yannakoudakis, and Shutova (2019) suggest working with large corpora and training *simple gradient boosting classifiers* on representations of an utterance and its surrounding discourse learned with a variety of document embedding methods". Su et al. (2020) focus on token-level metaphor detection paradigm and propose using an end-to-end deep metaphor detection model. Authors like Razali et al. (2022) use machine learning to automatically detect metaphor instances in short texts by implementing Support Vector Machine algorithms, while other authors like Gutierrez et al. (2016) propose modelling metaphor explicitly within compositional distributional semantic models to improve the resulting vector representations. Those authors classify the already used methods in the following categories: clustering; topic modelling; topical structure and imageability analysis; semantic similarity graphs and feature-based classifiers (Gutierrez et al., 2016). Recent approaches are more centred on using dense embedding methods (Vitez et al., 2022).

On the other hand, the study of metaphor-based terms in specialised discourse has been subject to scientific and cognitive studies. The automatic identification of metaphor-based terms is considered a substantial challenge. Some studies highlight the importance of automatic extraction of terms in specialised discourse (Rodríguez Penagos and others, 2005) while other studies, such as Urena Gomez-Moreno and Faber (2010), propose a semi-automatic method for term retrieval in the domain of Marine Biology. However, to the best of our knowledge, there have been no previous studies or methodologies which cover the automatic extraction of those terms from scientific discourse in other domains and no previous studies were carried out in the domain of Botany.

3 Data

Specialised discourse is rich in metaphor-based terms; Botany is no exception. The semantic motivations for plant names are usually influenced by the appearance of the plant, the place of its occurrence, the properties of the plant, its usage, as well as other motivations typical of a specific genus of species (Dębowiak and Waniakowa, 2019). Many studies have shown that metaphor is one of the most frequent techniques to coin flowers and plants names (Rastall, 1996); (Nissan, 2014); (Dębowiak and Waniakowa, 2019). This metaphoric use may give clues to cultural references related to legends and beliefs associated with plants in general, like their healing properties and supposed magical powers (Dębowiak and Waniakowa, 2019). At the same time, this shows that this metaphorical use may vary among languages and cultures. From another perspective, studies like Goodman (1963) highlight the importance of flower names based on metaphor for the study of colour and its comparison among languages. For this reason, we consider the study of metaphor-based terms in this domain relevant as a case-study.

The dataset we use to extract metaphor-based terms in English is the Encyclopaedia of Flowers and Plants, published by the American Horticultural Society (Brickell, 2012). We selected this edition as it is available in a digitalised format in the online library of the Internet Archive. This Encyclopaedia consists of 522,707 words. It contains a dictionary of names of flowers from around the world, with approximately 8000 terms referring to both scientific and common names and their origins, as well as 4000 images. It is divided into the following sections: firstly it has an introduction about how to use the book, plant names and origins and relevant information on how to create a garden and how to select plants. This introductory part shows that it is aimed at both professionals and laypersons. Secondly, it has a plant catalogue, subdivided into categories such as trees, shrubs, roses, climbers and wall shrubs, perennials, annuals, biennials and bedding, rock plants, bulbs, water and bog plants as well as tender and exotic plants. All those subsections contain rich contexts on each term, concerning the origin, uses, habitat, size, etc. Finally, the Encyclopaedia offers a dictionary section with an index of common names and glossary of terms. We benefited from this last section to extract and annotate terms. The advantage of using this Encyclopaedia is that it includes a wide range of varieties of flowers and plants from all around the world. For this reason, the obtained results may be useful to be ap-

plied in different contexts and in multidisciplinary studies.

