

Abs-167

Elena Semino (Department of Linguistics and English Language, Lancaster University)

Using corpus linguistic methods to compare a simplified version of *Romeo and Juliet* with the original play

Shakespeare's plays are among the canonical literary works that have been simplified in a variety of editions aimed at readers who cannot easily access the original versions, due to their (young) age, lack of expertise and/or insufficient mastery of English as a foreign or second language. Previous studies on simplified or 'graded' readers have been concerned with their role in the teaching of English as a second or foreign language, particularly with respect to the acquisition of vocabulary. In this paper I report the results of a study that combined corpus-based methods with traditional 'manual' analysis in order to arrive at a systematic account of the differences between the original version of *Romeo and Juliet* and a particular simplified version, published in a series entitled *Shakespeare Made Easy* (Nelson Thornes). The series includes twelve parallel-text editions of Shakespeare's plays, aimed primarily at older school-children (e.g. those preparing for GCSE-level tests in the UK). The software tool Wmatrix (<http://ucrel.lancs.ac.uk/wmatrix/>) was used to compare the original text of the play with the simplified version. The key word tool in Wmatrix identified approximately 100 overused items, while the semantic annotation tool identified approximately 10 overused items (in both cases the cut-off point was a log-likelihood value of 6.63, which corresponds to 99% significance). A detailed analysis of both lists of overused items makes it possible to identify a variety of systematic patterns of simplification, including particularly: the modernisation of Early Modern English spellings, morphology, and lexis; a reduction in the use of personification and body-part metonymies; a reduction in the use of particular metaphors. In other words, a comparison between the two versions of the play by means of two of the tools within Wmatrix does not simply reveal predictable differences such as the replacement of 'thou/thee' with 'you', but also helps to identify less predictable differences that can be explained, for example, as strategies to reduce the density of figurative language in the play, and to allow for differences in attitudes and background knowledge between Elizabethan audiences and the main readership of the simplified version (e.g. a reduction in the use of commercial metaphors for women). Overall, the employment of corpus methods helped to test the validity of hypotheses made on the basis of a textual analysis of the two versions of the play, and additionally produced further insights that could not easily have been reached by other means.