CLiC Dickens
A cognitive corpus stylistic approach to patterns of body language presentation in fiction

Johan de Joode, Michaela Mahlberg, and Peter Stockwell

Research context
- Prose fiction is typically treated as a single register without distinguishing between fictional speech and narration.
- Typically the focus of stylisticians is on the presentation of speech, thought, and writing (Leech and Short 2007).
- In previous work we have shown that patterns of body language play an important role as a textual technique of characterisation.
- This paper outlines how we investigate these patterns with the help of suspensions. We build on previous studies, but do so at scale.

Definitions and examples
- suspension: A narratorial interruption of character speech that does not end with sentence-final punctuation.
- non-quote: Any textual unit that is not a quote.
- quote: A textual unit that starts and ends with respectively single or double quotation marks. It may represent speech, writing, or thought.

Rosa has been crying and is yet in distress. On her coming in, the ironmaster leaves his chair, takes her arm in his, and remains with her near the door ready to depart.

"You are taken charge of, you see," says my Lady in her weary manner, "and are going away well protected. I have mentioned that you are a very good girl, and you have nothing to cry for."

"She seems after all," observes Mr. Tulkinghorn, loitering a little forward with his hands behind him, "as if she were crying at going away.

Corpus

<table>
<thead>
<tr>
<th>Content</th>
<th>Tokens</th>
<th>Tokens in quotes</th>
<th>Tokens in suspensions</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNov</td>
<td>3,837,284</td>
<td>1,326,893</td>
<td>108,634</td>
</tr>
<tr>
<td>19C</td>
<td>4,501,633</td>
<td>1,480,800</td>
<td>42,201</td>
</tr>
</tbody>
</table>

Suspension and quote counts
In DNov there are proportionally more words in suspensions, than in 19C. Dickens has a tendency to interrupt his characters.

Clusters in suspensions
These suspensions reveal interesting patterns in Dickens’s characterisation.

<table>
<thead>
<tr>
<th>n-gram alignment</th>
<th>1-grams</th>
<th>2-grams</th>
<th>3-grams</th>
<th>4-grams</th>
<th>5-grams</th>
</tr>
</thead>
<tbody>
<tr>
<td>key clusters suspensions vs. quotes, P &lt; 0.000001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Body Part Nouns
The percentage of body part clusters, for instance, increases with the length of the cluster.

The CLiC Interface

Future research
Other features that influence the characterisations of people and relationships:
- impertinences
- the intensity
- the keywords
- colloquialisms
- the length of the interventions

References

http://clic.nottingham.ac.uk
johan.dejoode@nottingham.ac.uk