The data was pre-processed by annotating the proper names and their metaphorical condition. The MIP criteria for metaphor identification (Group, 2007) was adapted to annotate the terms, considering a term as metaphor-based when one or more of the lexical units forming it or its etymology give evidence that they belong to different domains, based on its meaning in the dictionary. The annotated names represent both image metaphors and conceptual metaphors. An example of image metaphors, is the one-word name of the flower *Edelwiess* which is a combination between the two lexical units *edel* which means noble and *weiss*, which means white in German. This name represents an image metaphor where the flower is called as so as it symbolises purity. The scientific name of this flower is *Leontopodium Alpinum*, an MWE with Greek origin and etymology. It is also an image metaphor, as the lexical unit *Leontopodium* means lion's foot (Dweck, 2004), the resemblance is established between the for of the petals of the flowers and the aspect of the foot of a lion. Another example are the flowers *Sunburst* and *Moonlight*. The name of the flower *Sunburst* shows the resemblance between the colours of the flower and the colours of the sun, while the flower called *Moonlight*, alludes to the resemblance between the flower and the light of the moon. Other metaphor based-names represent a conceptual image, such as the MWE flower name *forget-me-not* which refer to the association between the heart-shaped blue flowers that reminds the person of his or her beloved one; or the one-word name of the flower *cascade* which associate the aspect of a flower with the whole process of the water falling in a real cascade.

Apart from the Encyclopaedia of Plants and Flowers, we also compiled a corpus of other resources related to Botany in English. It consists of 437,663 words. Some of the texts are monographs, others are journal articles, and some texts are retrieved from other online resources. The full list of references used to compile the English corpus are listed in Appendix 1. With respect to the Spanish dataset, we have annotated a list of flowers and plants names provided in selected monographs and glossaries following the same criteria as in the case of the English

terms. Above all, we used books and articles in the domain of Botany and botanical glossaries, such as the glossaries provided in *Los Áraboles en España* (de Lorenzo Cáceres, 1999), *Biología de la Conservación de Plantas en Sierra Nevada* (Peñas and Lorite, 2019) and the glossary of scientific names of plants and their vernacular names provided by the Entomological Museum in Leon in the Bio-Nica webpage[3]. The list obtained from this source consists of more than 5000 scientific and vernacular names of flowers and plants. As for the book *Los Áraboles en España*, it consists of almost 155,000 words with more than 600 terms in the section of Glossary. The book describes the details of each plant, its family names, its vernacular names and synonyms, its origin, etymology, description and cultivation information. It also provides illustrative images of each plant. The book *Biología de la Conservación de Plantas en Sierra Nevada* was also valuable as some of its chapters contained lists of scientific names of endemic flowers from Sierra Nevada and its common names too. In order to enhance the datasets, we also added more specialised, semi-specialised and informative texts in the domain of botany to obtain more rich contexts. It consists of 460,258 words. The full list of the sources used to compile the Spanish corpus are listed in Appendix.

With this paper, we release datasets of English and Spanish flower and plant names with their annotations metaphoric or not metaphoric. The English dataset consists of 6330 total plant and flower names as a combination of 1869 metaphorical names and 4461 non-metaphorical names. The Spanish dataset consists of 875 metaphoric names and 4,988 non-metaphoric names out of 5863 total.

Data Preparation Since we model the metaphoric name identification task as a token level classification task, we used IOB format tagging for our corpus. IOB format is widely used in token level classification tasks (Tjong Kim Sang and De Meulder, 2003) where B - Beginning, I - Inside and O - outside of a metaphoric flower or plant name; Table 1 shows an example IOB annotation. After tagging the sentences from the corpus, we identified that there were a very high number of sentences which do not have a single

---

[3]http://www.bio-nica.info/home/index.html

metaphoric name. In other words, the majority of the sentences only had 'O' as the tag for all their words. Since this has a negative impact on the model training process, we decided to balance the dataset by removing some sentences. Then we shuffled all the sentences and divided the training and test sets. Finally, we had 2020 total sentences divided 1500 and 520 in English training and test set respectively. For Spanish, we used only 250 sentences as the dataset.

Test sets were the same for discriminative and generative experiments. The only thing is that in the generative approach, we did not use the training set, since we cannot train ChatGPT.

