

# Multilingual Domain Ontologies of Rhetorical Figures and Their Applications

**Ramona Kühn**

University of Passau  
Germany

ramona.kuehn@uni-passau.de

**Jelena Mitrović**

University of Passau  
Germany

jelena.mitrovic@uni-passau.de

*Relevant UniDive working groups:* WG1, WG3

## 1 Introduction

Rhetorical figures are linguistic and rhetorical constructs that add flavor to language, making it more interesting, memorable, or persuasive. Common rhetorical figures include metaphors, similes (comparisons), hyperbole (exaggeration), and alliteration (consecutive words beginning with the same sound). Rhetorical figures often convey an implicit meaning which can only be understood if the readers or listeners make the correct inferences based on their background knowledge and experiences.

As figures are often ambiguous, it is challenging for computers to understand the intended meaning, leading to a literal interpretation of rhetorical figures. This becomes obvious in tasks of natural language processing (NLP) such as machine translation, hate speech detection, text summarization, or sentiment analysis: For example, the English expression “to get your feet under the table”<sup>1</sup> is often literally translated by modern translation systems. Considering rhetorical figures already led to an improvement in hate speech detection (Lemmens et al., 2021), text summarization (Alliheedi and Di Marco, 2014), and argument mining Mitrović et al. (2017), where the authors included characteristics of metaphors or figures of repetition, and showed the use of rhetorical concepts in arguments.

Still, there is a lack of formal systems that can describe rhetorical figures, their form and function, to enable computational detection or assist in the annotation process. As some rhetorical figures are not used too often, existing datasets are too small to reliably train machine learning algorithms to detect them. Dubremetz and Nivre (2017) encountered this problem when they tried to detect the figure chiasmus. We want to fill this gap by formalizing rhetorical figures. Furthermore, we do not only focus on the English language. As most NLP research is performed in English, which leads to inequalities (Søgaard, 2022), we want to promote

<sup>1</sup>Meaning: To establish yourself in a new job or situation.

the creation of ontologies of rhetorical figures in other languages to decrease this imbalance.

Our contributions are the following:

- We created formal domain ontologies of rhetorical figures in different languages.
- We present and compare existing ontologies in the domain of rhetorical figures.
- We identify their properties, especially in regard to the language they are modeled in (English, Serbian, German).
- We present methods and applications of how those ontologies can support the annotation of data.

## 2 Ontologies of Rhetorical Figures

An ontology is a formal representation of knowledge. In the case of domain ontologies, knowledge from a specific domain is represented. In our case, it is the domain of rhetorical figures. An ontology helps formalize and unify varying definitions. The ontologies contain classes and individuals which are connected by properties in a triplet relation, forming a knowledge graph. SPARQL is a query language to extract this knowledge from the ontology. In the following, we will present several ontologies of rhetorical figures in different languages and how they can be used for various NLP tasks.

The English **RhetFig** (Kelly et al., 2010) project was the first one to describe theoretical concepts of an ontology of rhetorical figures. They focused on grouping figures into linguistic and rhetorical classes. The Serbian **RetFig ontology** (Mladenović and Mitrović, 2013) implemented those concepts practically, describing the most common rhetorical figures in the Serbian language. They reused the structure that was suggested by the RhetFig project, as reusing is considered good ontological practice (Allemang and Hendler, 2011). The structure of RhetFig is also reused in the English **Ploke ontology** (Wang et al., 2021), an ontology that formalizes figures of perfect lexical repetition

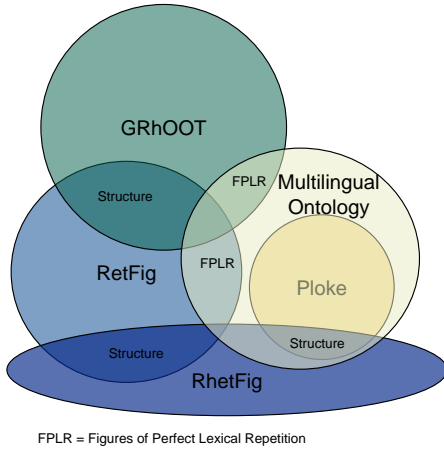


Figure 1: Connection of the different ontologies.

and includes cognitive affinities, e.g., that a certain figure is more memorable. We adapted the Serbian RetFig ontology to German and created the **GRhOOT ontology** (Kühn et al., 2022).

It is not only a translation of the Serbian ontology: We examined for each figure if a counterpart in the other language exists and if it has the same properties. We also identified figures that do not exist in Serbian. In addition, we included example sentences for each figure. Currently, the GRhOOT ontology contains the formal description of 110 rhetorical figures in the German language.

While the Serbian and German ones focus on the most known figures and their description, the Ploke ontology focuses on cognitive aspects. To unite all these aspects, we created the **multilingual ontology** (Wang et al., 2022) combining figures of perfect lexical repetition of the Serbian, German, and English Ploke ontology, which served as the basis. Fig. 1 shows the connection between the described ontologies.

The multilingual ontology benefits from each ontology it is based on: The formal description of RetFig and GRhOOT, and the cognitive aspects of the Ploke ontology. Furthermore, we analyzed differences in figures in the three languages, e.g., Serbian figures that do not exist in German, or the same name denoting different figures in other languages. Figure 2 illustrates the classes and relations for the figure epizeuxis in the multilingual ontology. It is an immediate repetition that triggers a neurocognitive affinity.

