ISSN 2007-9737

Lexical Function Detection in Spanish Collocations Using Transformer Architecture

Olga Kolesnikova*

Instituto Politécnico Nacional,
Centro de Investigación en Computación, Mexico City,
Mexico

kolesolga@gmail.com

Abstract. In this article, we study the abilities of transformer models to detect verb-noun lexical functions in Spanish collocations from context. The concept of lexical functions is a formalism to represent recurrent relations among words. A lexical function (LF) takes a word as input and outputs a set of words related to the input in a paradigmatic or syntagmatic way. For example, the syntagmatic LF Oper1 takes the noun *decision* as input and outputs the verb *make* with the semantics of 'Agent realizes the action denoted by the noun'. Oper1 captures the relation between the noun and the verb in many collocations such as *make a decision*, *take a walk*, *give a lecture*, *pay a compliment*, *keep a promise*, etc. The numeric part of the Oper1 notation represents that (1) the action of the verb is performed by the agent which is the first argument in the verb's subcategorization frame, (2) the syntactic function of the noun is subject. In general, lexical functions represent common semantic and syntactic patterns typical for certain word classes and can aid in many natural language processing tasks, especially in word sense disambiguation. In this article we report the results of our experiments with transformer models on the task of detecting verb-noun lexical functions.

Keywords. Collocation, lexical function, syntagmatic relations, transformer models, deep learning.

1 Introduction

Natural Language Processing (NLP) is a branch of artificial intelligence that enables computers to understand, interpret, and generate human language. It involves analyzing the structure and meaning of text or speech, allowing machines to perform tasks like translation, sentiment analysis, and information retrieval. NLP combines techniques from linguistics, computer science, and machine learning to process language in ways that mimic human understanding. Key challenges include dealing with ambiguity, context, and variability in language, as well as understanding idioms and slang. Modern NLP relies heavily on deep learning models, such as transformers, which are pre-trained on vast amounts of text data and fine-tuned for specific tasks like chatbots, summarization, or speech recognition. As NLP advances, it continues to improve human-computer interaction, making technology more intuitive and accessible. As we have mentioned, ambiguity of language items and lexicon diversity are one of the toughest issues in NLP. To contribute in solving these issues, various formal semantic concepts and models have been developed. In this article, we discuss one of these: the concept of lexical functions and study the abilities of modern transformer models to detect them from text. Lexical functions (LFs) are a central concept in the Meaning-Text Theory (MTT) proposed by Mel'čuk and Žolkovskij [11] where LFs are used to describe the systematic semantic and syntactic relationships between words and their meanings [12, 13]. LFs capture the predictable ways in which words combine and interact in language and are applied to model how certain words or phrases (the keyword) systematically evoke specific related words or expressions (the value) in a given context. In the next section we give more details on LFs.

2 Lexical Functions

A lexical function (LF) takes a word as input and outputs a set of words related to the input in a certain way on the paradigmatic or syntagmatic level. It is a function in the mathematical sense defined as a mapping from a word w_0 called the lexical function argument to the lexical function value which is a set of words $\{w_1, w_2, \ldots w_n\}$ where each word w_i, $1 \le i \le n$, has a particular (and the same) lexical relation with the argument w_0; so using mathematical notation, lexical function (LF) is represented as $LF(w_0) = \{w_1, w_2, \ldots w_n\}$.

Wanner [15] states that lexical function is a concept which can be used to systematically describe "institutionalized" lexical relations clarifying that "a lexical relation is institutionalized if it holds between two lexical units L_1 and L_2 and has the following characteristics: if L_1 is chosen to express a particular meaning M, its choice is predetermined by the relation of M to L_2 to such an extent that in case M and L_2 is given, the choice of L_1 is a language-specific automatism".

Institutionalized lexical relations can be of two types: paradigmatic and syntagmatic. Paradigmatic relations are observed between lexical units within a lexicon (examples of paradigmatic relations are synonymy, antonymy, hypernymy, hyponymy, etc.) and syntagmatic relations hold between lexical units that co-occur in texts (*make a decision*, *friendly attitude*, *rain cats and dogs*).