4 Methodology

Discriminative Models Transformers (Vaswani et al., 2017) have been a major breakthrough in Deep Learning research, since they provide a robust mechanism based on attention for the neural networks to flow information without recurrence and convolution. This architecture has produced state-of-the-art results in many NLP applications. With the introduction of BERT (Devlin et al., 2019), which employs the transformers architecture, the pre-trained large language models have played an important role in pushing the boundaries of all NLP tasks such as text classification (Ranasinghe, Zampieri, and Hettiarachchi, 2019) , (Uyangodage, Ranasinghe, and Hettiarachchi, 2021), question answering (Premasiri et al., 2022), text similarity (Mitkov et al., 2023) etc. and achieving new state-of-the-art. With this motivation, we use transformers as our primary experimental setup and evaluate multiple pre-trained language models. These models follow similar architectures to BERT (Devlin et al., 2019) while they are pre-trained on different corpora and different objectives. Figure 1 (Ranasinghe and Zampieri, 2021) shows the transformer architecture we used where we input sentences which contain metaphoric flower and plant names, then we obtain BIO tags from the output layer by adding a softmax layer on top of the last hidden state of the deep network to classify each token into one of I,O,B tags. We used several popular transformers based pre-trained language models.

For the experiments on English dataset, we used the cased and uncased variants of BERT base and BERT large versions. In order to establish the capabilities of multilingual models, we experimented with the multilingual-bert (Devlin et al., 2019) model with its cased and uncased variants and xlm-roberta-base (Conneau et al., 2020) model and xlm-roberta-large (Conneau et al., 2020) version. We further experimented with google/electra-base-discriminator (Clark et al., 2020) model which is different from BERT architecture. Finally, within these discriminative models we evaluate allenai/scibert_scivocab_cased (Beltagy, Lo, and Cohan, 2019) and allenai/scibert_scivocab_uncased (Beltagy, Lo, and Cohan, 2019) variants which are specifically pre-trained on scientific corpora. We assume that flower and plant names could appear in those corpora such that the model can leverage the learning to produce better results.

Since Spanish is low in resources on metaphoric flower and plants names corpora, we experimented zero-shot learning for Spanish on English data. We specifically used the multilingual-bert (Devlin et al., 2019) and xlm-roberta (Conneau et al., 2020) for our experimental setting as these models provide multilingual capabilities.

All the models were trained for three epochs, learning rate 4e-5 with 32 training batch size and for the hardware we used a GeForce RTX 3090 GPU.
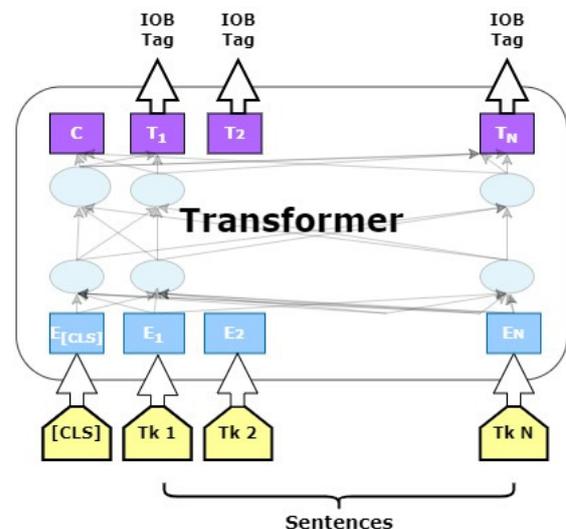

Figure 1: Transformers architecture for token level classification.

calliandra	haematocephala	(Red	powder	puff)	is	an	evergreen,	spreading	shrub
O	O	B	I	I	O	O	O	O	O

Table 1: BIO annotation example.

Model	Precision	Recall	F1
bert-base-uncased	92.8204	89.4824	91.0784
bert-base-cased	93.4157	90.8295	92.0801
bert-large-uncased	92.8424	90.6789	91.7219
bert-large-cased	93.4157	90.8295	92.0801
bert-base-multilingual-uncased	91.7655	89.6286	90.6648
bert-base-multilingual-cased	93.3662	91.1718	**92.2349**
xlm-roberta-base	90.1220	89.6020	89.8560
xlm-roberta-large	90.8455	89.4348	90.1220
xlnet-base-cased	89.8189	90.8769	90.3402
roberta-base	91.9779	89.8922	90.9025
google/electra-base-discriminator	92.0412	91.1617	91.5898
allenai/scibert_scivocab_uncased	91.7084	90.3453	91.0071
allenai/scibert_scivocab_cased	92.3408	90.6466	91.4750
ChatGPT	62.1516	45.1943	48.1392

Table 2: Resutls for English metaphoric flower and plant names identification; the Model column represents the model we experimented, the Precision column shows the macro precision, the Recall column shows macro recall and the F1 column shows macro F1 value for the results.

