Embracing different semiotic modes in undergraduate assignments

Conference or Workshop Item

How to cite:

© 2012 The Author

Version: Accepted Manuscript

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
Embracing different semiotic modes in undergraduate assignments

Maria Leedham  
m.e.leedham@open.ac.uk
Outline

1. Background

2. Corpus data and methods

3. Findings
   - from corpus linguistic keywords and counts
   - from comparison of pairs of assignments
   - from interview data

4. Implications
Outline

1. Background

2. Corpus data and methods

3. Findings
   - from corpus linguistic keywords and counts
   - from comparison of pairs of assignments
   - from interview data

4. Implications
Deficit approach

- student writing is ‘remedial’, ‘immature’ and contains ‘problems’ or ‘errors’ - especially L2 English student writing
  (Chen and Baker, 2010; Paquot, 2010)

Vs.

Academic literacies approach

- writing within the academy is a set of social practices in which genre, context and culture are highly significant
- highlights ‘the variety and specificity of institutional practices, and students’ struggles to make sense of these’ (Lea and Street, 2006: 376).
- All student writers are in a constant struggle to establish the preferred ways of making meaning within their particular context (e.g. Lillis, 2006).
Background: Undergraduate assignments

• high-stakes, occluded and under-researched

• from monolithic to research on disciplinary differences (e.g. Hewings, 1999; Hyland, 2008).

• increasing awareness of range of genres required at UG level. (Leedham, 2009; Nesi & Gardner, 2006).

• rise in new genres – (e.g. e-posters, reflective blogs, website evaluations, press releases).

‘unprecedented amount of innovation in assessment’ (Gibbs, 2006:20).
Outline

1. Background

2. **Corpus data and methods**

3. Findings
   - from corpus linguistic keywords and counts
   - from comparison of pairs of assignments
   - from interview data

4. Implications
Corpus data

British Academic Written English (BAWE)

- 6.5 million words
- 2,761 assignments
- 1,039 writers
- 30+ disciplines
- Variety of L1s
- All proficient writing

The corpora for this study

- L1 English & L1 Chinese UG texts from BAWE & beyond
- 5 disciplines
- =>
  - 104 texts from Chinese students
  - 295 texts from British students
- Plus insights from lecturer interviews

ESRC project number
RES-000-23-0800
Methods

1. Corpus linguistic keywords and counts
2. Comparison of pairs of assignments
3. Bawe lecturer interviews
Outline

1. Background

2. Corpus data and methods

3. Findings
   - from corpus linguistic keywords and counts
   - from comparison of pairs of assignments
   - from interview data

4. Implications
## Keywords relating to visuals and lists

<table>
<thead>
<tr>
<th>L1&amp; discipline</th>
<th>Chi-Biol</th>
<th>Chi-Bus</th>
<th>Chi-Econ</th>
<th>Chi-Engin</th>
<th>Chi-Food</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selected keywords</td>
<td>#</td>
<td>growth</td>
<td>#</td>
<td>curve</td>
<td>#</td>
</tr>
<tr>
<td></td>
<td>table</td>
<td>curve</td>
<td>eq.</td>
<td>according</td>
<td>curve</td>
</tr>
<tr>
<td></td>
<td>data</td>
<td>refer</td>
<td>referring</td>
<td>referring</td>
<td>statistical</td>
</tr>
<tr>
<td></td>
<td>equation</td>
<td>model</td>
<td>figure</td>
<td>deviation</td>
<td>numbers</td>
</tr>
<tr>
<td></td>
<td>figure</td>
<td>per</td>
<td>output</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Keywords relating to visuals and lists

Biology
• All "Phases" are labeled on the graph. The curve of the Exponential Phase was straight, though some point lay outside this best straight line of fit (0041a).

Economics
• Actually the total loss resulting from the lower monopoly output \((Q \, M)\) is the grey triangle. The part of the grey triangle above \(P \, C\) is the loss of consumer surplus (6008q).