LFs have the following key features:

– Abstract relationships: Lexical functions represent abstract semantic relationships between words. For example, they describe how a verb might relate to its typical subject, object, or adverb, or how a noun might relate to its typical adjective or verb.

– Standardized notation: Lexical functions are denoted by symbols (e.g., Magn for intensification, Oper for a verb that relates to a noun, Syn for synonyms, etc.).

– Language-independent: While the specific realizations of lexical functions depend on the language, the abstract relationships they represent are universal, making them useful for cross-linguistic analysis and machine translation.

Here we give some examples of LF:

– Magn (Intensification): Represents an intensifier or a word that strengthens the meaning of the keyword, e.g., for the keyword *rain*, the value of Magn might be *heavy* (as in *heavy rain*).

– Oper (Support Verb): Represents a verb that typically accompanies a noun to form a standard collocation, e.g., for the keyword *decision*, the value of Oper might be *make* (as in *make a decision*).

– Syn (Synonym): Represents a word with a similar meaning to the keyword, e.g., for the keyword *happy*, the value of Syn might be *joyful*.

– Anti (Antonym): Represents a word with the opposite meaning of the keyword, e.g., for the keyword *happy*, the value of Anti might be *sad*.

– Incep **(**Inceptive**):** Represents the beginning of an action or state, e.g., for the keyword *rain*, the value of Incep might be *start* (as in *the rain started*).

LFs can be applied in many areas, for instance:

– Lexicography: LFs help create more detailed and systematic dictionaries by capturing predictable word relationships.

– Machine translation: LFs provide a framework for translating collocations and idiomatic expressions between languages.

– Natural language generation: LFs guide the selection of appropriate words and phrases to produce coherent and natural-sounding text.

– Language learning: LFs help learners understand how words systematically combine in a language.

In summary, LFs are a powerful tool for modeling the predictable and systematic relationships between words, enabling a deeper understanding of language structure and facilitating applications in computational linguistics and language technology.

ISSN 2007-9737

Table 1. Lexical functions in Spanish verb-noun collocations in our dataset, in their full notation. For each lexical function, we give its description, an example of a collocation from our dataset, and its English translation

Lexical Function	Description	Examples	
		Spanish	English translation
AntiReal3	Failure to fulfill the typical purpose of the event (noun) with respect to the patient of the action (verb)	*violar el derecho*	violate the right
Caus1Func1	Causation of the realization of the event (noun) by the agent	*sacar provecho*	take advantage
Caus1Oper1	Causation of the event (noun) by the agent	*dar un resultado*	give a result
Caus2Func1	Experiencing of the event (noun) caused by a non-agent of the situation	*dar miedo*	cause fear
CausFunc0	Existence of an entity (noun) caused by an unidentified participant of the situation	*el plan se elabora*	the plan is developed
CausFunc1	Existence of an entity (noun) caused by the agent	*ofrecer servicio*	provide a service
CausManifFunc0	Existence and exhibition of an entity (noun) caused by an unidentified participant of the situation	*el concurso se anuncia*	the competition is advertised
CausMinusFunc0	Decrease of the realization of an entity (noun) caused by an unidentified participant of the situation	*el riesgo se reduce*	the risk is reduced
CausMinusFunc1	Decrease of the realization of an entity (noun) caused by the agent	*reducir el número*	reduce the number
CausPerfFunc0	Existence and complete realization of an entity (noun) caused by an unidentified participant of the situation	*el derecho se garantiza*	he right is guaranteed
CausPlusFunc0	Increasing realization of an entity (noun) caused by an unidentified participant of the situation	*el desarrollo se favorece*	the development is favored
CausPlusFunc1	Increase of the realization of an entity (noun) caused by the agent	*promover el desarrollo*	promote the development
ContOper1	Continuation of performing the event (noun) by the agent	*mantener la relación*	keep the relation
Copul	Linking verb	*ser parte*	be a part of
FinFunc0	Termination of the realization of an event (noun)	*el plazo transcurre*	the time period elapsed
FinOper1	Termination of the realization of an event (noun) by the agent	*perder control*	lose control
Func0	Realization of an event (noun)	*tiempo pasó*	time passed
Func1	Realization of an event (noun) by the agent	*(me) quedó duda*	a doubt remained
IncepFunc0	Commencement of realization of an event (noun)	*la hora llega*	the hour comes
IncepOper1	Commencement of realization of an event (noun) by the agent	*iniciar una sesión*	start a session
IncepReal1	Commencement of realization of the typical purpose an event (noun) by the agent	*abordar un problema*	attack a problem
LiquFunc0	Abortion of the realization of an event (noun)	*el problema se evita*	the problem is avoided
Manif	Exhibition of an event (noun)	*mostrar interés*	show interest
ManifFunc0	Existence and exhibition of an entity (noun)	*la pregunta se plantea*	the question is raised
MinusReal1	Decrease of realization of the typical purpose an event (noun) by the agent	*gastar dinero*	spend money
Oper1	Perform an event (noun) by the agent	*prestar atención*	pay attention
Oper2	Experiencing an event (noun) by the recipient	*recibir atención*	receive attention
Oper3	Experiencing an event (noun) by the patient	*contener información*	contain information
PerfFunc0	Complete realization of an event (noun)	*el momento llega*	the moment comes
PerfOper1	Perform an event (noun) to its full extent by the agent	*tomar precaución*	take precaution
PermOper1	Allow to perform an event (noun) by the agent	*permitir acceso*	permit access
Real1	Fulfillment of the typical purpose of the event (noun) with respect to the agent	*contestar una pregunta*	answer a question
Real2	Fulfillment of the typical purpose of the event (noun) with respect to the recipient	*merecer atención*	deserve attention
Real3	Fulfillment of the typical purpose of the event (noun) with respect to the patient	*reconocer el derecho*	recognize the right
PerfOper1	Perform an event (noun) to its full extent by the agent	*tomar precaución*	take precaution
PermOper1	Allow to perform an event (noun) by the agent	*permitir acceso*	permit access

