Jorge Reyes-Magaña¹, Gerardo Sierra², Gemma Bel-Enguix², Helena Gómez-Adorno³

> ¹ Universidad Autónoma de Yucatán, Facultad de Matemáticas, Mexico

² Universidad Nacional Autónoma de México, Instituto de Ingeniería, Mexico

³ Universidad Nacional Autónoma de México, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Mexico

helena.gomez@iimas.unam.mx, jorge.reyes@correo.uady.mx, {gsierram, gbele}@iingen.unam.mx

This paper presents Lexical Association Abstract. Networks (LexAN), which entail the development of a mathematical model comprising a collection of words derived from a textual corpus. The interconnections between word tokens are represented by weighted edges within a non-directed graph structure. The construction process of LexAN involves 6 stages: 1) Lemmatization 2) Multi-word expressions 3) Stopwords removal 4) Co-ocurrence graph 5) Word Co-ocurrence norms, and 6) LexAN construction. We employed a Medical text corpus containing 574,011 words to build our graphs. To assess the efficacy of our LexAN, these graph structures were implemented within a tool designed to address the lexical access problem, specifically functioning as a reverse dictionary. This application resulted in favorable and promising results.

Keywords. Network, co-occurrence, lexical access.

1 Introduction

Graph Theory has been used for several centuries to address various real-life problems, such as finding the shortest path between two points [15], identifying the most influential individuals within a social network [16], and detecting critical elements in a system that could potentially disrupt the environment [4]. In Natural Language Processing (NLP) graph theory has been used to deal with different tasks. Semantic networks [32] are graph structures that establish relationships between words [1], serving as more than just tools for organizing vocabulary.

They also capture the structure of knowledge. A well-known example of a semantic network is WordNet [21], which connects English nouns, verbs, adjectives, and adverbs by associating them with sets of synonyms and semantic relations that define word meanings.

Ferret [10, 11] and Zock et al. [34] proposed a matrix-based approach to address the challenges of topical detection and collocation links, encompassing both syntactic and semantic contexts in which individual words manifest.

Zock's proposition involves the utilization of intricate double-processing matrices. Employing networks is an alternative and more intuitive strategy to tackle this issue. A direct solution to establish well-balanced syntagmatic-paradigmatic relationships between words is by applying collocation networks [8]. 956 Jorge Reyes-Magaña, Gerardo Sierra, Gemma Bel-Enguix, et al.

Devenuetev	Value			
Parameter	Value			
graph_type	'Graph'			
window_size	10			
parallel_exec	True			
apply_preprocessing	False			
language	'es'			
output_format	'networkx'			

Table 1. Parameters for the co-occur	rrence graph
--------------------------------------	--------------

The researchers used the BNC corpus to construct two distinct graphs, namely G1 and G2. Initially, they created a co-occurrence graph (G1) wherein words are connected if they co-occur within a sentence within a maximum span of three tokens.

Subsequently, a collocation graph (G2) was derived from G1, consisting of only those links whose terminal vertices co-occur more frequently than what would be expected by chance. Bel-Enguix et al. [5] used graph analysis techniques to calculate associations from large collections of texts.

The objective of this study is to introduce Lexical Association Networks (LexAN) through a systematic approach. These graphs consist of nodes representing words and weighted edges that quantify the degree of semantic relatedness among them within a given text corpus.

The primary purpose of constructing these LexAN is to facilitate lexical searches, enabling the identification of terms that correspond to a specific concept based on their surrounding tokens, thus aiding in the retrieval of missing words.

This technique proves valuable in addressing the Tip-of-the-tongue problem, a phenomenon where an individual struggles to recall the exact word that accurately represents a particular meaning.

This challenge is commonly called a lexical access problem [35]. Moreover, this issue can be linked to a reverse dictionary task, as dictionaries of this nature operate by progressing from the definition to the corresponding concept.

Various strategies have been employed to address this challenge, including using graph techniques [26, 25], applied in broad linguistic contexts. However, the LexAN introduced in this study is specifically tailored to identify concepts within highly specialized language domains, showing promising results in the search for specific terms in areas like medicine, engineering, agriculture, etc.

Furthermore, through the application of the proposed technique, we aim to enhance the search precision in a reverse dictionary task by leveraging the principles and techniques of graph theory.

The presented methodology in this paper was developed and validated using a Medical information corpus known as MedGIL. All the stages involved in constructing LexAN can be applied to various types of text corpora, particularly those pertaining to specialized domains.