Generative Models While all above methods rely on the discriminative approach, which tries to identify boundaries in the data space, generative models attempt to model the placement of the data throughout the space. This approach attracted huge attention in the research community with the release of ChatGPT[4] by openAI[5]. The research on Generative Pre-trained Transformer (GPT) (Radford et al., 2018) models have produced multiple versions of it including GPT-3, GPT-3.5 and GPT-4. The free version of ChatGPT only supports GPT-3.5 for the time being and all our experiments are based on ChatGPT free version. According to OpenAI, the most cost capable and cost effective models out their models is gpt-3.5-turbo, which we used to our experiments.

Since ChatGPT is a generalised conversational application, it does not essentially provide IOB tags as outputs. After experimenting with different prompts to retrieve IOB tags from ChatGPT, we decided it would be easier to retrieve the metaphoric flower or plant name in the sentence from the API[6] and *No* otherwise. Prompt we used: *Is there a metaphoric flower name or metaphoric plant name included in the following sentence, say yes or no, if yes what is the metaphoric flower or metaphoric plant names in the sentence separately : {sentence goes here}.* The outputs of ChatGPT are not uniform, and we had to post process the outputs using regular expressions to re-generate the IOB tags for evaluation.

Since this is a token classification task, we use macro averaged Precision, Recall and F1 score as our evaluation metrics.

$$Precision = TP/(TP + FP) \qquad (1)$$

$$Recall = TP/(TP + FN) \qquad (2)$$

$$\text{F1} = 2 * (\text{Precision} * \text{Recall})/(\text{Precision} + \text{Recall}) \qquad (3)$$

5 Results and Discussion

5.1 English

The results in table 2 show the competitive performance of transformer models, in the flower and plant names classification task. Despite the fact that most of the transformer models we experimented with are not specifically pre-trained on botanic corpora, almost all discriminative models were able to produce more than 90% F1 score in the task. Interestingly, the multilingual bert model could

---

[4]https://chat.openai.com/

[5]https://openai.com/

[6]https://bit.ly/3OLCWFn

Model	P	R	F1
bert-base-multilingual-uncased	59.2957	40.3103	43.0472
bert-base-multilingual-cased	54.0904	52.1401	**52.8657**
xlm-roberta-base	67.4035	36.5622	37.4988
xlm-roberta-large	64.1040	47.4813	51.8174
ChatGPT	63.1887	46.6820	51.4120

Table 3: Results on metaphoric flower and plant names identification in Spanish; P - The macro averaged precision, R - The macro averaged Recall, F1 - The macro averaged F1 score.

surpass the other models and mark the top results at 92.2349% F1 score.

Another noteworthy observation in our study was that cased models outperformed all the respective uncased models. Even though the xlm-roberta-base was the least performer in discriminative models, the performance gap to the best performer is only 2.3789% which shows the competitiveness of the transformers in token level classification tasks.

Even though scibert models are specifically trained on scientific corpus, these models were not able to outperform the bert multilingual model, which shows that the general knowledge could play a significant role in metaphoric identification task.

While ChatGPT seems very good at handling general text, it does not perform well in metaphoric names identification in flower and plant names. Given that we cannot further fine-tune the GPT model with our corpus, the ChatGPT is struggling to identify and generate text with metaphoric flower and plant names. Another important observation was, ChatGPT was not producing consistent results because we could observe different results for the same sentence if we retrieve twice. This shows that ChatGPT is uncertain about its answers on metaphoric flower and plant names, maybe with GPT-4 it may have a better understanding with more data. We leave it for future work.

5.2 Spanish

Table 3 shows the results on Spanish data in zero-shot configuration on English data. We note that in all models, learning from English data has lead to decent results on Spanish metaphoric flower and plant names identification. Interestingly, bert-base-multilingual-cased model performs better in both languages marking over 52% F1 score on Spanish. It was noted that there is a significant difference between English and Spanish results, as expected because the English models were fine-tuned on English metaphoric data, but we were not able to do that in Spanish due to lack of resources.