ISSN 2007-9737

In this article, we consider several types of lexical functions found in verb-noun collocations. Our objective is to automatically detect LFs in text, using Spanish verb-noun collocations as a case study.

3 Lexical Functions in Verb-Noun Collocations

In this section, we consider lexical functions in verb-noun collocations. Table 1 presents the lexical functions in Spanish verb-noun collocations found in our dataset.

As it can be seen in Table 1, LFs can be simple and compound. A simple LF represents a single semantic unit and is denoted with an abbreviated Latin word reflecting the function's meaning. A compound LF includes more than one semantic unit. For example, Oper (Latin, *operor*, perform) and Incep (Latin, *incepere,* begin) are simple LFs meaning to perform and to begin, respectively. They are used to construct a compound LF IncepOper meaning to begin to perform (an action), e.g., as in *acquire a habit*, *run into trouble*. LFs describe not only semantics in collocations, specifically for verb-noun collocations in our dataset, but also the syntactic relations among collocational elements using subscript numbers to identify semantic roles of the arguments in the verb's subcategorization frame.

The number 1 denotes the agent, 2 is used for the recipient, 3 for the patient, and the order of the numbers explains the syntactic functions of the semantic roles. For example, Oper1 means to perform an action, the agent is the subject in sentences where Oper1 is used: *The professor applied the exam*. In Oper2, the patient of the action is the subject: *The student passed the exam*. Oper12 means that the subject in a sentence is the agent, and the recipient is the object: "*I feel enormous sympathy for people that live in poverty and fear*."[1] If the number is zero, there is no agent neither recipient, the action realizes itself, e.g., Func0 in *snow falls*.

[1] https://www.europarl.europa.eu/doceo/document/CRE-6-2006-04-06_EN.html?redirect/

4 Automatic Detection of Lexical Functions

Lexical functions represent common semantic and syntactic patterns of certain word classes and can aid in many tasks of natural language processing, lexical and syntactic disambiguation being the most fundamental one among them. In this section, we review some research works on automatic detection of LFs in texts.

The work in [4] explores the use of supervised learning algorithms to classify Spanish verb-noun collocations according to the LFs typology. The authors aim to predict the meaning of previously unseen collocations by identifying semantic patterns in a manually annotated training set. The problem was defined as classification of Spanish collocations into nine semantic classes (eight LFs and one class for free word combinations, FWC) using supervised learning. The goal was to identify the best-performing classifiers for each semantic class. Their dataset consists of 1,000 frequent verb-noun pairs from the Spanish Web Corpus[2], annotated with LFs and word senses from Spanish WordNet. The authors used 68 classifiers, representing each verb-noun pair with binary features based on hypernyms from WordNet[3] [3]. Performance was evaluated using precision, recall, and F1-score with 10-fold cross-validation. Different classifiers performed best for different LFs, with no single classifier dominating across all classes. The highest F1-score of 0.87 was achieved by the BayesianLogisticRegression classifier for Oper1 and the average F1-score was 0.74.

The research in [7] provides a comprehensive list of LFs describing their meaning and giving examples of English. It presents an overview of state of the art methods for LFs detection om English and in Spanish highlighting the best result of Oper1 detection with SimpleCart algorithm: its F1-score reached 0.876. High F1-scores were shown also for Func0 (0.824 by AttributeSelectedClassifier) and ContOper1 (0.800 by LWL classifier). A lot of attention in this work is given to application of LFs in different areas of

[2] https://www.sketchengine.eu/estenten-spanish-corpus/
[3] https://wordnet.princeton.edu

Table 2. Dataset statistics

Category	Train	Test	Total per category
CausFunc0	7,052	1,763	8,815
CausFunc1	12,706	3,177	15,883
FWC	8,950	2,237	11,187
Oper1	13,457	3,364	16,821
Real1	12,799	3,200	15,999
Total per subset	**54,964**	**13,741**	
Total in dataset			**68,705**

Table 3. BERT results on detecting lexical functions and free word combinations

Category	Precision	Recall	F1-score
CausFunc0	0.32	0.27	0.29
CausFunc1	0.56	0.57	**0.57**
FWC	0.38	0.39	0.39
Oper1	0.50	0.52	0.51
Real1	0.55	0.57	0.56

Table 4. RoBERTa results on detecting lexical functions and free word combinations

Category	Precision	Recall	F1-score
CausFunc0	0.34	0.26	0.29
CausFunc1	0.57	0.61	**0.59**
FWC	0.42	0.37	0.40
Oper1	0.52	0.55	0.53
Real1	0.55	0.59	0.57

Table 5. DistilBERT results on detecting lexical functions and free word combinations

Category	Precision	Recall	F1-score
CausFunc0	0.30	0.24	0.27
CausFunc1	0.55	0.60	**0.57**
FWC	0.40	0.36	0.38
Oper1	0.49	0.52	0.51
Real1	0.55	0.56	0.56

natural language processing, linguistics and language teaching and learning.

In the paper [8], the use of word2vec embeddings and supervised machine learning to detect LFs in Spanish verb-noun collocations without relying on manually annotated resources was examined. The dataset consists of 240 Spanish verb-noun collocations (60 for each of the