The paper's structure is organized as follows: Section 2 provides an overview of relevant research related to this type of graph, the lexical access problem, and its potential solutions.

Section 3 outlines the step-by-step methodology employed for constructing LexAN. Sections 4 and 5 present the experimental setup, results, and ensuing discussions. Finally, Section 5 presents the concluding remarks and outlines potential avenues for future research.

2 Related Work

Reyes et al. [26, 25] used Word Association Norms (WAN) to construct semantic networks represented as mathematical graphs, facilitating lexical searches using centrality algorithms.

WANs are assemblies of word associations, typically gathered by presenting a stimulus word to individuals and requesting them to produce the initial word that comes to mind, either verbally or in written form.

The findings report the efficacy of WANs as a solution to the lexical access problem, which is very suitable due to their apt representation of word connections and the interconnectedness of concepts within the human mind.

This methodology has been successfully applied using WAN in both English and Spanish, demonstrating the feasibility of this approach across different languages.

WANs are available in various languages, with the most common ones in English being the Edinburgh Associative Thesaurus ¹ (EAT) [17] and the collection of the University of South Florida² (USF) [22].

For Mexican Spanish, the Corpus de Normas de Asociación de Palabras para el Español de México³ [3] exists.

Additionally, the internet's universality has facilitated the gathering of WANs with the assistance of online users, as demonstrated by the multilingual dataset called Small World of Words⁴.

Some efforts [27] have also been made to generate artificial WANs, using the Diccionario del Español de México as the primary corpus to derive automatic word association norms.

The necessity to advance solutions to the lexical access problem is interconnected with the domain of reverse (onomasiological) dictionaries, a challenge that has been addressed through diverse approaches:

- Psychology. The difficulty in lexical access is regarded as a problem of search [35], prompting interdisciplinary approaches to address this issue. Previous studies have proposed various solutions to tackle this problem [18, 12].
- Linguistics. Writers seeking to find the appropriate word corresponding to a particular meaning or concept can benefit from resources such as thesauri, reverse dictionaries, synonymy and antonymy dictionaries, and pictorial dictionaries. For instance, Roget's Thesaurus of English Words and Phrases [20]. These resources are categorized based on the type of information they provide, the structure of the wordbook, and the search methods employed.
- Computing. Using modern approaches, some new technologies have emerged. The online dictionary called OneLook Reverse Dictionary⁵ for English enables searches in natural language and regular expressions.



Fig. 1. Lexical association network of MedGIL

WantWords is an Open-source Online Reverse Dictionary System developed by Qi et al. [24] using Deep Learning. In the case of Spanish, Sierra [30] proposes a dictionary that accepts user queries in natural language and employs a search engine improved by Hernández [14].

3 Methodology

Our methodology relies on two specific Spanish language corpora, which are described as follows:

3.1 Main Corpus

Plain text containing medical data information. We obtained our corpus MedGIL from MedlinePlus, a service provided by the National Library of Medicine (NLM), the world's largest medical library and a part of the National Institutes of Health (NIH).

We rigorously compiled Spanish documents covering various medical domains, resulting in a corpus comprising 574,011 words. We stored the MedGIL in a Corpus Management System called GECO (GEstor de COrpus) [31] to facilitate our process.

¹rali.iro.umontreal.ca/rali/?q=en/Textual%20Resources/EAT ²web.usf.edu/FreeAssociation

³www.labpsicolinguistica.psicol.unam.mx/Base/php/general.php ⁴smallworldofwords.org/

⁵www.onelook.com/reverse-dictionary.shtml

958 Jorge Reyes-Magaña, Gerardo Sierra, Gemma Bel-Enguix, et al.

Metric	LexAN_MedGIL			
#nodes	14,895			
#edges	243,862			
Diameter	6			
Average degree	32.74			
Average clustering coefficient	0.429			
Barycenter	dolor			
Radius	3			

Table 2. Metrics of LexAN over MedGIL corpus

This decision was motivated by the requirement of having an organized corpus for evaluation purposes. It should be noted that GECO is capable of accommodating various types of text-based corpora.

3.2 Evaluation Corpus

It encompasses a collection of terms and corresponding definitions sourced from MedGIL. The definitions were acquired through a method of supervised identification, focusing on extracting definitional contexts [2].

