A Question of Style: corpus building and stylistic analysis of the *Edinburgh Review* and the *Quarterly Review*, 1814-1820

Dr Francesca Benatti and Dr David King, The Open University
Introduction

A Question of Style in brief
A Question of Style

- Winner, 2016 RSVP Field Development Grant

Research question:
- Did the *Edinburgh Review* create a “transauthorial discourse” (Jon Klancher) that hid individual authorial voices behind an impersonal corporate style?
INTRODUCTION
The Edinburgh Review

- Founded in 1802 by members of Whig intelligentsia
- Book reviews, but actually essays with political aims
- All articles published anonymously
- Most articles by now attributed
- Extremely influential, so Quarterly founded in 1809 as Tory counterpart
Editorial interventions

• *ER* editor Francis Jeffrey and *QR* editor William Gifford edited articles

• “retrenchments and verbal alterations” (Jeffrey)

• “tampering with articles” (Ugo Foscolo)

• Do these editorial interventions produce a uniform *ER* style?
INTRODUCTION
Additional questions

• Are reviews of the same text similar to one another? (10 pairs of reviews)

• Does the genre of the text being reviewed influence the style of the review? (3 genres)
Corpus selection

The composition and rationale of our corpus
Size, composition, rationale

- Articles from three genres:
  - Reviews of literature
  - Reviews of travel
  - Reviews of history

- Written by most frequent contributors in these fields
  - ER: Jeffrey, Brougham, Smith, Hazlitt, Mackintosh, Moore, Palgrave, Allen, Playfair
  - QR: Croker, Scott, Wilson, Barrow, Southey

- Reviewing significant literary or historical works
OCR correction

Post- Optical Character Recognition processing
Post-OCR processing

- Challenge: Need to correct OCR from Google Books
- Problem: OCR errors too inconsistent for automation
- Individual spelling choices
  - Publick
- Regional identities
  - Perswaded
- Language transformation
  - Shakspear, Shakspeare, Shakespear, Shakespeare
- Solution: David reviewed all proposed automated corrections and “spelling mistakes” against the digitised source image
Quotations

Or, what is in a review?
QUOTATIONS

Or, what is in a review?

![Chart Title]

- ERQR Corp: 65% non-quote, 35% quote
- ER Corp: 63% non-quote, 37% quote
- QR Corp: 71% non-quote, 29% quote
Or, what is in a review?

### Quotations % Max and Min

#### Corpus

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Quote % Max</th>
<th>Quote % Min</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERCorp</td>
<td>79%</td>
<td>1%</td>
</tr>
<tr>
<td>QRCorp</td>
<td>65%</td>
<td>11%</td>
</tr>
</tbody>
</table>
Or, what is in a review?

- The presence of quotations is a problem for our chosen analytical methods
- e.g. we want to analyse Francis Jeffrey’s style (reviewer), not Walter Scott’s (reviewed)
- Solution:
  - David marked quotations using TEI XML <quote> element
  - Then we removed quotations using XSL transformation
- This reduces the total size of the ERQR corpus to 512,702 words
Analysis

Stylometry
Stylometry in brief

- The study of how hidden stylistic traits can be measured through statistical methods to trace an author's voice

- Delta method introduced by John Burrows in his 2001 Busa Award lectures and beyond

- Generally concerned with authorship attribution but increasingly used to study style more broadly

- Burrows’ s Delta method implemented by Eder, Rybicki and Kestemont’s Stylo software package

- Improved method Cosine Delta developed by University of Würzburg

- Based on analysis of most frequent words
Stylometry: by journal using Stylo (Eder, Rybicki, Kestemont); Cosine Delta; 300 MFW
Stylometry: by genre (using Stylo (Eder, Rybicki, Kestemont); Cosine Delta; 300 MFW)
Stylometry: by author using Stylo (Eder, Rybicki, Kestemont); Cosine Delta; 300 MFW
Next steps
• Some traces of “house style”

• Influence of genre of text being reviewed

• Influence of text being reviewed (to an extent)
And more questions

• Stylometry with:
  • Character n-grams
  • Positive vs. negative reviews

• Corpus stylistics with:
  • Keywords
  • N-grams

• Assessment of the benefits of curation:
  • Keeping quotations
  • Using “raw” OCR

• Archival research
  • Comparing edited vs. unedited versions of articles
“Many interesting things cannot be counted, but many others can.”

—John Burrows
THANK YOU!

Download our corpus from ORDO (search for RSVP)
https://doi.org/10.21954/ou.rd.6850865.v1

Follow us on
http://www.open.ac.uk/blogs/styleproject/
@rhymesontheroad
@dh_ou