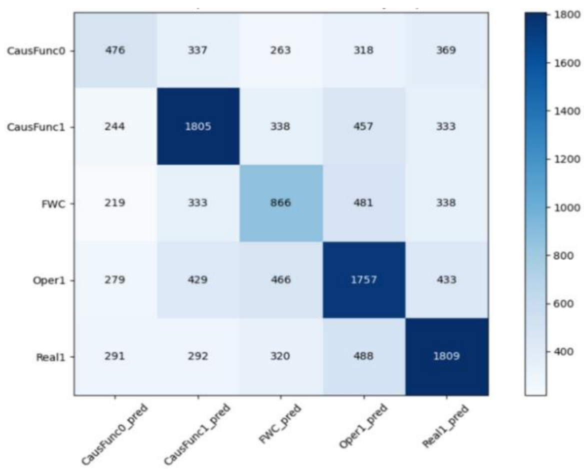

Fig. 1. Confusion matrix for BERT

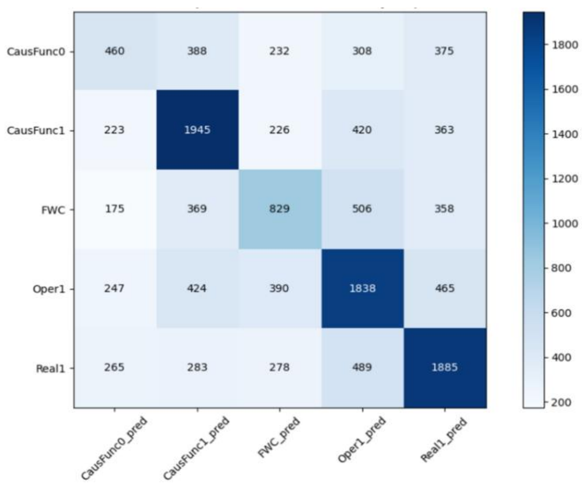

Fig. 2. Confusion matrix for RoBERTa

four LFs) and 60 free word combinations (FWC, non-collocations). The corpus used was a collection of 1,131 issues of the Excelsior newspaper. The authors lemmatized the corpus and removed stopwords before training word2vec embeddings. Six supervised learning algorithms were used: Support Vector Machine, Multi-layered Perceptron, k-Nearest Neighbors, Decision Tree, Random Forest, and Ada Boost.

The best results were achieved with word2vec embeddings of 90 to 180 dimensions. Multi-layered Perceptron achieved the highest F1-score of 0.72 for Oper1. Random Forest achieved the highest F1-score of 0.79 for Real1. Support Vector Machine reached its highest F1-score of 0.84 for CausFunc1. The results showed that word2vec embeddings outperform traditional bag-of-words representations, capturing LFs contextual characteristics more effectively.

5 Task and Experimental Setup

In this work we use the same dataset as in [7]: 240 Spanish verb-noun collocations (60 for each of the four lexical functions: CausFunc0, CausFunc1, Oper1, Real1) and 60 free word combinations (FWC), that is, non-collocations.

For each collocation, we collected its context from the texts of 1,131 issues of the Excelsior newspaper[4] using the window of eight (four words to the left of the verb and four words to the right of the noun) after lemmatization and stopwords deletion.

Our task was to fine-tune transformer models on the training subset of our dataset in order to detect lexical functions in unseen collocations from the test subset based on their contexts.

The dataset includes 68,705 samples of the five categories: CausFunc0, CausFunc1, FWC, Oper1, Real1. We divided the dataset into training and text subsets as shown in Table 2.

We used the following models from the Hugging Face platform[5], for each model we indicate the version applied in our experiments:

1. BERT [2]: ‘bert-base-uncased’,
2. RoBERTa [10]: ‘roberta-base’,
3. DistilBERT [14]: ‘distilbert-base-uncased’,
4. AlBERT [9]: ‘albert-base-v2’,
5. BETO [1]: ‘dccuchile/bert-base-spanish-wwm-uncased’

All the models mentioned above were implemented using the class ‘ClassificationModel’ from the library ‘simpletransformers’[6].

For all the models, we fine-tuned them on the training set with the following configuration:

[4] The dataset is accessible upon request to the author on Creative Commons License.

[5] https://huggingface.co

[6] https://simpletransformers.ai/docs/classification-models/

Table 6. AlBERT results on detecting lexical functions and free word combinations

Category	Precision	Recall	F1-score
CausFunc0	0.25	0.03	0.06
CausFunc1	0.48	0.56	**0.51**
FWC	0.32	0.12	0.18
Oper1	0.39	0.57	0.46
Real1	0.43	0.54	0.48

Table 7. BETO results on detecting lexical functions and free word combinations

Category	Precision	Recall	F1-score
CausFunc0	0.37	0.33	0.35
CausFunc1	0.58	0.62	**0.60**
FWC	0.47	0.43	0.45
Oper1	0.55	0.57	0.56
Real1	0.59	0.60	**0.60**