ChatGPT has kept similar performance for Spanish recording over 51% F1 score. This is very close value to the best discriminative model but could not outperform bert-base-multilingual-cased model. Unlike ChatGPT, since discriminative models are able to fine-tune, we conjecture that their performance could be boosted with a fine-tuning step with more data.

6 Conclusions

The detection of metaphorical terms is an important research area for many NLP applications. Detecting metaphor-based terms of flowers and plants may give birth to different multidisciplinary research and applications. On the one hand, it may help in overcoming the so-called plant awareness disparity or plant blindness (Parsley, 2020) as the metaphoric factor would help in remembering the names of flowers and plants and their aspect. It may also give insightful information to Cognitive Studies towards understanding phenomena such as metaphor and metonymy, and even towards a more comprehensive understanding of conceptual complexes (Ruiz de Mendoza Ibáñez, 2017). This may be carried out by comprehending the associations between metaphoric names and the image of the flower and plant representing them, and how the resemblance of images or the metonymic aspect is conceptualised through the coinage of terms. On the other hand, this information is also helpful for the studies of representation of abstract phenomena in art and its comprehension across languages. The automatic extraction of those terms is a step towards achieving more comprehensive and accurate results. In addition, this may help rendering texts more accessible to people with ASD. At the same time, these types of studies may also help in the development of software or mobile

applications to be used by both laypersons and professionals.

In conclusion, we show that the state-of-the-art transformers are well capable of performing excellently in identifying metaphoric flower and plant names.

7 Acknowledgements

Bibliography

Beltagy, I., K. Lo, and A. Cohan. 2019. SciBERT: A pretrained language model for scientific text. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3615–3620, Hong Kong, China, November. Association for Computational Linguistics.

Brickell, C. 2012. *Encyclopedia of plants and flowers.* Dorling Kindersley, Santa Fe, New Mexico, USA.

Clark, K., M.-T. Luong, Q. V. Le, and C. D. Manning. 2020. ELECTRA: Pre-training text encoders as discriminators rather than generators. In *ICLR*.

Coll-Florit, M. and S. Climent. 2019. A new methodology for conceptual metaphor detection and formulation in corpora: A case study on a mental health corpus. *SKY Journal of Linguistics*, 32.

Conneau, A., K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, Online, July. Association for Computational Linguistics.

de Lorenzo Cáceres, J. M. S. 1999. *Los Árboles en España: Manual de Identificación.* Mundi-Prensa, Spain.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June. Association for Computational Linguistics.

Dweck, A. 2004. A review of edelweiss. *Sofw Journal*, 130(9):65–71.

Dębowiak, P. and J. Waniakowa. 2019. Semantic motivation of plant names as a part of their etymology. page 173–200, 08.

Goodman, J. S. 1963. Malayalam color categories. *Anthropological Linguistics*, pages 1–12.

Group, P. 2007. Mip: A method for identifying metaphorically used words in discourse. *Metaphor and symbol*, 22(1):1–39.

Gutierrez, E. D., E. Shutova, T. Marghetis, and B. Bergen. 2016. Literal and metaphorical senses in compositional distributional semantic models. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 183–193.

Jang, H., S. Moon, Y. Jo, and C. Rose. 2015. Metaphor detection in discourse. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 384–392.

Lakoff, G. and M. Johnson. 2008. *Metaphors we live by.* University of Chicago press.

Leong, C. W., B. B. Klebanov, C. Hamill, E. Stemle, R. Ubale, and X. Chen. 2020. A report on the 2020 vua and toefl metaphor detection shared task. In *Proceedings of the second workshop on figurative language processing*, pages 18–29.

Mitkov, R., H. M. Le An Ha, T. Ranasinghe, and V. Sosoni. 2023. Automatic generation of multiple-choice test items from paragraphs using deep neural networks. *Advancing Natural Language Processing in Educational Assessment*, page 77.

Mu, J., H. Yannakoudakis, and E. Shutova. 2019. Learning outside the box: Discourse-level features improve metaphor identification. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 596–601, Minneapolis, Minnesota, June. Association for Computational Linguistics.

