# ELEXIS Transnational Research Visit Grant

## Final report – Ana Salgado, NOVA CLUNL

**Travel Grant:** Call 1
**Researcher:** Ana Salgado
**Hosting institution:** ELEXIS-ES: Real Academia Española (RAE, Spain), Centro de Estudios de la RAE y de la Asociación de Academias de la Lengua Española
**Institutional affiliation:** Faculdade de Ciências Sociais e Humanas da Universidade NOVA de Lisboa/Centro de Linguística da Universidade NOVA de Lisboa, Portugal
**Current position:** PhD Student in Terminology
**Project title:** Comparison of the macro/microstructure of the Iberian Academy Dictionaries: the *Diccionario de la Lengua Española* (DLE, RAE) and the *Dicionário da Língua portuguesa* (DLP, ACL)
**Period of stay:** November 11-30, 2018

## Introduction

Since I am the person in charge of the coordination of the new Portuguese Academy Dictionary, specifically as regards planning the macrostructure and microstructure of this dictionary, it is crucial that I learn how other structures are devised, such as the *Diccionario de la lengua española* (DLE) published by an analogous academy – *Real Academia Española*.

This research is aimed at defining guidelines for the inclusion and description of terms in general language dictionaries in order to help lexicographers in this specific task. Combining lexicographical and terminological methods is definitely a plus for the planning of a dictionary's macrostructure and microstructure, improving the organization and description of lexicographical articles as a whole.



RAE



ACL

## Research goals

Our main goal is to analyse the macro/microstructure of the DLE (RAE) and compare it with the Portuguese Dictionary (DLP) of the *Academia das Ciências de Lisboa* (ACL), especially to rethink the theoretical and methodological methods of the lexicographical tradition, such as the lexicographical labels that identify specialized lexicon.

Since my PhD project focuses on the lexicographical work methodology, there is great interest in getting to know and understand the methodology adopted by the team working at the *Instituto de Lexicografía* (ILex) of the RAE.

In addition to the methodology itself, it would also be highly relevant to exchange views on how scientific terminology is dealt with, discuss different approaches, exchange impressions on lexicographical works, get to know its infrastructures and computational linguistic tools, and share knowledge.

Placing the structures of two dictionaries – RAE vs. ACL – side by side and comparing them will aid in the decision-making process, namely by allowing us to:

- combine lexicographical/terminological methodologies;
- create a methodology for the selection, description/definition of linguistic and conceptual information (ISO TC37);
- improve the macro- and microstructure of lexicographical works;
- increase the quality of lexical databases.

## Methodological plan

- Establishing preparatory contacts with the host organisation
- Preparing an interview guide to become familiarized with the departments: ILex, Corpes, Tecnología (Desarollo, Lingüística Computacional)
- Meetings with experts
- Scheduling research interviews
- Analysing interviews
- Analysing infrastructures
- Working on a final presentation on the Portuguese Academy Dictionary

## Research questions

- How do lexicographers include terminological units (terms) in Iberian Academy Dictionaries?
- What is the percentage of terms in Iberian Academy Dictionaries?
- What do domain labels tell the user? Do they indicate a technical word? Are they useful?
- Which terms can you find in Iberian Academy Dictionaries? Are they all marked?
- What domains have already been labelled?
- Is it possible to standardize those labels?
- How can you organize conceptual information?
- What standards should be used? (TEI, Oxygen)

## First impressions

As our main goal was to get acquainted with the whole structure that supports the DLE, I only focused on the platforms and tools directly related to the DLE working process, although, of course, there are other features that have a direct influence on the dictionary. I spent a few days exploring the Entorno de Redacción Integrado (ERI), a software platform in JAVA and XML, in which lexicographical work is developed. Regarding the lexicographical work itself, I learned how

the dictionary is updated (documentation, academic proposals, commissions, external queries…). From Enclave, a new service platform, I would like to highlight one of its modules: the *Diccionario avanzado*. The DLE contains a lot of information hidden in its database because of print edition limitations. However, this module now enables combined searches and access the list of words related to a specific domain – e.g., when searching for the word *fútbol*, I was able to find all the definitions in which that word appears.

## First results

Having access to the DLE's underlying structure made it possible for us to exhaustively survey domain labels in order to determine the number of domains represented, those shared between the two dictionaries, the ones that differ, and which domains are the most frequent.

This study allowed us to analyse, describe, and compare domain labelling in general language dictionaries, which points to specialized lexicon in those general language dictionaries. We have examined whether label application is systematic or there is an imbalance, and whether there are recent and relevant domains that do not appear. We verified the criteria for differential selection and treatment of terms, which differ between the dictionaries under review. Finally, we looked at other lexicographical resources used in the definitions as domain indications.

Stemming from this work, in particular from the topic of domain labelling, we are pleased to announce the acceptance of a submitted paper (peer reviewed) to the *III Jornadas internacionales sobre investigaciones lexicográficas y lexicológicas* (inLÉXICO2019), which will be held at the *Universidad de Jaén*, on April 4-5 2019. (Salgado, A. & Costa, R., *Marcas temáticas nos dicionários académicos ibéricos: estudo comparativo*).

## Conclusions and future work

It has been very beneficial to understand how the DLE is updated since it is a more complex process than what is done at the ACL. The RAE has academic sessions and different committees to approve any changes. On top of that, the RAE also interacts with the *Asociación de Academias de la Lengua Española* (ASALE), present all over the Hispanic world.

It is our intention to propose an agreement between academies leading to a systematic labelling of the specialized use of a particular entry/meaning and its representation.

## Acknowledgments

I want to thank the ELEXIS project for the opportunity given to me. In the digital age, common standards are highly needed to create structured and interoperable lexical databases that will optimize the lexicographical work. I would like to thank Jordi Porta and Elena Zamora. And also Paz Battaner, who gently came to assist me in my final presentation.