Engineering
• According to the program and refer to the figure 4.1.1, it is easy to find… (6107d).

Food Science
• According to the 3 sets of data calculated above… (6150d).
A year 3 Engineering assignment
Photograph 9 – Tree shrew pollinating the male flower

[Attenborough, 1995]

<table>
<thead>
<tr>
<th>Carbon Content %</th>
<th>Classification</th>
<th>Classification Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3-0.4</td>
<td>Low Carbon Steel</td>
<td>General for weld</td>
</tr>
<tr>
<td>0.3-0.7</td>
<td>Medium Carbon Steel</td>
<td>Used for nuts, shadil</td>
</tr>
<tr>
<td>0.7-1.7</td>
<td>High Carbon Steel</td>
<td>Use application</td>
</tr>
</tbody>
</table>

Table 1

Visuals
Lists and ‘listlikes’

- Sales promotion:
  - Monthly promotions, according to customer sales and current interest
  - Discounts for bulk orders
  - Free P&P on orders over £25, encouraging bulk buying
  - Discounts for new businesses using us for the first time, on condition they use us for a minimum of two more orders

- Public relations:
  - User friendly website
  - Easy search tools within website, enabling you to find the exact bulb you want even for the engineering minded
  - Extensive “Help” and “FAQ” pages
  - 12 hour guaranteed reply to email queries

Conclusions

The experiment yielded the following conclusions:

- The efficiency of a single stage centrifugal pump at high pump speed (3000 RPM) is better than it at low pump speed (2000 RPM).
- The input power with high pump speed increases faster than the one with low pump speed as discharge increases.
- The relationship between total head and discharge is not affected by pump speed, but higher pump speed provides higher total head.
## Counts of visual and list items

<table>
<thead>
<tr>
<th></th>
<th>Tables</th>
<th>Figures</th>
<th>Lists</th>
<th>Listlikes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Biol</td>
<td>15****</td>
<td>25****</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Eng-Biol</td>
<td>5</td>
<td>13</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Chi-Econ</td>
<td>1</td>
<td>14****</td>
<td>2*</td>
<td>25****</td>
</tr>
<tr>
<td>Eng-Econ</td>
<td>0</td>
<td>12</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Chi-Bus</td>
<td>2</td>
<td>2</td>
<td>6*</td>
<td>129****</td>
</tr>
<tr>
<td>Eng-Bus</td>
<td>6**</td>
<td>6**</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>Chi-Food</td>
<td>20*</td>
<td>6</td>
<td>5</td>
<td>82****</td>
</tr>
<tr>
<td>Eng-Food</td>
<td>14</td>
<td>6</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>Chi-Engin</td>
<td>10*</td>
<td>21</td>
<td>7</td>
<td>53****</td>
</tr>
<tr>
<td>Eng-Engin</td>
<td>7</td>
<td>21</td>
<td>10</td>
<td>24</td>
</tr>
</tbody>
</table>

* p<.05
** p<.01
**** p<.0001

per 10,000 words
Outline

1. Background

2. Corpus data and methods

3. Findings
   - from corpus linguistic keywords and counts
   - from comparison of pairs of assignments
   - from interview data

4. Implications
The role of maternal effect genes in the development of the nematode Caenorhabditis elegans

ABSTRACT

Caenorhabditis elegans (C. elegans) has been used as one of the favourite model organisms for developmental studies. Embryogenesis of C. elegans extensively relies on maternal effect genes for intrinsically asymmetric cell division and cell-cell interactions. In this review, the early embryogenesis of C. elegans, from the establishment of Anterior-Posterior polarity initiated by sperm entry to the asymmetrical cell divisions and different cell lineages induced by a variety of cell fate determinants is summarized. Some of the molecular mechanisms carried out by the crucial maternally expressed cell fate determinants underlying these processes are described.

