

## Preface

This slender volume contains the papers that were accepted for publication at the IWCS Workshop on Models for Modality Annotation (MOMA 2015), organised as a satellite event of the 11th International Conference on Computational Semantics (IWCS 2015) at the Queen Mary University of London, on April 14 2015.

The notion of modality involves a spectrum of phenomena that are pervasive in language but still far from being formalised. For an exhaustive formalisation, a joint effort by computational, corpus, and formal linguists as well as language typologists is required.

Computationally, the automatic identification and interpretation of modalised statements is a prime concern in a large number of applications, especially with the recent attention to opinion mining and social networks. Indeed, recent years have witnessed the development of annotation schemes and annotated corpora for different aspects of modality in different languages. While there have been efforts towards finding a common avenue for modality annotation, (the CoNLL-2010 Shared Task, ACL thematic workshops and a special issue of Computational Linguistics), the computational linguistics community is still far from having developed working, shared standards for converting modality-related issues into annotation categories.

In corpus linguistics studies of modality-related phenomena, researchers use an incremental method based on redefining categories after assessing agreement through several rounds of manual annotation, with the aim of finding the right balance between feasibility and expressivity of categories.

Formally, and from a comparative linguistics perspective, characterisations are sought of the range of modal types and their marking across the languages of the world, towards a complete classification of modal functions. This would yield a thorough understanding of the relations holding between modal categories, and an understanding of the grammatical vs. lexical nature of modal markers across languages. Insights from this tradition are crucial for the advancement of computational work on modality, since a comprehensive scheme for producing reliable annotated data must obviously be usable from a computational perspective, but it also has to rely on a solid theoretical base. In other words, a balance must be found between accuracy and detailing in the description of the phenomenon, and preventing proliferation of labels which might cause data to be too sparse to learn from, and also lower agreement among annotators.

Considered all of the above, the main aim behind the organisation of this workshop was bringing together researchers from the involved fields to join efforts in defining exhaustive and at the same time usable representations of modality, towards working, implementable annotation standards. Albeit small in number, the contributions that are published in this volume do indeed cover the topics we intended to touch upon, namely a general model of modality for a given language (Portuguese in this case), issue related to annotating specific modality phenomena, such as epistemic phrases in informal language, and, from a practical point of view, developing tools to support modality annotation, as *GraphAnno*. In addition to contributed papers, the workshop also features an invited talk by James Pustejovsky (Brandeis University) on reference and point of view, a round table discussing modality issues from several perspectives, including opinion mining, and a session devoted to actual planning future strategies towards shared standards in the annotation of modality.

The organisers

*Malvina Nissim*

*Paola Pietrandrea*