A Digital Database of Manuscripts and Language Use in Post-Conquest England

Dr. Krista A. Murchison, Leiden University

In the traditional linguistic model of medieval England, the Norman Conquest of 1066 caused English, which previously had been an acceptable language of high society, to be sidelined in aristocratic and courtly domains and replaced by French, only regaining its status after about two hundred and fifty years. A growing body of research has pointed to the significant structural problems with this traditional linguistic model, and it is now generally accepted that French persisted as an important domestic and aristocratic language for centuries after the Conquest (Stein 2007; Butterfield 2009; Waters 2015). In light of this increasingly important body of research, the status of French in this period, including the contexts and implications of its use, is being re-examined (Ormrod 2003; Baswell 2007; Watson 2009; Waters 2015).

The project explored in this short presentation, funded by a Europeana Research Grant and produced with support from Leiden University's Centre for Digital Scholarship, takes a new approach to exploring the status of French writing following the Conquest. In particular, it will create and analyze a digital database of all c. 958 manuscripts containing French literature from medieval England. Manuscripts have been chosen as the focal point of this project because they provide unmatched large-scale data about language use during the post-Conquest period.

The first stage of the project is centered on the compilation of information about the distribution of languages in each manuscript. Since this information is being drawn from a variety of sources, including manuscripts and catalogues, it is being gathered manually. In the case of Cambridge University Library, MS Ee.1.20, for example, I have used both catalogue data and in-person manuscript consultation to determine the start and end points of the texts and the languages used in the manuscript (French, in this case). Then, this information, along with provenance information about each manuscript, will be marked up in Oxygen XML editor. In keeping with the FAIR data principles, the compiled data will be hosted online and made available to others through an online catalogue. The statistical analysis of this database, which will take place in the final phase of the project, will seek to identify and analyze the distribution of languages in each manuscript, the distribution of multilingual and monolingual manuscripts, and how these statistical distributions differ between time periods, between male and female audiences, and between lay and clerical audiences.

The digital approach used in this investigation will lead to the first large-scale quantitative information about the distribution of French writing across medieval society. By providing unprecedented quantitative evidence for the persistence of French among lay and clerical audiences alike in the centuries following the Conquest, this project will contribute to an urgent and growing body of linguistic scholarship (Lusignan 2011; Ingham 2012; Wogan-Browne 2015) and socio-historical studies (Rothwell 1994; Georgianna 1998; Butterfield 2009; Ingham 2010) that has exposed factual and ethical flaws in the 'triumph' of English narrative and highlight medieval England's multifaceted intercultural exchanges.