TNA ACTIVITY REPORT

Metaphors of Crisis – Metaphorical Patterns Across Discourses in Literature and Newspapers.

Author: Dario Del Fante

Current position: Assistant Professor – Researcher

Affiliation: University of Ferrara – Italy

Contact: dario.delfante@unife.it – dariodelfante41291@gmail.com

Host institution: Laboratorio de Innovaciòn en Humanidades Digitales – UNED – Madrid - Spain

Mentor: Prof. Salvador Ros

Period of stay: 2nd September 2022 - 7th October 2022

Introduction

I applied for the TNA fellowship to give a decisive boost to my project on the study of figurative language. Given the necessity of improving the methodology for metaphor detection and extraction from large datasets and the lack of computational and technical support from the department where I was working at the time of the application, the CLS-INFRA project seemed to be a suitable opportunity for my career.

Objectives

The project had two main aims:

1. to improve my research by extending my knowledge in programming to develop an automatic method for metaphor identification that uses automatic computational techniques: e.g., LDA Topic Modelling, BERT topic modeling, or different computational approaches.
2. to build an annotated corpus (FAIR resource if possible) made of samples of parliamentary, literary, and newspaper texts which can be served as:
    a. test for the automatic method for metaphor identification by launching some experiments over models like Bert.
    b. to explore the hypothesis about the literary origin of 'usual' metaphorical mappings.

Detailed report on the stay

I stayed in Madrid from 2nd September to 7th October 2020. I have been working at Laboratorio de Innovaciòn en Humanidades Digitales LINEDH, which is located within the

Central Library Building of the UNED university. The following sections represent a detailed description of my visiting period. I was supervised by Professor Salvador Ros.

First week – 2nd September – 8th September

I first met Prof. Salvador Ros who introduced me to the LINEDH laboratory and to the Central Library facilities. We had two meetings in which we got familiar with the other' background and research activity. The third meeting has been then devoted to the definition of a research plan for the stay. We ended up deciding that the main topic during the visiting period would have concerned the development of an automatic approach for metaphor detection and extraction. I spent the rest of the week conducting a preliminary literature review on this specific topic.

Second Week – 9th September – 15th September

Firstly, I took part in the audience at the International Conference "Libraries & Digital Humanities: Projects and Challenges" which took place between 13th September and 16th September. After the first round of Literature Review and another meeting with Prof. Ros, we decided to use a BERT-based algorithm to automatically recognize and extract metaphors. We decided to restrict the scope of the research and embark on a case study regarding the metaphors of migration and the metaphors of coronavirus. The aim was to train a BERT model at detecting metaphors and at recognizing between coronavirus-related metaphors and migration-related metaphors.

Third Week – 16th September – 22nd September

I was involved in the compilation of two datasets. The first one contains parliamentary debates between 2015-2016 and 2020-2021, retrieved from the Parlamint project[1]. The second contains newspaper texts published between 2015-16 and 2020-2021. Both datasets are in English and Italian. I annotated the dataset according to three values: the presence of a migration-related metaphor; the presence of a coronavirus-related metaphor; absence of metaphorical language. Some preliminary tests were performed on these datasets.

Fourth Week – 23rd September – 30th September

The first preliminary results showed that the Bert-algorithm detects and recognizes metaphors in texts. However, there was a high percentage of false positives. For this reason, we decided to enlarge as much as possible the size of the dataset and then train the algorithm. Additional preliminary tests confirmed the previous results. However, the percentage of false positives was still high.

Fifth Week – 1st October – 7th October

We have scheduled the further steps of this research. We attended an important meeting with the Linguistics research group at the Department of Philology at UNED. I presented my research and the small project developed with Prof. Salvador Ross. We critically reviewed the methodology with the help of a variety of experts from the research group at UNED. We then

---

decided to collaborate on the development of the dataset with the addition of a Spanish section.

Concluding remarks and further research

The visit has been extremely useful for my future work on developing an automatic methodology for metaphor detection and interpretation in English and Italian.

Despite it needs great improvements, the results fulfill my expectations. Further research will be concerned with the enlargement of the dataset and the addition of the Spanish language. In this sense, the project's cross-cultural nature will be enforced. We plan to present the results of the project to several conferences over the following year and to continue the collaboration, which might be concretized into the publication of a couple of articles.

Another idea regards the application of alternative methodologies to the same dataset such as machine learning and to compare the results.

Acknowledgements