In the evaluation phase of LexAN, we carefully curated a test corpus comprising 2,720 definitions and their corresponding terms.

Graph Construction

Having the two corpora stated before, we constructed the Lexical Association Networks. We executed a series of processes in the following sequence:

- 1. Lemmatization: We applied the Freeling tool [23] to perform lemmatization on both corpora.
- Multi-word Expressions: In cases where the terms identified consisted of multiple words, we merged the tokens using an underscore symbol(_). For instance, terms such as diabetes mellitus or hemorragia subconjuntival were treated in this manner. This process was applied to both corpora.

- 3. Stopwords Removal: We eliminated stopwords from both corpora using the stopword list in Spanish provided by the Natural Language Toolkit (NLTK) [7].
- Co-occurrence Graph Construction: We created a co-occurrence graph called MedGILCo using the main corpus. To achieve this, we used a Python library, text2graphapi⁶, configuring the parameters as outlined in Table 1.
- 5. Word Co-occurrence Norms (WCN): Based on the terms in the evaluation corpus, we generated a set of files known as "Words Co-occurrence Norms" (WCN).

These files represent a reinterpretation of the previously mentioned "Word Association Norms" collections. Each file, named as a term, in the WCN corpus, contains three columns:

- Response. Lists the neighboring words adjacent to the main term in the MedGILCo graph.
- Frequency (F). Indicates the frequency of co-occurrence between the term and the response, as derived from the edges in the MedGILCo graph.
- Association Strength (AS). Represents a normalization of the frequency:

$$AS = \frac{F}{\Sigma F}.$$
 (1)

6. Lexical Association Networks Construction: Finally, using the WCN corpus, we constructed the Lexical Association Network, a weighted undirected graph.

It is formally defined as: $G = \{V, E, \phi\}$ where:

- $V = \{v_i | i = 1, ..., n\}$ is a finite set of nodes of length $n, V \neq \emptyset$, that corresponds to the Terms and their responses.
- $E = \{(v_i, v_j) | v_i, v_j \in V, 1 \le i, j \le n\}$, is the set of edges.
- $\phi: E \to \mathbb{R}$, is a function over the weight of the edges.

⁶pypi.org/project/text2graphapi/

Precision	Window Size									
FIECISION	1	2	3	4	5	6	7	8	9	10
p@1	0.006	0.011	0.012	0.012	0.012	0.027	0.027	0.027	0.025	0.022
p@3	0.017	0.029	0.036	0.038	0.043	0.074	0.076	0.071	0.069	0.063
p@5	0.023	0.053	0.063	0.071	0.077	0.111	0.118	0.109	0.105	0.101

Table 3. Precision of the reverse dictionary using LexAN_F_MedGIL

The weight of the edges in the graph is determined by two different functions:

 Inverse Frequency (IF). The weight is calculated by taking the inverse proportion of the original frequency (F):

$$IF = \frac{1}{F}.$$
 (2)

 Inverse Association Strength (IAS). Similarly to the frequency weight, the inverse association is calculated for the weight:

$$IAS = 1 - AS. \tag{3}$$

The reason for recalculating the weights is to perform the execution of certain graph-based algorithms that rely on the geodesic distance.

Fig. 1 presents an example of a LexAN based on the MedGIL Corpus, generated using the above methodology with IAS as the edge weight. For visualization purposes, the node and edge labels are not shown.

Graph Metrics

Tab. 2 provides several metrics associated with the LexAN of the MedGIL Corpus without considering edge weights. The graph diameter of the LexAN_MedGIL is calculated to be 6.

This value represents the maximum shortest path length d(u, v) between any two vertices (u, v) in the graph, where d(u, v) denotes the graph distance.

The average degree of a graph is characterized as a graph invariant that corresponds to the arithmetic average of all individual vertex degrees in the graph [9], which is 32.74 for LexAN_MedGIL. The average clustering coefficient is determined to be 0.429. This value is obtained by calculating the mean of local clusterings in the graph.

The local clustering of a node measures the proportion of existing triangles among all possible triangles in its neighborhood.

To approximate the average clustering coefficient, the following experiment is repeated n times: a node is randomly selected, two neighbors are chosen randomly, and their connection is checked.

The approximate coefficient is the ratio of discovered triangles to the number of trials [29].

The barycenter, also known as the median [33], refers to the subgraph induced by the set of nodes v that minimizes the objective function:

$$\sum_{u \in V} d_G(u, v), \tag{4}$$

where d_G is the path length. In the case of the LexAN_MedGIL, this subset includes only the node "dolor" (pain).

