Classical Philology Goes Digital. Working on Textual Phenomena of Ancient Texts
Universität Potsdam, February 16-17, 2017

# The Syntax of Similes
## A Treebank-Based Exploration of Simile in Greek Poetry

Francesco Mambrini
(DAI, Berlin)

In this presentation we explore whether treebank annotation can be used to identify and study similes in Ancient Greek poetry. We use data from the last version (2.1) of the Ancient Greek Dependency Treebank (AGDT) published by the Perseus Project https://perseusdl.github.io/treebank_data/), and we focus our attention on the Iliad.

This research program entails the fascinating question of how to use morphosyntactic information to study semantic and discursive features of literary texts. We discuss here the methodological aspects involved, by focusing in particular on the main linguistic and technical challenges.

Some structural aspects of the AGDT that may affect the study of Homeric similes include: the presence of possible multiple interpretations on any given passage, inconsistencies among the annotators in treating the same syntactic constructions, or the presence of annotation errors. More in general, however, the research is complicated by the fact that similes may be expressed by many different syntactic constructions; moreover, the notations that are used in the AGDT guidelines for each of these syntactic structures (such as ὡς governing a verb) are often polysemic, which means that they are used also to encode other linguistic phenomena unrelated to similes. Finally, while sentences in syntactic annotation are treated as discrete units, several long similes in the Iliad build a complex narrative that spans over multiple sentences.

On the other hand, treebanks are invaluable tools to retrieve specific constructions from corpora, get quantitative data on those syntactic phenomena, also allowing for powerful and informative visualization. But in particular, such treebank-based enquiries provide an excellent occasion to test the soundness, consistency and usefulness of an annotation schema, as well as to address errors and problems.