INTRODUCTION

The C. elegans and its life cycle

Caenorhabditis elegans (C. elegans) is a small (~1mm long) free living soil nematode that has a predominantly hermaphroditic adult life. (Figure 1)

Figure 1  Adult C. elegans [1] Upper diagram: differential interference contrast image of an adult C. elegans. Lower diagram: anatomical structures of adult C. elegans (schematic drawing). Middle Left scale bar: 0.1 mm

The life cycle of C. elegans contains an embryonic stage, four larval stages (L1-L4) and an adult stage. (Figure 2) Molt (apopthesis, new cuticle formation, and ec dysis) takes place at the end of each larval stage. Under certain external conditions such as starvation, a non-growing stage, dauer larva, may form through a facultative, reversible, arrest at the lethargus in the second of four cuticle moults. The life cycle is about 2 to 3 weeks. Each approximately 8,000 C. elegans proteins have already been matched to homologous human gene transcripts (Lai et al., 2000). Specific mutants may be produced by targeted deletion through transposon insertion or mutagens. Embryos may be manipulated by transformation or injection with transgenes and marker proteins such as green fluorescent protein (GFP) are easily visualised in the transparent embryos. RNA interference (RNAi) is a particularly useful technique for studying maternal effect genes by eliminating the expression of specific maternal or zygotic genes in offspring.

Reproduction

In hermaphrodite worms, fertilization occurs in the spermatheca – an organ where the sperm is stored – when mature oocytes pass from the ovary towards the vulva (Fig 1A–B). The point of sperm entry determines the posterior end of the embryo. After fertilization, a rigid, ovoid-shaped chitin eggshell called the chorion is made (Kemphues & Strome, 1997) and the long axis of this ovoid is termed the anteroposterior (a/p) axis of the embryo.
Visuals and extended captions in Biology

- Comparison of two Biology assignments

<table>
<thead>
<tr>
<th>Text feature</th>
<th>Chinese, text 0434a</th>
<th>English, text 0067b</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of pages excluding refs</td>
<td>15.5</td>
<td>9</td>
</tr>
<tr>
<td>No. of tokens (in WS)</td>
<td>3234</td>
<td>3201</td>
</tr>
<tr>
<td>No. of tables</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>No. of figures</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>Visuals as proportion of whole text</td>
<td>48% (7.5 pp)</td>
<td>22% (2 pp)</td>
</tr>
<tr>
<td>Layout</td>
<td>whole page</td>
<td>2 columns</td>
</tr>
</tbody>
</table>
Diagrams and extended captions in text by Chinese writer

• 186 words
• different font
• text wrapping
• full sentences
• same neutral stance as main prose (e.g. use of passives, no first person pronouns, formal language)
• freestanding text i.e. separate reading path
2. Bivariate Regression and Multivariate Regression

(a) Bivariate Regression

**QTMARK = β0 + β1ATTR + ε**

The following results are obtained after running the bivariate regressions in EVIEWS

**QTMARK = 61.67025 + 0.00231 ATTR**

Interpretation for the regression results:

- The increase of 0.00231 means that even students who did not attend any revision lecture, they could get 64.97 in the exam, which may not make much economic sense as revision lectures are designed to boost a student’s exam marks.
- Ridge coefficient of 0.00231 shows a strong relationship between the attendance of revision lectures and exam performance. It means that 1% increase in the proportion of revision lecture attendance would decrease 0.00231 in the exam.
- The calculated t-score of 2.314 shows that the coefficient is small, so the null hypothesis cannot be rejected.
- A value of 0.00000000000783 of the variance in QTMARK is explained by ATTR. Therefore, it could be concluded that the coefficient has a significant effect on exam performance that it could even be omitted.

Two-tailed t-test for the significance of the slope β:

H0: β = 0 (Proportion of revision lecture attended does not affect exam performance)
H1: β ≠ 0 (Proportion of revision lecture attended does affect exam performance)

Since the calculated t-score of 2.314 is lower than the critical value of 2.056 at 5% significance level with 370 d.f., we fail to reject H0. This case concludes that revision lecture attendance does not affect exam performance.