The eccentricity of a node v is the maximum distance from v to any other node in the graph [13].

The radius is the minimum eccentricity in the graph, and the center represents the set of nodes with an eccentricity equal to the radius [28].

In the LexAN_MedGIL, the radius is 3. The periphery consists of nodes with an eccentricity equal to the diameter [6].

For reading purposes, the set of nodes related to the center, eccentricity, and periphery can be found on Github⁷.

Computación y Sistemas, Vol. 27, No. 4, 2023, pp. 955–963 doi: 10.13053/CyS-27-4-4774

⁷github.com/jocarema/LexAN

960 Jorge Reyes-Magaña, Gerardo Sierra, Gemma Bel-Enguix, et al.

						, ,				
Precision	Window Size									
	1	2	3	4	5	6	7	8	9	10
p@1	0.022	0.059	0.095	0.119	0.140	0.234	0.244	0.237	0.233	0.248
p@3	0.054	0.119	0.175	0.239	0.295	0.430	0.465	0.489	0.484	0.504
p@5	0.073	0.150	0.241	0.319	0.386	0.528	0.571	0.595	0.598	0.618

Table 4. Precision of the reverse dictionary using LexAN_AS_MedGIL

4 Experiments

4.1 Lexical Search

Reyes et al. [26] introduced a lexical search model for constructing a reverse dictionary. In their original model, they worked with a graph built with Word Association Norms.

Our study used the Words Co-occurrence Norms (WCN) to build two distinct lexical association networks (LexAN) based on the MedGIL corpus. These graphs are denoted as LexAN_AS_MedGIL and LexAN_F_MedGIL, respectively, representing association strength and frequency weights.

We employed the Evaluation Corpus described in Section 3 to replicate the process. Additionally, we adopted the precision at k (p@k) evaluation metric [19].

For instance, p@1 indicates that the concept associated with a given definition was ranked correctly in the first position, p@3 signifies that the concept appeared within the top three results, and the same principle applies to p@5.

Considering the size of the graph in terms of the number of nodes and edges, we decided to implement graph pruning, as discussed in [25].

This involved constructing a subgraph for each definition in the Evaluation Corpus, retaining only the neighbors at a distance of 1 in the corresponding LexAN.

The results obtained using LexAN_F_MedGIL are presented in Table 3, while Table 4 displays the results obtained using LexAN_AS_MedGIL.

In both cases, we employed various window sizes during the construction of the LexAN to examine the impact of window co-occurrence in the MedGIL Corpus on the reverse dictionary. Figure 2 displays the graphical representation of the values presented in Tables 3 and 4.

Based on these preliminary findings, it is evident that the results obtained using association strength as the weighting scheme are superior. Code is available in our Github - LexAN⁷ repository.

5 Discussion

Our findings align with the results reported by Reyes et al. [26], indicating that using association strength leads to higher performance in the reverse dictionary task. It is important to note that calculating the frequency is necessary to obtain association strength, which is considered a normalization of the frequency.

As depicted in Figure 2, we observe that larger window sizes are associated with improved search performance. However, it is important to consider that larger window sizes also increase the graph's dimension and the lexical search's computational complexity.

As mentioned earlier, graph pruning was necessary to limit the search space. The maximum window size used in our study was 10, as we believe this value balances search complexity and the graph's large size.

Concerning the parameters established during the construction of the co-ocurrence graph, the use of parallel execution produces faster processing during the co-ocurrence detection.

The preprocessing procedures were executed externally to co-ocurrence API with the aim of managing certain specialized elements (lemmatization, removing stopwords, etc.) during this particular phase.

The output data format was configured as networkx, enabling the execution of centrality algorithms within this designated library.



Fig. 2. Precisions using LexAN_F_MedGIL (left) and LexAN_AS_MedGIL (right)

The best performance achieved was a precision at 5 (p@5) value of 0.618 when employing association strength as the weight.

Notably, a precision at 3 (p@3) value of 0.504 was obtained, indicating that at least half of the 2720 definitions in the evaluation corpus were correctly identified within the first three potential concepts presented by the reverse dictionary. Regarding exact matches, 674 samples were correctly positioned in the first place.

6 Conclusion and Future Work

This paper presents Lexical Association Networks, using a methodology for constructing a graph from plain text, to help tackle the lexical access problem. The proposed methodology involves several more intricate stages than simply creating a co-occurrence graph.

