What can a diachronic corpus of modern British fiction tell us about changes in (British) society?

WiP, ICAME39, Tampere 2018

Jarle Ebeling
Signe Oksefjell Ebeling
Aims

• First explorations of a new resource

1. In what ways does a keyword analysis reflect changes in (British) society?

2. In what ways does a general and ambiguous keyword (*poor*) reflect changes in (British) society?
Content

- (Extracts of) Novels and short story collections published between 1900 and 2018 by writers born and/or educated in the United Kingdom
- No children's/juvenile fiction, science fiction, fantasy literature or drama
- Texts published between 1900-1940 mainly harvested from Project Gutenberg (.org, .ca, .net.au) and Fadedpage
- More recent texts are from the BNC and other corpora/sources
- March 2018: 510 texts (190 female / 320 male), approx. 37 mill. words and 321 different writers
Method

- **Periods**
  - P1: 1900-1939
  - P2: 1940-1979
  - P3: 1980-2018

- **Generate keyword lists** (WordSmith Tools 6 (Scott 2012))
  - P1 vs. ref. corpus (P2+P3)
  - P2 vs. ref. corpus (P1+P3)
  - P3 vs. ref. corpus (P1+P2)

- **Classification of keywords into broad semantic categories**

- **Focused study of the meanings and use of the word *poor*, which is key in P1**
Keyword analysis / Semantic categories

<table>
<thead>
<tr>
<th>P1</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family relations</td>
<td>War/Military</td>
<td>Technology/inventions</td>
</tr>
<tr>
<td>Forms of address</td>
<td>Food &amp; drink</td>
<td>Leisure</td>
</tr>
<tr>
<td>Religion</td>
<td>Technology</td>
<td>Indulgence</td>
</tr>
</tbody>
</table>

Examples:
- P1: divine, religion, cousin, aunt, marry, servant(s), captain, lady, lord, honour, bowed
- P2: enemy/ies, smoke, war, uniform, cigarette(s), eating, whisky, telephone(d), receiver
- P3: screen, camera, plastic, TV, pub, sex(ual), shopping, weekend, coffee, wine,

Changes in keywords reflect changes in society and involve a shift from
- a family-centered class society occupied with religion and the divine to
- a war-ridden society followed by post-war optimism to
- a technology-centered, pleasure-seeking, affluent society
The case of *poor*

- Keyword in P1 (1910-1939), but frequent in the CBF overall

- Sharp decrease in the use of *poor* over the century:
  - P1: 393 pmw
  - P2: 292 pmw
  - P3: 169 pmw

- Is there any particular use or meaning that can account for the decrease, and if so, can it be explained from a socio-cultural perspective?
**Poor: meaning and use**

- Attributive and predicative uses
  - lacking sufficient money – … because they’re poor …
  - low quality – … made a poor witness …
  - deserving pity/sympathy (attributive only): My poor girl …

- Nominalized adjective
  - lacking sufficient money: … the poor get poorer…

Hypothesis: there will be a marked drop in the use of the «lacking sufficient money»-meaning, which may account for the decrease from P1 to P3.

Random sample of 200 concordance lines per period
Proportions of meanings and uses in the sample

Distributions of meanings and uses fairly stable across the three periods

- lacking sufficient money (poverty): P1: 38, P2: 30, P3: 28
- low quality: P1: 14, P2: 11, P3: 22

• **Attributive**: P1: 178, P2: 181, P3: 177
• **Predicative**: P1: 9, P2: 16, P3: 14
• **Nominalized ADJ**: P1: 13, P2: 3, P3: 9
Attributive pity/sympathy use

Although not perfect precision, the following search was performed:

- poor (NN P) and excluding instances such as too/so poor, the poor, poor followed by full stop, BE poor, etc., etc.

Decrease in the use of the pity/sympathy meaning: in line with the overall keyword analysis, where P1 was seen to be more community/family-centered. P3's focus on leisure/indulgence («me-me-me» society) generates less sympathy for other people.

(Alt. explanation: P2/P3 use other ways of expressing pity/sympathy)

Typical examples of combinations that have seen a drop:
- poor Myles;
- poor lady Vivian;
- poor cousin;
- poor woman;
- poor devil;
- poor thing

> summary(pity_P1$Per)
  Min. 1st Qu. Median    Mean 3rd Qu.    Max.
  0.130  1.700  2.990   3.499   4.500  17.250

> summary(pity_P2$Per)
  Min. 1st Qu. Median    Mean 3rd Qu.    Max.
  0.140  1.020  2.030   2.599   3.340  13.890

> summary(pity_P3$Per)
  Min. 1st Qu. Median    Mean 3rd Qu.    Max.
  0.170  0.740  1.215   1.630   2.005   8.790
Concluding remarks

• These first explorations of the CBF show some of the potential of a large diachronic corpus of one broad genre in one variety of English.

• Gives a snapshot of society through literature over a century.
  – A keyword analysis of a corpus of fiction texts seems to mirror society.
  – Even the analysis of the ambiguous keyword *poor* may shed some light on changes in society.

• This WiP report has only scratched the surface of how literature reflects society. Further study is needed to validate the (preliminary) findings presented.

“Literature reflects society and society shapes literature.”
-- Oscar Wilde
Quiz: exclusive to P1, P2 or P3?

\textit{poor}

- bastard
- boy
- bugger
- child
- creature
- dear
- devil
- fellow
- girl
- man
- sod
- soul
- thing
- woman
<table>
<thead>
<tr>
<th>Rank</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>thing</td>
<td>man</td>
<td>man</td>
</tr>
<tr>
<td>2</td>
<td>man</td>
<td>devil</td>
<td>girl</td>
</tr>
<tr>
<td>3</td>
<td>dear</td>
<td>thing</td>
<td>thing</td>
</tr>
<tr>
<td>4</td>
<td>child</td>
<td>girl</td>
<td>woman</td>
</tr>
<tr>
<td>5</td>
<td>devil</td>
<td>boy</td>
<td>bastard</td>
</tr>
<tr>
<td>6</td>
<td>fellow</td>
<td>fellow</td>
<td>sod</td>
</tr>
<tr>
<td>7</td>
<td>girl</td>
<td>dear</td>
<td>child</td>
</tr>
<tr>
<td>8</td>
<td>boy</td>
<td>child</td>
<td>bugger</td>
</tr>
<tr>
<td>9</td>
<td>soul</td>
<td>woman</td>
<td>boy</td>
</tr>
<tr>
<td>10</td>
<td>woman</td>
<td>soul</td>
<td>creature</td>
</tr>
</tbody>
</table>

Rank 11-14
P1: chap, creature, father, people
P2: chap, people, darling, bastard
P3: soul, devil, people, chap
The open part of the corpus (1900-1940):

- https://nabu.usit.uio.no/hf/ilos/oelc/
References


Scott, M. (2010). Problems in investigating keyness, or clearing the undergrowth and marking our trails... In M. Bondi & M. Scott (Eds.), *Keyness in text* (pp. 43-57). Amsterdam: Benjamins.