(b) Multivariate Regression

**QTMARK = β0 + β1ATTR + β2ABILITY + β3HSQT + ε**

Modelling by OLS, we get

**QTMARK = 56.52575 + 0.16584 ATTR + 0.849900 ABILITY + 0.417658 HSQT**

Interpretation of the regression results:

- The coefficient of ATTR is 0.16584, which is significant at the 5% level. This means that for every 1% increase in the proportion of revision lecture attendance, the exam score increases by 0.16584.
- The coefficient of ABILITY is 0.849900, which is also significant at the 5% level. This means that for every 1% increase in the proportion of revision lecture attendance, the exam score increases by 0.849900.
- The coefficient of HSQT is 0.417658, which is significant at the 5% level. This means that for every 1% increase in the proportion of revision lecture attendance, the exam score increases by 0.417658.

As reported in question 1, the correlation coefficient was 0.67.
## Bulleted lists vs. connected prose in Economics

<table>
<thead>
<tr>
<th>Text feature</th>
<th>Chinese, 0155a</th>
<th>English, 0202j</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of pages excluding refs</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>No. of tokens (in WS)</td>
<td>3731</td>
<td>4242</td>
</tr>
<tr>
<td>No. of formulae</td>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td>No. of lists</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>No. of listlikes</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>Lists and listlikes as % of whole text</td>
<td><strong>90%</strong></td>
<td><strong>0%</strong></td>
</tr>
</tbody>
</table>
2. Bivariate Regression and Multivariate Regression

(a) Bivariate Regression

**QTMARK** = b4ATR + b5ATT

The following results are obtained after running the bivariate regressions in EViews.\(^2\)

**QTMARK** = 64.67023 - 0.002371ATT

Interpretation for the regression results:

- The intercept 64.67 means that even students did not attend any revision lecture, they could get 64.97 in the exam, which may not make much economic sense as revision lectures are designed to boost a student’s exam marks.
- High coefficient of -0.002371 shows an inverse relationship between the attendance of revision lectures and exam performance. It means that 1% increase in the proportion of revision lecture attendance would decrease grade by 0.002371 in the exam.
- Estimation of the intercept for the QTMARK model, which is the coefficient of A, shows that the coefficient is small, we could nearly omit its effect.
- Unadjusted value of about 0.002371 means that only 0.002371% of the variation in QTMARK is explained by ATR. Therefore it could be concluded that ATR has such a trivial effect on exam performance that it could even be omitted.

Two-tailed t-test for the significance of the slope b

H0: β = 0 (Proposition of revision lecture attended does not affect exam performance)

H1: β ≠ 0 (Proposition of revision lecture attended does affect exam performance)

Since the calculated t-test -0.13 is lower than the critical value of t-test at 5% significance level with 370 d.o.f, we fail to reject H0. This case and the conclusion is that revision lecture attendance does not affect exam performance.

(b) Multivariate Regression

**QTMARK** = b0 + b1ATR + b2ABILITY + b3HSQ + b4

Modeling by GLS, we get:

**QTMARK** = 56.32537 + 0.165849ATT + 0.84905ABILITY + 0.417685HSQ

Interpretation of the regression results:

- \(b_0\) represents the intercept of the model, and it is the expected value of QTMARK when the independent variables are zero.
- \(b_1\) represents the coefficient of ATR, which indicates the change in QTMARK for a one-unit change in ATR, holding all other variables constant.
- \(b_2\) represents the coefficient of ABILITY, which indicates the change in QTMARK for a one-unit change in ABILITY, holding all other variables constant.
- \(b_3\) represents the coefficient of HSQ, which indicates the change in QTMARK for a one-unit change in HSQ, holding all other variables constant.

L1 Chinese writer


These qualitative variables would have some impacts on the QTMARK, but whether they are statistically significant will be investigated later.