Additionally, we introduce the concept of Word Co-occurrence Norms, which can be applied to any corpus using the techniques described in Section 3.

While our approach was applied to the Spanish language, specifically utilizing the medical corpus MedGIL, the construction of LexAN can be extended to any language with the availability of stop-word lists and a Term Extractor, which are commonly found in various languages.

In future work, we plan to explore additional techniques, such as reordering the outcomes through weighting mechanisms, to improve further the precision achieved in finding words related to the lexical access problem.

Moreover, we plan to extend the experiments to other specialized corpora and target terms obtained with different methods.

Acknowledgments

This work is funded by projects Conahcyt CF-2023-G-64 and PAPIIT IT100822.

References

- 1. Aitchison, J. (2012). Words in the mind: An introduction to the mental lexicon. John Wiley and Sons.
- 2. Alarcón, R., Sierra, G., Bach, C. (2007). Developing a definitional knowledge extraction system. Conference Proceedings of Third Language and Technology Conference.
- Arias-Trejo, N., Barrón-Martínez, J. B., López Alderete, R. H., Robles Aguirre, F. A. (2015). Corpus de normas de asociación de palabras para el español de México [NAP]. UNAM.

Computación y Sistemas, Vol. 27, No. 4, 2023, pp. 955–963 doi: 10.13053/CyS-27-4-4774

- 962 Jorge Reyes-Magaña, Gerardo Sierra, Gemma Bel-Enguix, et al.
 - 4. Arulselvan, A., Commander, C. W., Elefteriadou, L., Pardalos, P. M. (2009). Detecting critical nodes in sparse graphs. Computers and Operations Research, Vol. 36, No. 7, pp. 2193–2200. DOI: 10.1016/j.cor.2008.08.016.
 - 5. Bel-Enguix, G., Rapp, R., Zock, M. (2014). A graph-based approach for computing free word associations. Proceedings of the 9th edition of the Language Resources and Evaluation Conference, pp. 3027–3033.
 - 6. Bielak, H., Syslo, M. M. (1983). Peripheral vertices in graphs. Studia Scientiarum Mathematicarum Hungarica, Vol. 18, pp. 269–275.
 - Bird, S. (2006). NLTK: The natural language toolkit. Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions, pp. 69–72. DOI: 10.48550/ARXIV.CS/0205028.
 - 8. Cancho, R. F. I., Solé, R. V. (2001). The small-world of human language. Proceedings of the Royal Society of London, Series B: Biological Sciences, Vol. 268, No. 1482, pp. 2261–2265. DOI: 10.1098/rspb.2001.18 00.
 - Delen, S., Demirci, M., Cevik, A. S., Cangul, I. N. (2021). On omega index and average degree of graphs. Journal of Mathematics, Vol. 2021, pp. 1–5. DOI: 10.1155/2021/556 5146.
- **10. Ferret, O. (2002).** Using collocations for topic segmentation and link detection. Proceedings of the 19th International Conference on Computational Linguistics, pp. 260–266.
- **11. Ferret, O. (2006).** Building a network of topical relations from a corpus. Proceedings of the Fifth International Conference on Language Resources and Evaluation.
- 12. Ghosh, U., Jain, S., Soma, P. (2014). A two-stage approach for computing associative responses to a set of stimulus words. Proceedings of the 4th Workshop on Cognitive Aspects of the Lexicon and 25th International Conference on Computational Linguistics, pp. 15–21.