Nacey, S., A. G. Dorst, T. Krennmayr, and W. G. Reijnierse. 2019. *Metaphor identification in multiple languages: MIPVU around the world*, volume 22. John Benjamins Publishing Company.

Nissan, E. 2014. Multilingual lexis, semantics, and onomasiology. terminological database modelling, by using the cupros metarepresentation language: An xml-compatible xml-precursor enabling flexible nested-relation structures. *Language, Culture, Computation. Computational Linguistics and Linguistics: Essays Dedicated to Yaacov Choueka on the Occasion of His 75th Birthday, Part III*, pages 122–173.

Orăsan, C., R. Evans, and R. Mitkov. 2018. Intelligent text processing to help readers with autism. *Intelligent Natural Language Processing: Trends and Applications*, pages 713–740.

Parsley, K. M. 2020. Plant awareness disparity: A case for renaming plant blindness. *Plants, People, Planet*, 2(6):598–601.

Peñas, J. and J. Lorite. 2019. *Biología de la conservación de plantas en Sierra Nevada*. Universidad de Granada.

Premasiri, D., A. H. Haddad, T. Ranasinghe, and R. Mitkov. 2022. Transformer-based detection of multiword expressions in flower and plant names. In *Proceedings of the 5th Workshop on Multi-word Units in Machine Translation and Translation Technology (MUMTTT 2022)*, September.

Premasiri, D. and T. Ranasinghe. 2022. Bert (s) to detect multiword expressions. In *Proceedinsg of the 4th International Conference on Computational and Corpus-Based Phraseology, Europhras 2022*. Springer Cham.

Premasiri, D., T. Ranasinghe, W. Zaghouani, and R. Mitkov. 2022. DTW at qur'an QA 2022: Utilising transfer learning with transformers for question answering in a low-resource domain. In *Proceedinsg of the 5th Workshop on Open-Source Arabic Corpora and Processing Tools with Shared Tasks on Qur'an QA and Fine-Grained Hate Speech Detection*, pages 88–95, Marseille, France, June. European Language Resources Association.

Radford, A., K. Narasimhan, T. Salimans, I. Sutskever, et al. 2018. Improving language understanding by generative pre-training.

Ranasinghe, T., D. Sarkar, M. Zampieri, and A. Ororbia. 2021. WLV-RIT at SemEval-2021 task 5: A neural transformer framework for detecting toxic spans. In *Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021)*, pages 833–840, Online, August. Association for Computational Linguistics.

Ranasinghe, T. and M. Zampieri. 2021. MUDES: Multilingual detection of offensive spans. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations*, pages 144–152, Online, June. Association for Computational Linguistics.

Ranasinghe, T., M. Zampieri, and H. Hettiarachchi. 2019. BRUMS at HASOC 2019: Deep Learning Models for Multilingual Hate Speech and Offensive Language Identification. In *Forum for Information Retrieval Evaluation*, pages 199–207.

Rastall, P. 1996. Metaphor and the names of plants. *English Today*, 12(2):30–31.

Razali, M. S., A. A. Halin, Y.-W. Chow, N. M. Norowi, and S. Doraisamy. 2022. Deep and contextually engineered features for metaphor detection.

Rodríguez Penagos, C. et al. 2005. *Metalinguistic information extraction from specialized texts to enrich computational lexicons*. Ph.D. thesis, Universitat Pompeu Fabra.

Ruiz de Mendoza Ibáñez, F. J. 2017. Conceptual complexes in cognitive modeling. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, 30(1):299–324.

Štajner, S., V. Yaneva, R. Mitkov, and S. P. Ponzetto. 2017. Effects of lexical properties on viewing time per word in autistic and neurotypical readers. Association of Computational Linguistics.

Su, C., F. Fukumoto, X. Huang, J. Li, R. Wang, and Z. Chen. 2020. Deepmet: A reading comprehension paradigm for token-level metaphor detection. In *Proceedings of the second workshop on figurative language processing*, pages 30–39.

Tjong Kim Sang, E. F. and F. De Meulder. 2003. Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition. In *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003*, pages 142–147.

Turney, P., Y. Neuman, D. Assaf, and Y. Cohen. 2011. Literal and metaphorical sense identification through concrete and abstract context. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 680–690.

