

individual relation. Furthermore the shared typology can be used as a conceptual framework for in-depth comparison between the different meaning relations. It also greatly facilitates the transfer of knowledge and resources between paraphrasing, textual entailment, and semantic similarity research.

9.2 Resources

Throughout my research, I have created several language resources that I have made available to the Natural Language Processing and Computational Linguistics community. These resources are also part of the contributions of this thesis:

- The Extended Paraphrase Typology (EPT) and annotation guidelines with examples for creating a corpus annotated with EPT.
- The Single Human-Interpretable Typology for Annotating Meaning Relations (SHARel) and annotation guidelines with examples for creating a corpus annotated with SHARel.
- The WARP-Text web-based annotation interface for a fine-grained annotation of pairs of text.
- The Extended Typology Paraphrase Corpus (ETPC) - the first Paraphrase Identification corpus annotated with paraphrase types.
- The first corpus explicitly annotated with Paraphrasing, Textual Entailment, Contradiction, Semantic Textual Similarity, and Textual Specificity.
- The first corpus of multiple meaning relations, annotated with SHARel.

During my research and in collaboration with the Language Technology Lab at the University of Duisburg-Essen, I co-organized the first RELATIONS workshop¹ [Kovatchev et al., 2019a], bringing together researchers on textual meaning relations. The workshop was collocated with the 13th International Conference on Computational Semantics (IWCS) in Gothenburg, Sweden, May 23 2019.

9.3 Future Research Directions

This thesis gives two new perspectives on the research of textual meaning relations within Natural Language Processing and Computational Linguistics.

¹<https://sites.google.com/view/relations-2019>

- The importance of linguistic knowledge in the automatic processing of meaning relations.
- The presentation of the first empirical research on multiple textual meaning relations.

My work opens several new directions for future research.

In Part II of this thesis I applied paraphrase typology to the task of Paraphrase Identification. One of my findings was that the existing corpora for Paraphrase Identification is not well balanced in terms of paraphrase types. I argued that this imbalance can introduce a bias in the task and decrease the generalizability of automated PI systems. A promising research direction in this area is to work towards creating better corpora for the recognition tasks on textual meaning relations. Following what I did in my dissertation, I am currently working on a new corpus for Recognizing Textual Entailment for Spanish. The objective behind the corpus creation is to obtain a corpus more balanced in terms of certain under-represented phenomena, such as negation and named entity reasoning. My work would also help expand the research on textual meaning relations to languages other than English.

In Chapter 6 I presented the advantages of a qualitative evaluation framework over traditional measures such as Accuracy and F1. An important future research in this area would be to extend the qualitative evaluation framework to other empirical tasks on textual meaning relations, such as recognizing textual entailment or semantic textual similarity. My work in Part III of this thesis and in particular the SHARel typology is a first step towards extending the coverage of the qualitative evaluation. I showed that the typology can be applied to multiple textual meaning relations. The next step in this research direction would be the creation of a larger corpus and the corresponding software package for a general qualitative evaluation framework for textual meaning relations.

The qualitative evaluation of Paraphrase Identification systems showed that some phenomena, such as negation, ellipsis, and named entity reasoning are challenging across all evaluated systems. As a future research, I believe that each of these phenomena has to be analyzed in more detail in the context of the importance it has for the automatic processing of textual meaning relations. In a continuation of my thesis, I am currently investigating the role of negation for paraphrase identification, recognizing textual entailment, semantic textual similarity, and question answering. The preliminary results indicate that negation is extremely challenging across multiple automated systems.

Finally, my work in Part III of this thesis opens several new directions on the joint research of multiple textual meaning relations. One potential area is the creation of datasets and automated systems for the simultaneous processing of

multiple textual meaning relations. Some preliminary experiments that I carried out on the corpus presented in Chapter 7 indicate it is possible to use one meaning relation to predict the others. The next step in this research direction would be the creation of larger datasets with multiple textual meaning relations and the creation of more sophisticated automated systems.