- Gupta, S., Singh, M., Madan, A. K. (2000). Connective eccentricity index: A novel topological descriptor for predicting biological activity. Journal of Molecular Graphics and Modelling, Vol. 18, No. 1, pp. 18–25. DOI: 10.1016/s1093-3263(00)00027-9.
- 14. Hernández, L. (2012). Creación semi-automática de la base de datos y mejora del motor de búsqueda de un diccionario onomasiológico. Universidad Nacional Autónoma de México.
- **15. Javaid, M. A. (2013).** Understanding Dijkstra algorithm. SSRN Electronic Journal. DOI: 10.2 139/ssrn.2340905.
- 16. Khrabrov, A., Cybenko, G. (2010). Discovering influence in communication networks using dynamic graph analysis. IEEE Second International Conference on Social Computing, pp. 288–294. DOI: 10.1109/socialcom.2010.48.
- 17. Kiss, G. R., Armstrong, C., Milroy, R., Piper, J. (1973). An associative thesaurus of English and its computer analysis. Edinburgh University Press.
- **18. Lafourcade, M., Joubert, A. (2015).** TOTAKI: A help for lexical access on the TOT problem. pp. 95–112. DOI: 10.1007/978-3-319-08043-7 _7.
- **19. Manning, C. D., Raghavan, P., Schutze, H.** (2009). Introduction to information retrieval. Cambridge University Press.
- **20. Mark-Roget, P. (1911).** Roget's thesaurus of english words and phrases: Classified and ARR. so as to facilitate the expression of ideas, and assist in literary composition.
- **21. Miller, G. A. (1995).** WordNet. Communications of the ACM, Vol. 38, No. 11, pp. 39–41. DOI: 10.1145/219717.219748.
- Nelson, D. L., McEvoy, C. L., Schreiber, T. A. (2004). The university of South Florida free association, rhyme, and word fragment norms. Behavior Research Methods, Instruments, and Computers, Vol. 36, No. 3, pp. 402–407. DOI: 10.3758/bf03195588.

- **23.** Padró, L., Stanilovsky, E. (2012). Freeling 3.0: Towards wider multilinguality. Proceedings of the Eighth International Conference on Language Resources and Evaluation, European Language Resources Association, pp. 2473–2479.
- 24. Qi, F., Zhang, L., Yang, Y., Liu, Z., Sun, M. (2020). WantWords: An open-source online reverse dictionary system. Proceedings of the Conference on Empirical Methods in Natural Language Processing: System Demonstrations, pp. 175–181. DOI: 10.18653 /v1/2020.emnlp-demos.23.
- 25. Reyes-Magaña, J., Bel-Enguix, G., Sierra, G., Gómez-Adorno, H. (2019). Designing an electronic reverse dictionary based on two word association norms of english language. Proceedings of Electronic Lexicography in the 21st Century Conference, pp. 865–880.
- 26. Reyes-Magaña, J., Bel-Enguix, G., Gómez-Adorno, H., Sierra, G. (2019). A lexical search model based on word association norms. Journal of Intelligent and Fuzzy Systems, Vol. 36, No. 5, pp. 4587–4597. DOI: 10.3233/jifs-179010.
- 27. Reyes-Magaña, J., Martínez, G. S., Bel-Enguix, G., Gómez-Adorno, H. (2020). Automatic word association norms (AWAN). Proceedings of the Workshop on the Cognitive Aspects of the Lexicon, pp. 142–153.
- 28. Roditty, L., Vassilevska Williams, V. (2013). Fast approximation algorithms for the diameter and radius of sparse graphs. Proceedings of the Forty-Fifth annual ACM symposium on

Theory of computing, pp. 515–524. DOI: 10.1 145/2488608.2488673.

- **29.** Schank, T., Wagner, D. (2005). Approximating clustering coefficient and transitivity. Journal of Graph Algorithms and Applications, Vol. 9, No. 2, pp. 265–275.
- **30.** Sierra, G., McNaught, J. (2000). Design of an onomasiological search system. Terminology, Vol. 6, No. 1, pp. 1–34. DOI: 10.1075/term.6.1 .02sie.
- Sierra, G., Solórzano-Soto, J., Curiel-Díaz, A. (2017). GECO, un gestor de corpus colaborativo basado en web. Linguamática, Vol. 9, No. 2, pp. 57–72.
- **32. Sowa, J. F. (1992).** Conceptual graphs as a universal knowledge representation. Computers & Mathematics with Applications, Vol. 23, No. 2, pp. 75–93. DOI: 10.1016/0898 -1221(92)90137-7.
- **33. Wilson, R. J. (2001).** Introduction to graph theory. Prentice hall Upper Saddle River.
- 34. Zock, M., Ferret, O., Schwab, D. (2010). Deliberate word access: An intuition, a roadmap and some preliminary empirical results. International Journal of Speech Technology, Vol. 13, No. 4, pp. 201–218. DOI: 10.1007/s10772-010-9078-9.
- Zock, M., Schwab, D., Rakotonanahary, N. (2010). Lexical access, a search-problem. Proceedings of the 2nd Workshop on Cognitive Aspects of the Lexicon, pp. 75–84.

Article received on 11/06/2023; accepted on 21/09/2023. Corresponding author is Gerardo Sierra.