Urena Gomez-Moreno, J. M. and P. Faber. 2010. Strategies for the semi-automatic retrieval of metaphorical terms. *Metaphor and Symbol*, 26(1):23–52.

Uyangodage, L., T. Ranasinghe, and H. Hettiarachchi. 2021. Can multilingual transformers fight the COVID-19 infodemic? In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, pages 1432–1437, Held Online, September. INCOMA Ltd.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Vitez, A. Z., M. Brglez, M. Robnik-Šikonja, T. Škvorc, A. Vezovnik, and S. Pollak. 2022. Extracting and analysing metaphors in migration media discourse: towards a metaphor annotation scheme. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 2430–2439.

Yaneva, V. 2016. Assessing text and web accessibility for people with autism spectrum disorder.

A Appendix 1: English And Spanish Corpus

List of references used to compile the corpus in English and Spanish.

References
Brickell, Christopher. Encyclopedia of Plants and Flowers. New York, Dorling Kindersley, 2012
Vigneron, Jean Pol, et al. "Optical Structure and Function of the White Filamentary Hair Covering the Edelweiss Bracts." Physical Review E, vol. 71, no. 1, 19 Jan. 2005
Maghiar, Lăcrămioara M., et al. "Integrating Demography and Distribution Modeling for the Iconic Leontopodium Alpinum Colm. In the Romanian Carpathians." Ecology and Evolution, vol. 11, no. 18, 25 Aug. 2021, pp. 12322–12334
Blanco-Pastor, J. L., et al. "Past and Future Demographic Dynamics of Alpine Species: Limited Genetic Consequences despite Dramatic Range Contraction in a Plant from the Spanish Sierra Nevada." Molecular Ecology, vol. 22, no. 16, 12 July 2013, pp. 4177–4195
Ni, Lianghong, et al. "Migration Patterns of Gentiana Crassicaulis, an Alpine Gentian Endemic to the Himalaya–Hengduan Mountains." Ecology and Evolution, vol. 12, no. 3, Mar. 2022
Pink, Alfred. Dictionary of Flowers and Plants for Gardening. Teresa Thomas Bohannon, 2008

Table 4: English Corpora.

Reference
Gómez García, Daniel. Flora y Vegetación de La Jacetania. De la Naturaleza. Diputación General de Aragón, 2004
M. López Guadalupe, et al. "Comunidades, Hábitat Y Tipos de Suelos Sobre Los Que Se Desarrolla La Manzanilla de Sierra Nevada." Ars Pharmaceutica (Internet), vol. 26, no. 4, 1985, pp. 255–263
Pugnaire, Francisco, et al. "Facilitación de Las Especies Almohadilladas Y Cambio Global En Las Comunidades Alpinas Del Parque Nacional de Sierra Nevada." Proyectos de Investigación En Parques Nacionales: 2010-2013, by Ministerio de Agricultura, Ministerio de Agricultura, Alimentación y Medio Ambiente. Organismo Autónomo de Parques Nacionales, 2015, pp. 91–104
Montserrat, P, and R Balcells. "LA FLORA DEL PIRINEO." Sinergia (Publicación Paramédica de Sociedad General de Farmacia, SA), vol. 14, no. 6675, 1960
Paúl Gonzáles, et al. Las Plantas Comunes Del Bosque Seco Del Marañón: Biodiversidad Para Las Comunidades Locales	Bosques Andinos. Bosques Andinos, Lima- Perú, Biblioteca Nacional del Perú, 2020
Blanca, Gabriel. Flora Amenazada Endémica de Sierra Nevada. Granada, Universidad de Granada, 2001
Peñas, Julio, and Juan Lorite. "Biología de La Conservación de Plantas En Sierra Nevada. Principios Y Retos Para Su Preservación - Universidad de Granada." Editorial.ugr.es, Universidad de Granada, 2019
Ministerio de Salud y Protección Social de Colombia. "Vademecum Colombiano de Plantas Medicinales." Orasconhu.org, Ministerio de Salud y Protección Social de Colombia, 2008
Sánchez de Lorenzo Cáceres, Jose Manuel. Los Árboles En España: Manual de Identificación. Ediciones Mundi-Prensa, 1999

Table 5: Spanish Corpora.