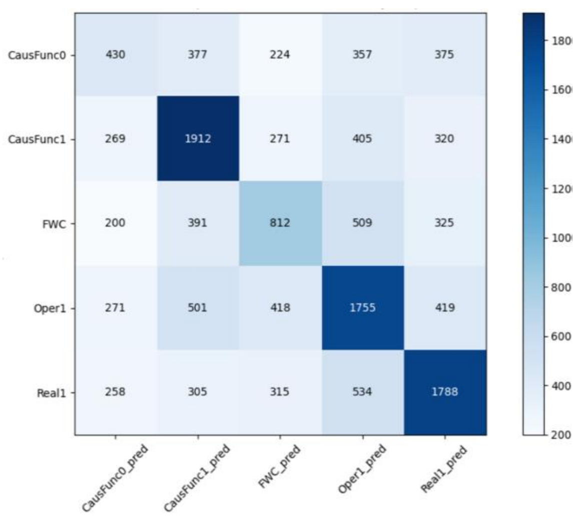

Fig. 3. Confusion matrix for DistilBERT

The number of epochs: 5,

Learning rate: 2e-5,

Optimizer: AdamW,

Loss function: Cross-Entropy.

To evaluate the performance of the models on detecting lexical functions and free word combinations in the test set, we used confusion matrix, precision, recall, and F1-score.

6 Results and Discussion

In this section we report and discuss the results we obtained in the experiments with the models mentioned in the previous section.

BERT model showed an accuracy of 0.49, values of precision, recall, and F1-score are given in Table 3, Figure 1 shows the confusion matrix. The best F1-score produced by BERT is 0.57 for CausFunc1.

RoBERTa model showed an accuracy of 0.51, values of precision, recall, and F1-score are given in Table 4, Figure 2 shows the confusion matrix.

The best F1-score produced by RoBERTa is 0.59 for CausFunc1

DistilBERT model showed an accuracy of 0.49, values of precision, recall, and F1-score are given in Table 5, Figure 3 shows the confusion matrix. The best F1-score produced by DistilBERT is 0.57 for CausFunc1.

AlBERT model showed an accuracy of 0.42, values of precision, recall, and F1-score are given in Table 6, Figure 4 shows the confusion matrix. The best F1-score produced by AlBERT is 0.51 for CausFunc1.

BETO model showed an accuracy of 0.54, values of precision, recall, and F1-score are given in Table 7, Figure 5 shows the confusion matrix. The best F1-score produced by BETO is 0.60 for CausFunc1 and Real1.

The experimental results demonstrate varying performance levels across different transformer-based models in detecting lexical functions and free word combinations.

Among the models evaluated, BETO achieved the highest overall accuracy of 0.54, along with the best F1-scores for CausFunc1 and Real1 (0.60 each).

This suggests that BETO, a Spanish-language variant of BERT, may be particularly well-suited for this task, possibly due to its specialized training on Spanish corpora. RoBERTa also performed competitively, with an accuracy of 0.51 and an F1-score of 0.59 for CausFunc1, indicating that its optimized pretraining approach contributes to robust performance.

ISSN 2007-9737

BERT and DistilBERT showed similar results, with accuracies of 0.49 and identical best F1-scores of 0.57 for CausFunc1. This similarity is noteworthy, as DistilBERT is a distilled version of BERT designed for efficiency. The comparable performance suggests that DistilBERT retains much of BERT's capability while being more lightweight. However, AlBERT lagged behind the other models, with the lowest accuracy (0.42) and F1-scores, particularly for CausFunc0 and FWC, where its performance dropped significantly (0.06 and 0.18, respectively). This may indicate that AlBERT's parameter-sharing mechanism, while reducing model size, could compromise its effectiveness for certain lexical function categories.

Across all models, CausFunc1 consistently achieved the highest F1-scores, suggesting that this category is more distinguishable or better represented in the training data. In contrast, CausFunc0 and FWC were consistently challenging for all models, with lower precision and recall values. This could reflect inherent ambiguities or overlaps between these categories and others in the dataset. The confusion matrices (Figures 1–5) likely provide further insights into these misclassifications, though specific details would require visual inspection.

