Contemporary research on late antiquity is gradually transforming our understanding of the cultural and religious developments of this period. Archaeological (e.g. Lavan 2011) and literary studies (e.g. Cameron 2006) have recently shown that in order to appreciate the complex interactions of the late Roman world, we need to deconstruct a long tradition of polemical writing practiced by Christians and other biased groups, and take a fresh look at the primary evidence. As a result, the old conception of a conflict between early Christianity and the classical antiquity is giving way to a more accurate view of a sophisticated dialogue between Christianity, other ancient religions and the classical tradition.

My research project explores these interactions in a corpus of Greek poetry written between c. 300 and 700 CE. With Shorrock (2011), I consider texts with both ”pagan” (denoting all religions other than Judaism and Christianity) and Christian subject matter to spring inherently from the same classical roots, intersecting at several points with each other as well as their classical predecessors. Following Shorrock’s characterization, I draw the distinction between explicitly Christian and non-Christian modes of writing, without making any assumptions about the religious beliefs of the authors. My aim is to apply statistical tools to examine the relations between aesthetic ideals and the different modes of writing, using an array of data discussed below. The primary focus is on classical intertextuality, here understood in a broad sense comprising both intentional and unintentional relations between late ancient poetry and its classical models. With this approach, I wish to be able to contribute to the recent discussion (e.g. Agosti 2013) about the relationship between religion and classicism.

Poetic discourse underwent a renaissance in late antiquity, becoming the most significant form of literary communication in the highest spheres of society. This can be illustrated by the huge poetic output of Gregory of Nazianzus (c. 18 000 lines, late 4th century CE), containing a wide range of subjects from theology to biography. The high status of poetry in late antiquity is also apparent in its official applications: it was typically chosen over prose for such occasions as monumental dedications, panegyrics and Christian congregations. This, together with the fact that a relatively large amounts have survived, makes Greek poetry an interesting object of study in view of the discussion outlined above.

The digital corpus of Greek poetry is provided by the Thesaurus Linguae Graecae project (http://tlg.uci.edu), which has made available most surviving literary Greek texts from Homer until late antiquity and well beyond. However, due to a strict
copyright license, no text materials owned by the TLG may be copied into an intermediary database, nor may its own databases be queried via a remote script for a large-scale analysis. One of the methodological advances I propose in this project is a new way to leverage TLG’s resources. In my approach, data collection happens client-side through TLG’s web interface and includes the collection of morphological analyses, lemmas, word/sentence boundaries as well as metrical and phonetic data. These data, in virtue of excluding the original text, may legally be saved to a database to be used in the actual analysis, opening up new possibilities for a quantitative study of TLG’s materials.

Methodologically, my project draws from recent developments in using quantitative models to automatically detect intertextuality (e.g. Coffee et al. 2013). Most existing methods have a full-text corpus at hand, whereas this project (for the reasons discussed above) is constrained to a relatively limited collection of data. In many cases, however, an algorithm can be modified to suit different needs. For example, Forstall et al. (2011) have proposed a "functional n-gram" feature to detect influence using character-level n-grams; but as the authors point out, the same approach can be used with other feature types as well. My aim is to employ a similar method with the data that can be extracted from a TLG text, such as metrical patterns, word boundaries and phonetic classes.

This poster will present new tools to help collecting a number of data from a poetic text provided by the TLG. These include an algorithm which computes the most probable metrical reading of a given line, an integrated morphological analysis tool to speed up the collection of lemmata, as well a front-end to manage the collected datasets in a Text Encoding Initiative (TEI)-compliant XML format.

References


