

Citation Practices in Jerome's Letters as *vestigia* of Late Antique Identity Construction

Marie Revellio
(Universität Konstanz)

The church father St. Jerome (ca. 347-420) – an intermediary between the different cultures of the Greek and increasingly Christian influenced East of the Roman Empire and the still more pagan Latin West – formed the pointed phrase: *Ciceronianus es, non Christianus* (Jer. ep. 22,30,4). This striking formulation expresses the immanent strained relationship of classical education and late antique Christianity. The line of distinction in this dichotomous definition of identity follows explicitly the reading of literature. As Jerome frequently embeds references from classical-pagan as well as Christian texts into his own work, his texts are predestined to entangle intertextuality research with questions of group and identity construction.

The project deals with intertextual references as literary stages to discuss questions concerning cultural hybridity. It examines verbal text re-use through the combination of computational analysis and hermeneutic close-reading.

Starting with Jerome's letters, reused text passages of classical-pagan authors will be traced automatically (e.g. of the in these times still canonical Vergil). Therefore, the whole corpus of letters has to be properly digitized. This step has been done in cooperation with the *Open Greek and Latin Project* at the University of Leipzig. Within this project, the standard printed text version of the letters (CSEL) was already digitized in TEI-XML, but the electronic text needed some manual proofreading and emendation.

The detection of re-used text passages is conducted with the *Tesserae*-Project. In order to get the best possible impression of verbal intertextuality and to be on the same time able to compare the findings with traditional commentaries, a very unrestricted *Tesserae* search is run. This proceeding has the advantage that as many results as possible are found. But for further discrimination between valuable (in sense of meaningful) instances of text reuses and only coincidentally lexical matches additional analyses are needed. This methodical problem between *precision* and *recall* is addressed through adequate filter setting via a Python implementation. The training of the classification system of meaningful and not-meaningful results follows three steps: calibration, evaluation and use. Because not all meaningful results are known by start, the correct calibration of the classification system has to deal with the lack of an exact training benchmark. Therefore, two optimization rules are chosen: firstly, do not fall behind the status quo of research and secondly, provide new results while close-reading is still feasible. The filter criteria are derived from the implicit applied techniques of traditional commentaries.

The evolved filter setting generates a quite acceptable number of estimated positive results. This combination of digital and non-digital methods can help to further explicit the theoretical questions entangled to meaningfulness in respect of re-used texts but also serves as a starting point for a methodical reflection on both methods.