A multidimensional analysis of student writing across levels and disciplines

Hilary Nesi  h.nesi@coventry.ac.uk
www.coventry.ac.uk/bawe

The BALE corpus:

- 6,506,995 words
- 2,896 texts
- 2,761 assignments
- 1,039 contributors
- 30+ disciplines
- 13 genre families

Tagged for 67 linguistic features, grouped into 16 grammatical/functional categories:

1. Tense and aspect markers
2. Place and time adverbials
3. Pronouns and pro-verbs
4. Questions
5. Nominal forms
6. Pasives
7. Subordination features
8. Stative forms
9. Prepositional phrases, adjectives and adverbs
10. Lexical specificity
11. Lexical classes
12. Modals
13. Specialized verb classes
14. Reduced forms and dispreferred structures
15. Coordination
16. Negation

Distribution of factors across genres

<table>
<thead>
<tr>
<th>Involved</th>
<th>Narrative</th>
<th>Explicit</th>
<th>Abstract</th>
<th>Persuasive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essay</td>
<td>-14.3</td>
<td>-2.5</td>
<td>6.2</td>
<td>5.9</td>
</tr>
<tr>
<td>Meth. Recount</td>
<td>-15.8</td>
<td>-3.7</td>
<td>4.5</td>
<td>7.3</td>
</tr>
<tr>
<td>Critique</td>
<td>-14.8</td>
<td>-3.1</td>
<td>6.0</td>
<td>6.4</td>
</tr>
<tr>
<td>Explanation</td>
<td>-15.4</td>
<td>-3.6</td>
<td>5.0</td>
<td>5.8</td>
</tr>
<tr>
<td>Case Study</td>
<td>-16.4</td>
<td>-2.9</td>
<td>5.8</td>
<td>4.3</td>
</tr>
<tr>
<td>Exercise</td>
<td>-12.1</td>
<td>-3.9</td>
<td>4.6</td>
<td>5.7</td>
</tr>
<tr>
<td>Design Spec</td>
<td>-13.9</td>
<td>-4.0</td>
<td>4.1</td>
<td>6.8</td>
</tr>
<tr>
<td>Proposal</td>
<td>-16.4</td>
<td>-3.8</td>
<td>6.3</td>
<td>4.8</td>
</tr>
<tr>
<td>Narrative Recount</td>
<td>-4.8</td>
<td>-1.1</td>
<td>3.8</td>
<td>4.0</td>
</tr>
<tr>
<td>Research Report</td>
<td>-16.2</td>
<td>-3.1</td>
<td>5.5</td>
<td>7.2</td>
</tr>
<tr>
<td>Problem Question</td>
<td>-12.0</td>
<td>-2.8</td>
<td>5.2</td>
<td>6.4</td>
</tr>
<tr>
<td>Literature Survey</td>
<td>-17.9</td>
<td>-2.6</td>
<td>6.3</td>
<td>5.0</td>
</tr>
<tr>
<td>Empathy Writing</td>
<td>-11.5</td>
<td>-2.2</td>
<td>4.5</td>
<td>4.5</td>
</tr>
</tbody>
</table>

**Multiple Range Test Scores for levels**

- **Involved**: General academic prose = -15. Fewer verbal features at each level of progression – less “author-centred”, more “object centred”: Significant differences between levels 1/2 and 3/4.
- **Narrative**: Academic prose typically falls between -2 and -3. A steady decrease in narrative features
- **Explicit**: Increasingly explicit.
- **Persuasive**: General academic prose = 0 (neutral). BALE corpus writing scores are consistently negative.
- **Abstract**: Highly passivized, scores similar to general academic prose. A steady increase in abstraction, reversed at Masters level.

**Different colours = significant differences**

**Involved / Narrative**
- Entirely negative scores on the ‘involved’ and ‘narrative’ dimensions indicate a high informational focus and a low level of narration.

**Explicit / Abstract**
- Entirely positive scores on the ‘explicit’ and ‘abstract’ dimensions indicate a lexically dense text containing passivess, past participle clauses, and other features typical of academic prose.

**Persuasive**
- Mixed scores on the “persuasive” dimension indicate variation in the degree of argumentation (Problem questions and Proposals being the most persuasive, and Literature Surveys the least).

An Investigation of Genres of Assessed Writing in British Higher Education
ERRC Research Award RES-000-23-4880
November 2004 – December 2007

Acknowledgement
Multidimensional analysis for this study was carried out by Douglas Biber and his research team at Northern Arizona University. Basic information about the five dimensions can be found in Biber (1988) and Conrad & Biber (eds) (2001).