Overall, the results highlight the importance of model selection for tasks involving lexical function detection. BETO's superior performance underscores the potential benefits of language-specific pretraining, while RoBERTa's strong results emphasize the value of optimized training procedures. The challenges faced by AlBERT and the similar performance of BERT and DistilBERT suggest trade-offs between model efficiency and effectiveness that warrant further investigation. Future work could explore additional fine-tuning strategies or dataset augmentation to improve performance on the weaker categories.

7 Conclusions and Future Work

The study explored the automatic detection of lexical functions in Spanish verb-noun collocations using transformer-based models, demonstrating the potential and limitations of current approaches. BETO emerged as the top-performing model, achieving the highest accuracy (0.54) and F1-scores (0.60 for CausFunc1 and Real1), demonstrating the advantage of language-specific pretraining for tasks involving nuanced semantic relationships. RoBERTa also delivered strong results, suggesting that optimized pretraining methodologies can enhance performance even in multilingual contexts.

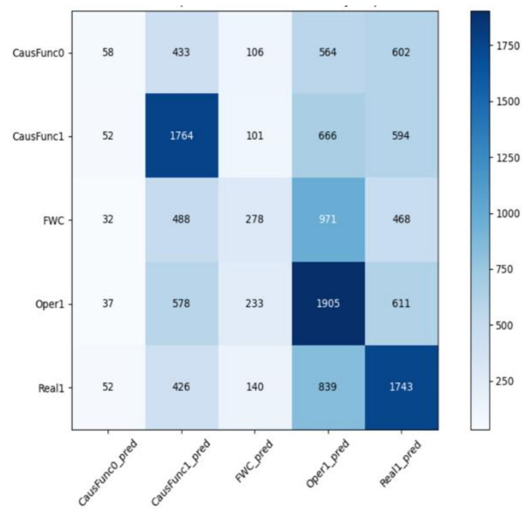

Fig. 4. Confusion matrix for AlBERT

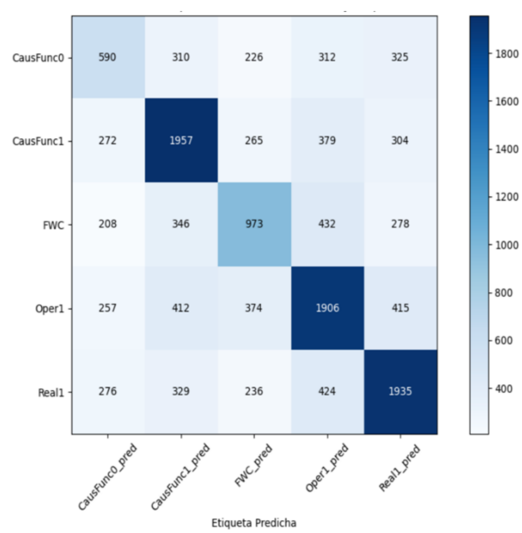

Fig. 5. Confusion matrix for BETO

The comparable performance of BERT and its distilled variant, DistilBERT, indicates that model efficiency can be improved without significant trade-offs in effectiveness, while AlBERT's weaker results highlight the challenges of balancing parameter reduction with task-specific accuracy.

The consistent difficulty in classifying CausFunc0 and free word combinations (FWC) across all models points to inherent ambiguities or data sparsity in these categories. This suggests that lexical function detection may benefit from richer contextual representations or additional linguistic features to disambiguate such cases. The success in identifying CausFunc1 and Real1, however, validates the feasibility of using transformer models to capture systematic semantic patterns in collocations, aligning with the theoretical framework of Meaning-Text Theory [12]. Note that similar methods are applied in other areas, for example, in music [5].

Future work could explore several directions to improve performance and applicability. First, expanding the dataset to include more examples of underperforming categories, such as CausFunc0 and FWC, could address data imbalance and improve model generalization. Second, incorporating syntactic or dependency-based features alongside contextual embeddings might enhance the models' ability to discern subtle semantic-syntactic interactions. Third, investigating hybrid approaches—combining transformer models with rule-based methods or knowledge graphs—could leverage the strengths of both symbolic and statistical paradigms. Finally, extending this research to other languages and collocation types would test the universality of the findings and contribute to cross-linguistic NLP applications, such as machine translation or lexicography.

