

## Innovators of Early Modern English spelling change: Using DICER to investigate spelling variation trends

Through the use of new tools able to deal with the inherent spelling variation observed in Early Modern English (EModE) (Görlach, 1991; Nevalainen, 2006), recent research has allowed for the large-scale analysis of EModE corpora, quantifying the high levels, and gradual standardisation, of spelling during the period (Baron et al., 2009a). Whilst a decline over time in the quantity of spelling variants present was observed in all of the corpora studied, the levels of variation differed considerably between corpora at concurrent time-points.

Here, we investigate these differences by exploring spelling variant patterns uncovered by a web-based tool, DICER (Baron et al., 2009b). DICER analyzes spelling variant standardisations to discover and quantify the presence of character edit rules. These rules represent spelling decisions made by authors, scribes, editors and publishing houses. Some of the rules revealed are well defined in descriptions of EModE language (e.g. Fisher, 1977), such as the inconsistency of a final <e> and the interchanging of <u>/<v> and <i>/<y>. Whilst other rules revealed are less well known, such as the doubling and singling of letters, such as <l> and <t>, plus the insertion, substitution and deletion of particular vowels in specific contexts.

The aim of our study is to quantify these spelling patterns over time, genres and text types. The use of this analysis will be two-fold. Firstly, it will assist in the underlying development of an historical spelling standardisation tool (see Baron & Rayson, 2009), with custom rule sets being made available depending on the meta-data of the text being standardised. Secondly, it will aid in the investigation of differences in spelling variant levels between corpora. Ultimately, we wish to reveal when spelling patterns first appear and when they are no longer present and also uncover if any groups are ahead of the trend in using (or not using) specific patterns, i.e. the innovators of spelling change.

### References

Baron, A. and Rayson, P. (2009). Automatic standardization of texts containing spelling variation, how much training data do you need? In M. Mahlberg, V. González-Díaz and C. Smith (eds.) *Proceedings of the Corpus Linguistics Conference, CL2009*, University of Liverpool, UK, 20-23 July 2009.

Baron, A., Rayson, P. and Archer, D. (2009a). Word frequency and key word statistics in historical corpus linguistics. In *Anglistik: International Journal of English Studies*, 20 (1), pp. 41-67.

Baron, A., P. Rayson, and D. Archer. (2009b). Automatic standardization of spelling for historical text mining. In *Proceedings of Digital Humanities 2009*, Maryland, USA. University of Maryland. pp. 309–312.

Fisher, J.H. (1977). Chancery and the emergence of standard written English in the fifteenth century. *Speculum*, 52(4): 870–899.

Görlach, M. (1991). *Introduction to Early Modern English*. Cambridge University Press, Cambridge.

Nevalainen, T. (2006). *An Introduction to Early Modern English*. Edinburgh Textbooks on the English Language, Edinburgh University Press, Edinburgh.