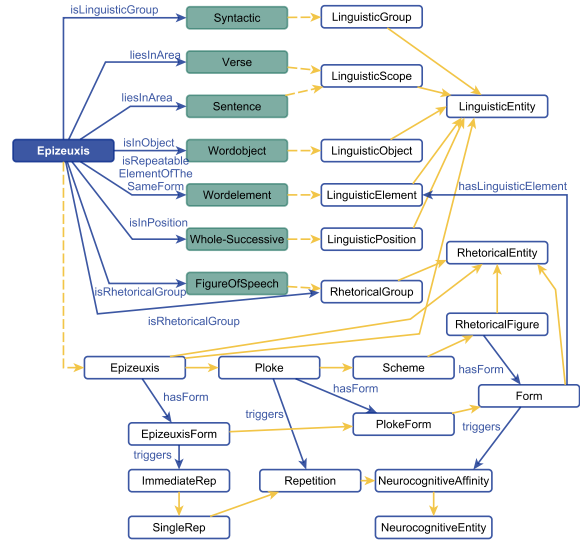


Figure 2: Multilingual ontology, adapted from Wang et al. (2022).

### 3 Applications for Rhetorical Ontologies

In this section, we want to show different applications and practical use cases for ontologies of rhetorical figures. First of all, the presented ontologies are a collection of knowledge from different sources, processed and formalized, providing a clear overview of the topic and the different categories. The users can utilize the ontologies as reference work and browse them to determine rhetorical figures and to find out their intention in a text.

We built a decision tree within the scope of the GRhOOT ontology, allowing the user to identify rhetorical figures of repetition. Furthermore, the ontologies can be queried with SPARQL queries. Knowledgeable users can formulate queries searching for figures fulfilling specific properties, e.g., when searching for figures with a repetition of a word, the respective query looks like this:

```
SELECT distinct ?Figure
WHERE {
  ?Figure ontology:IsRepeated ?Element .
  ?Element rdfs:label ?ElementName .
  Filter (?ElementName = "Word" )
}
```

However, this approach requires knowledge of SPARQL and the properties and classes of the ontology which can only be achieved with extensive training. The ultimate goal is to make the ontologies accessible to users without prior knowledge of ontologies or query languages.

A graphical user interface that guides users through the process of identifying figures can help here: Our current work focuses on the implementation of such an application. Users that suspect a rhetorical figure in a sentence but cannot name it can select the salient properties in the graphical interface, e.g., repetition of a word in the beginning and end of a sentence. The properties that the users can select are based on the properties of the ontologies. No knowledge about the ontologies or their structures is required. The selected properties are then automatically translated into SPARQL queries and executed on the ontology, which delivers the names of matching rhetorical figures, here “symploke”. The users can then annotate the sentence. This process does not require any special training and the annotation of datasets can be performed without linguistic experts. If more annotated datasets exist, the training of machine and deep learning algorithms becomes more reliable, which again leads to an improvement of various NLP tasks, as described in Section 1.

## 4 Conclusion

We presented and compared ontologies of rhetorical figures in different languages, namely Serbian, German, and English, and described how they benefit from each other. Their advantage is not only the formalization of rhetorical figures but also their ability to support users in an annotation task. As mentioned, the performance of various NLP tasks can be increased when rhetorical figures are considered. In the future, we will focus on further development of the graphical application to support the annotation of rhetorical figures. Furthermore, we want to create similar ontologies in other languages.

## Acknowledgments

The project on which this report is based was funded by the German Federal Ministry of Education and Research (BMBF) under the funding code 01IS20049. The author is responsible for the content of this publication.



## References

- Dean Allemang and James Hendler. 2011. *Semantic web for the working ontologist: effective modeling in RDFS and OWL*. Elsevier.
- Mohammed Alliheedi and Chrysanne Di Marco. 2014. Rhetorical figuration as a metric in text summarization. In *Advances in Artificial Intelligence: 27th Canadian Conference on Artificial Intelligence, Canadian AI 2014, Montréal, QC, Canada, May 6-9, 2014. Proceedings 27*, pages 13–22. Springer.
- Marie Dubremetz and Joakim Nivre. 2017. Machine learning for rhetorical figure detection: More chiasmus with less annotation. In *Proceedings of the 21st Nordic Conference on Computational Linguistics*, pages 37–45.
- Ashley R Kelly, Nike A Abbott, Randy Allen Harris, Chrysanne DiMarco, and David R Cheriton. 2010. Toward an ontology of rhetorical figures. In *Proceedings of the 28th ACM International Conference on Design of Communication*, pages 123–130.
- Ramona Kühn, Jelena Mitrović, and Michael Granitzer. 2022. **GRhOOT: Ontology of rhetorical figures in German**. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 4001–4010, Marseille, France. European Language Resources Association.
- Jens Lemmens, Ilija Markov, and Walter Daelemans. 2021. Improving hate speech type and target detection with hateful metaphor features. In *Proceedings of the Fourth Workshop on NLP for Internet Freedom: Censorship, Disinformation, and Propaganda*, pages 7–16.
- Jelena Mitrović, Cliff O’Reilly, Miljana Mladenović, and Siegfried Handschuh. 2017. Ontological representations of rhetorical figures for argument mining. *Argument & Computation*, 8(3):267–287.
- Miljana Mladenović and Jelena Mitrović. 2013. Ontology of rhetorical figures for serbian. In *Text, Speech, and Dialogue: 16th International Conference, TSD 2013, Pilsen, Czech Republic, September 1-5, 2013. Proceedings 16*, pages 386–393. Springer.
- Anders Søgaard. 2022. Should we ban english nlp for a year? In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 5254–5260.
- Yetian Wang, Randy Allen Harris, and Daniel M Berry. 2021. An ontology for plope: Rhetorical figures of lexical repetitions.
- Yetian Wang, Ramona Kühn, Randy Harris, Jelena Mitrović, and Michael Granitzer. 2022. Towards a unified multilingual ontology for rhetorical figures. In *14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - KEOD*, pages 117–127.