“Judge Dredd” then and now: A corpus-assisted diachronic analysis of a comic strip.

Judge Dredd, the comic strip used for this study, has appeared in the comic 2000AD since 1977 without a break, making it one of the longest running comic strips in UK comic history.

The aim of the study: develop a picture of how language is used in the comic strip through corpus analysis.

Also: by collecting data from two different points in time, investigate whether any features of language usage have changed.

Two corpora were created of comic strip data of around 10000 words each. One represented “Judge Dredd” from 1977/78 (then) and the other from 2002/03 (now).

Table 1. Corpus profiles

<table>
<thead>
<tr>
<th>Corpus name</th>
<th>Year</th>
<th>No. of words</th>
<th>no. of editions</th>
<th>avg words/edition</th>
</tr>
</thead>
<tbody>
<tr>
<td>JD7778C</td>
<td>1977/78</td>
<td>11127</td>
<td>17</td>
<td>655</td>
</tr>
<tr>
<td>JD0203C</td>
<td>2002/03</td>
<td>10664</td>
<td>19</td>
<td>561</td>
</tr>
</tbody>
</table>

Wordlists and Keywords

Wordlists - AKA frequency lists. Often an “essential starting point” Stubbs 2001:126. It is often a good idea to compare the frequency list you generate from the corpus under investigation with one from a reference corpus.

Wordlists from BNC available at: http://ucrel.lancs.ac.uk/bncfreq/ (or see Leech, Rayson and Wilson 2001)

Keywords - Words which occur in a text or corpus unusually frequently (positive keywords) or unusually infrequently (negative keywords) when compared to a (larger) reference corpus (e.g. the BNC).

Note: You need to decide the most suitable reference corpus for your study.

Parts of speech (POS) tagging - You might want to carry out searches at a grammatical level (rather than the word level). This is possible using Wordsmith (Scott 1996) and AntConc (Anthony 2008), but first you will need to create a POS tagged version of your corpus. You can do this by running your data through the CLAWS tagger (developed here at Lancaster University).

Automatic tagging service using the CLAWS Tagger: http://ucrel.lancs.ac.uk/claws/trial.html

The C7 tag set used by the CLAWS Tagger: http://ucrel.lancs.ac.uk/claws7tags.html

Or you could use Wmatrix (Rayson 2008), which automatically tags for POS using CLAWS when you load your texts into it.

A POS tagged corpus can also useful if you want to differentiate between words that are spelt the same but belong to a different grammatical class. For example, TO as an infinitive marker vs. TO as a preposition:

TO as an infinitive would be tagged: to_TO where to = the word, and _TO = the tag. (Assuming C7 tagset)

TO as a preposition would be tagged: to_II where to = the word, and _II = the tag (Assuming C7 tagset)

Analyzing JD7778C with JD0203C

Comics are made up of a number of components (see Table2), and part of my analysis was interested in the structure of the comics and where the words occurred.

Task: Using your knowledge of mark-up and annotation, consider what information you could add to the Judge Dredd corpora that might make a component analysis possible.
Table 2. Comic components

<table>
<thead>
<tr>
<th>Component</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech Balloons</td>
<td>Interaction between characters; ‘running commentary’ by characters helps to tell reader what is happening</td>
</tr>
<tr>
<td>Captions</td>
<td>Often open a story; brings the reader up to date with events; In JD0203C, used more for internal monologue or free indirect thought; in JD7778C more like 3rd person narration.</td>
</tr>
<tr>
<td>Sound Effects</td>
<td>Used much more extensively in JD0203C than in JD7778C. Increased usage could be due to better quality pictures which allow more things to be happening in the panel without seeming cluttered.</td>
</tr>
<tr>
<td>Picture Text</td>
<td>Totals very low - difficult to draw any conclusions, but still an important component in the telling (and understanding) of the story. Possibly needs further research.</td>
</tr>
<tr>
<td>Thought Balloons</td>
<td>Used more in JD7778C – create tension/drama (high frequency of but). In JD0203C only used in one story.</td>
</tr>
</tbody>
</table>

Some references: