Tracking Literary Reputation with Text Analysis Tools[*]

Maite Taboada

Mary Ann Gillies

Associate Professor
Department of Linguistics

The question of why writers' works, and by extension their literary reputations, fall in and out of critical and popular favour has long fascinated literary critics. In 1905, Marie Corelli was the best-known and most successful novelist in Britain. By 1950 she had been consigned to literary obscurity and few read her books. D.H. Lawrence did not enjoy wide recognition during his lifetime, yet he is now part of the English literary canon. How do we account for such shifts in literary reputation? These two questions form the core of our project, on literary reputation in Britain between 1900 and 1950.

We are currently conducting a pilot project with two authors: John Galsworthy and D.H. Lawrence. We have in mind a larger project, with

The final goal of our project is to be able to determine what in a reviewer's text seems to influence the literary reputation of a particular author, and whether what reviewers say can be mapped to the author's reputation trajectory. The most novel aspect of the project it the use of objective tools to extract highly subjective material.

In the process of constructing the database and creating tools we need to address a number of issues relating to the nature of the project itself and to the application of existing methodologies. First of all, we focus on quantifiable aspects of reputation, and how it varies according to audience. Each text is annotated with the reviewer's name, source, place of publication, and audience type. Thus, we can measure the impact of specific reviews.

Secondly, we are applying tools created for present-day texts (e.g., part-of-speech taggers and parsers). Our impression is that some modifications are necessary to account for the slightly different style of a different time and register. A related problem is the heavy use of irony and external references as indicators of evaluation.

Finally, the most important challenge lies with the evaluation of the system. The current evaluation of our present-day movie review system relies on whether the author recommended the movie or not. In literary reviews, the overall evaluation is much more subtle. We need to determine whether the system is flagging each revi