In conclusion, this study advances the automatic detection of lexical functions, offering practical insights for NLP tasks reliant on semantic collocation analysis. While challenges remain, the results open the way for more sophisticated, linguistically informed models capable of capturing the systematicity of human language.

Acknowledgments

The work was done under partial support of Mexican Government: SNII, COFAA-IPN, BEIFI-IPN, and SIP-IPN grants 20241816, 20241819, 20240951.

References

1. **Cañete, J., Chaperon, G., Fuentes, R., Ho, J.H., Kang, H., Pérez, J. (2023)**. Spanish Pre-trained Bert Model and Evaluation Data. DOI: 10.48550/ arXiv.2308.02976.
2. **Devlin, J., Chang, M.W., Lee, K., Toutanova, K. (2019)**. Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1 (long and short papers), pp. 4171–4186. DOI: 10.18653/v1/N19-1423.
3. **Fellbaum, C. (1998).** WordNet: An Electronic Lexical Database, Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/7287.001.0001.
4. **Gelbukh, A., Kolesnikova, O. (2010).** Supervised Learning for Semantic Classification of Spanish Collocations. Advances in Pattern Recognition: Second Mexican Conference on Pattern Recognition, Proceedings, Vol. 2, pp. 362–371. DOI: 10.1007/978-3-642-15992-3_38.
5. **Gelbukh, A., Pérez-Alvarez, D.A., Kolesnikova, O., Chanona-Hernandez, L., Sidorov, G. (2024).** Multi-Instrument Based N-Grams for Composer Classification Task. Computación y Sistemas, Vol. 24, No. 1, pp. 85–98. DOI: 10.13053/CyS-28-1-4903.
6. **Kolesnikova, O., Gelbukh, A. (2019)**. Dictionary and Corpus-Based Study of Lexical Functions in Spanish. POLIBITS, Vol. 60, pp. 43–56. DOI: 10.17 562/PB-60-6.
7. **Kolesnikova, O. (2020).** Automatic Detection of Lexical Functions in Context. Computación y Sistemas, Vol. 24, No. 3, pp.1337–1352. DOI: 10.13 053/cys-24-3-3774.
8. **Kolesnikova, O., Gelbukh, A. (2020).** A Study of Lexical Function Detection with Word2vec and Supervised Machine Learning. Journal of Intelligent & Fuzzy Systems, Vol. 39 No. 2, pp. 1993–2001. DOI:10.3233/JIFS-179866.
9. **Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., Soricut, R. (2019)**. Albert: A Lite Bert for Self-supervised Learning of Language

Representations. arXiv preprint arXiv:1909. 11942. DOI: 10.48550/arXiv.1909.11942.

10. **Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Stoyanov, V. (2019).** Roberta: A Robustly Optimized Bert Pretraining Approach. DOI: 10.48550/arXiv.1907.11692.
11. **Mel'čuk, I.A., Žolkovskij, A.K. (1970).** Towards a Functioning 'Meaning-Text' Model of Language. Linguistics, Vol. 8, No. 57, pp. 10–47. DOI: 10.1515/ ling.1970.8.57.10.
12. **Mel'čuk, I.A. (1996).** Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon. In **Wanner, L. (Ed.).** Lexical Functions in Lexicography and Natural Language Processing, pp. 37–102, Amsterdam, Philadelphia, PA: Benjamins Academic Publishers.
13. **Mel'čuk, I. (2023).** Lexical Functions. General Phraseology, pp. 215–228. John Benjamins Publishing Company. DOI:10.1075/lis.36.app.
14. **Sanh, V., Debut, L., Chaumond, J., Wolf, T. (2019)**. DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter. arXiv: 1910.01108. DOI: 10.48550/arXiv.1910. 01108.
15. **Wanner, L. (2004).** Towards Automatic Fine-Grained Semantic Classification of Verb-Noun Collocations. Natural Language Engineering, Vol. 10, No. 2, pp. 95–143. DOI: 10.1017/S1351324 904003328.

Article received on 14/01/2025; accepted on 21/03/2025.
Corresponding author is Olga Kolesnikova.