Contextual Information and Metadata

Sheena Gardner and Dawn Hindle
CELTE, University of Warwick

Corpus Linguistics Birmingham July 2005
Contextual vs. Linguistic

“the social context (place, time, participants) ..

is arguably .. as significant as ..

intrinsic linguistic features”

Burnard 2004:8
External vs. Internal

- External: essentially non-linguistic
- Internal: essentially linguistic

“…the interrelation between them is one of the areas of study for which a corpus is of primary value. In general external criteria can be determined without reading the text.”

Atkins et al. 1992:5
<table>
<thead>
<tr>
<th>Arts &amp; Humanities</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life Sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Contextual perspectives: from the discourse community

a. Department documentation

b. Tutor interviews & surveys

c. Student submission forms
Sociology Assignment Types
(interview)

- Essays
- Book reviews
- Book reports
- Projects
- Urban Ethnography Assignments
- Fieldwork Reports
- Dissertations
‘Linguistic’ perspectives: from text analysis

- Multidimensional analysis will tell us about the language of texts in our 4 x 4 grid. Broadly, a description of register.

- SFL Genre analysis will tell us about the generic stages of texts. Broadly, a description of text organization and purpose.
Are essays in Sociology written in the same general academic register as essays in History?

Are essays in Sociology organised in the same way as essays in Biology?

And if the answers are ‘No’, do we still call them all ‘essays’? i.e. Can we blend contextual and linguistic categories?
Atkins et al. assume realist categories:

- **style** = *prose/verse/rhyme* "as determined by surface features of the text or author’s claims" (1992:8)
- **region** = *regional type* - ‘to be refined by internal evidence’

Burnard suggests it is not possible to entirely distinguish contextual and linguistic information (2004:5)
Diverse Issues in Blending

- Linguistic choices can be interpreted differently by members of different discourse communities e.g. business and EFL (Forey 2004)
- Same labels (e.g. ‘essay’), different values - whose meaning predominates?
- Genre analysis is of language in context – requires both
Three Discrete classifications

1. ‘Contextual’ Assignment Types (from the discourse community)

2. ‘Linguistic’ Registers (from analysis of surface linguistic features)

3. ‘Contextual + Linguistic’ Genres (from analysis of language in context)
Issues in contextual information

“[Metadata] extends from straightforward labelling and identification of individual items to the detailed representation of complex interpretive data associated with their linguistic components”

Burnard 2004: 15
Year of study?

In terms of university administration, a third year student could be:

- in their third year of successfully completed study;
- a student in the final year of their course, but who spent their “third year” doing an industrial placement (intercalated year).

How will we classify this?
Year of study?

Categories: 1, 2, 3, 3*, 4, 4*

- 3 = 3rd year, no intercalated year
- 3* = 3rd year, intercalated year
- 4 = postgraduate
- 4* = 4th year undergraduate Masters
Language and educational background

- Native vs. non-native English speakers

- Submission form data:
  1. first language
  2. secondary education
Is it in or is it out?

- What counts as ‘assessed’?
- What counts as English?
- What counts as ‘writing’?
- Creative writing / Fiction
Tradition vs. innovation

“We’re quite a traditional department in that we still use mainly essays, we’re very conscious that we would like to, and perhaps need to, do something about that”

Mike Neary, Director of Undergraduate Studies, Sociology, Warwick, 26/04/2005
<table>
<thead>
<tr>
<th>Contextual Info as Metadata (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Student</strong></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td><strong>First Language</strong></td>
</tr>
<tr>
<td><strong>UK Education</strong></td>
</tr>
<tr>
<td><strong>Course</strong></td>
</tr>
<tr>
<td><strong>Contextual Info as Metadata (2)</strong></td>
</tr>
<tr>
<td>-----------------------------------</td>
</tr>
<tr>
<td><strong>Assignment</strong></td>
</tr>
<tr>
<td><strong>Title</strong></td>
</tr>
<tr>
<td><strong>Date written (year-month)</strong></td>
</tr>
<tr>
<td><strong>Level (of student)</strong></td>
</tr>
<tr>
<td><strong>Module title</strong></td>
</tr>
<tr>
<td><strong>Module code</strong></td>
</tr>
<tr>
<td><strong>Assignment Type</strong></td>
</tr>
<tr>
<td><strong>JACS Code</strong></td>
</tr>
<tr>
<td><strong>Disciplinary Grouping</strong></td>
</tr>
<tr>
<td><strong>Grade</strong></td>
</tr>
<tr>
<td><strong>Authors</strong></td>
</tr>
</tbody>
</table>