A Question of Style: individual voices and corporate identity in the *Edinburgh Review*, 1814-1820

Conference or Workshop Item

How to cite:


For guidance on citations see FAQs.

© [not recorded]

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
A Question of Style: individual voices and corporate identity in the *Edinburgh Review*, 1814-1820

Francesca Benatti and David King
The Open University
Research question

Did the *Edinburgh Review* create a “transauthorial discourse” (Klancher 1987) that hid the voices of individual contributors behind a corporate style?

**Funded by the Research Society for Victorian Periodicals Field Development Grant (January-October 2017)**
The *Edinburgh Review*

Most influential periodical in early 19th C.

Edited by Francis Jeffrey, who could make alterations to any article

All articles published anonymously
The Quarterly Review

Edinburgh staunchly Whig (liberal)

Tory rival, Quarterly Review started 1809 to counter it

Often reviewed same texts

Some authors wrote for both
Corpus at project end

Edinburgh Review:
- 325,000 ‘words’
- 60 articles

Quarterly Review:
- 175,000 ‘words’
- 20 articles

Chosen from reviews of
literature, travel writing, history,
politics and economics
Corpus preparation

Corpus selection

OCR correction

TEI encoding

Analysis with computational tools

Interpretation of results
Corpus selection
Corpus selection

...
## Corpus selection

<table>
<thead>
<tr>
<th>We wanted</th>
<th>We got…</th>
</tr>
</thead>
<tbody>
<tr>
<td>pages</td>
<td></td>
</tr>
<tr>
<td>172–208</td>
<td></td>
</tr>
<tr>
<td>in a complete,</td>
<td></td>
</tr>
<tr>
<td>consecutive</td>
<td></td>
</tr>
<tr>
<td>sequence…</td>
<td></td>
</tr>
<tr>
<td></td>
<td>172</td>
</tr>
<tr>
<td></td>
<td>173</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>178</td>
</tr>
<tr>
<td></td>
<td>179</td>
</tr>
<tr>
<td></td>
<td>182</td>
</tr>
<tr>
<td></td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>181</td>
</tr>
<tr>
<td></td>
<td>186</td>
</tr>
<tr>
<td></td>
<td>187</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>207</td>
</tr>
<tr>
<td></td>
<td>208</td>
</tr>
</tbody>
</table>
OCR correction

Most scans are good. Most OCR is good.

Building a library of before and after texts from which we can prepare an alignment, from which we can identify error patterns.

Farewt'u! a won! ijiat must be, and liallí been __ A sound nliicb makes us linger;

Farewell! a word that must be, and hath been __ A sound which makes us linger;
OCR correction

Typical incorrect recognition of individual characters:
\[ w \rightarrow vv, vi, iv \]
\[ h \rightarrow li \]

Some run in two directions:
\[ rn \leftrightarrow m \]
\[ u \leftrightarrow n \]
\[ s \leftrightarrow a \]
BUT

We want stylistic quirks:
public or publick to day or to-day

We do not want words normalised:
surprized to surprised Shak[e]speare[e]
BUT

Croker reviewing Scott's *Guy Mannering* writes 'gipsys' in his analysis, whereas Scott writes 'gipsies' in the story itself.
OCR correction

Some good tools already exist for issues such as:
• normalisation
• long S
• hyphenation  

No point trying to reinvent them.

Instead:
• building a set of post-processing scripts to aid manual review, e.g., to highlight all occurrences of *die* and its possible confusion with *the* in a text, helped by n-grams to assess confidence.
• porting TEI header, footer and paragraph insertion scripts from previous work to speed up mark-up.
Analysis – Stylometry

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

Made better known by John Burrows in his 2001 Busa Award lectures and beyond

Perception of authorial “voice” is quite subjective
  • e.g. Duncan Wu (Introduction, *New Writings of William Hazlitt*, 2007)
Two interpretations of style*

Style as fingerprint

Unconscious elements in the way we write
(e.g. Van Halteren et al. "Existence of a human stylome." (2005))

Reflected by use of Most Frequent Words

Style as signature

Conscious choice of words, sentences, tone
(e.g. Van Dalen-Oskam Riddle of Literary Quality project)

Still unsure how to identify with stylometry

* as defined by Sarah Allison at DH2016, Stylistics workshop, 12 July 2016
Fingerprint - Delta method

“Delta is the mean of the absolute differences between the z-scores for a set of word-variables in a given text-group and the z-scores for the same set of word-variables in a target text.”

Delta – continued

Delta works on the Most Frequent Words present in a given set of texts

All authors use Most Frequent Words differently

Underpinned by solid mathematical and linguistic foundations
## Delta – example

<table>
<thead>
<tr>
<th>Word</th>
<th>Moore</th>
<th>Coleridge</th>
<th>Godwin</th>
<th>Southey</th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
<td>7.71%</td>
<td>6.40%</td>
<td>6.90%</td>
<td>7.69%</td>
</tr>
<tr>
<td>of</td>
<td>5.85%</td>
<td>5.06%</td>
<td>4.49%</td>
<td>3.54%</td>
</tr>
<tr>
<td>and</td>
<td>2.83%</td>
<td>3.95%</td>
<td>3.52%</td>
<td>3.15%</td>
</tr>
<tr>
<td>to</td>
<td>2.97%</td>
<td>3.04%</td>
<td>3.01%</td>
<td>3.11%</td>
</tr>
</tbody>
</table>
Signature – possible routes

Van Dalen-Oskam
- vocabulary richness?
- word length?
- sentence length?

Allison
- medium-frequency words?
- words used vs. words avoided?

Mahlberg
- Corpus stylistics
Analysis – false clusters

Female pronouns

- Moore_French_Novels_34_1820_corr: 36%
- Jeffrey_Edgeworth_28_1817: 33%
- anon_christabel_edinburgh_review_27_1816: 32%
- Jeffrey_Lalla_Rookh_29_1817: 23%
- Brougham_melanges_30_1818: 21%

…and 10 texts contained no female pronouns at all
Fingerprint vs Signature

Both attempt to remove the influence of content over style in the analysis

**Fingerprint – MFW**
- Frequent words
- Choose what to *include* in the analysis
- Unconscious style?

**Signature – TF:IDF**
- Significant words
- Choose what to *exclude* from the analysis
- Conscious style?
Ongoing work

- Enhance scripts
- Include more texts
- Expand reference corpora
- Share scripts and TEI texts
- Evaluate and critique
A Question of Style Project

Francesca Benatti
David King

Faculty of Arts and Social Sciences
Faculty of Science, Technology, Engineering and Mathematics
The Open University
Milton Keynes

Project blog:
http://www.open.ac.uk/blogs/styleproject/
Project outputs (in 2018):
https://ou.figshare.com/