Correlation matrix\(^1\)

- The QTMARK has strong positive relationships with variables ABILITY, ALEVELSA, ATTC, ATT, and strong negative relationships with variables EXPLC, TOPS.
- There is no correlation between ATTC, ATT, and ATTR are very high. Therefore, multicollinearity is an issue needed to think about later on.
- EXPLC has strong negative relationships with ATTC, ATT, and ATTR and strong positive relationships with TOPS. This could be explained that students who got drunk cannot get up early to attend lectures and classes. The more TOPS they attend, the more money they spend on alcohol.
- Generally, HSQ has a strong positive relationship with ALEVELS and ALEVELSA, indicating a negative relationship with ABILITY and TOPS. This could be explained that the more TOPS students went for, the fewer hours they spent on studying and the higher ability a student has the less study is required for him/her. On the other hand, students having a good A-level record maintained their hardworking attitude.

L1 English writer

**Question 2b**

**Interpretation of results (equations 4 and 5 appendix 2).

The coefficient on class attendance is 0.13, which implies that holding all other variables constant, if you increase class attendance by 1 unit (1% increase in class attendance in a year), then the exam mark will increase by 0.13 units (0.13% increase in your mark). The coefficient on lecture attendance is 0.06, meaning holding all other variables constant, attending 1% more lectures will increase your mark by 0.06%. The coefficient on revision lecture attendance is slightly surprising, at -0.04, implying that attending 1% more revision lectures, your mark will decrease by 0.04%. The intercept can be interpreted to mean that if you attended no classes, revision or standard lectures, you would score 49.35%.

**Tests** (shown in appendix 2):

The coefficient on class attendance was significant at the 0.01% level implying that in the multiple regression model, class attendance has a significant impact on test mark. The coefficient on lecture attendance was not significant at the 10% level, implying that lecture attendance does not have a significant impact in a multivariate framework. However, lecture attendance does not appear to have a reasonably high correlation with class attendance, so the regression may be suffering from multicollinearity, which has made the result not significant. However, multicollinearity can be occurring such another factor being ‘unhelpful’ for it to have a negative impact on the regression. The coefficient on revision lecture attendance was significant up to the 1% level, thus implying that while we can fairly sure that revision lecture have a significantly negative impact, there is scope for the fact that the null hypothesis is indeed correct (type I error) and that the result is not significant.

The F-test for the joint explanatory power of the independent variables yielded an F-statistic of 13.07. This is significant at the 0.01% level as it exceeds the critical value of 3.78. Hence we can reject the null hypothesis given in the appendix. This means that the explanatory variables have made a significant joint contribution to exam performance.

**Question 3**

To investigate whether there are differences in performance between the sub-sample of 2002 students and previous year’s students, I have created intercept dummy variables and added them to the original equation, as shown by equations 1 and 2 in appendix 3. The first equation is known as the restricted equation, as opposed to the unrestricted model in equation 2, because it imposes the F-test null hypothesis (see hypothesis 4, appendix 3) on equation 2. Hence in equation 2, the intercept is allowed to vary whereas it is not allowed to equation 1 and is assumed to be constant in all years.

**Interpretation of coefficients**

The intercept in equation 3 can be interpreted as before, meaning that if you attended no lectures and had no A’s at A level, you would score 56.97. This is slightly infeasible in the sense that you would not have got on course if you did not score any A’s at A level. The coefficient of -0.14 on lecture attendance means that if you attended 1% more lectures you would get 0.14 out of 100 more in the exam ceteris paribus. The coefficient of 0.04 on A’s at A level means that if you got an extra A at A-level you would get 0.04 more ceteris paribus. The dummy variables in this case have a slightly different interpretation.Basically they say how much the intercept will move up or down compared to the omitted category, the year 2000 students.

The dummy variable coefficient of -1.19 means that if you are a 1999 student, you would score a proportion of 1.19% less than if you were a 2002 student. The coefficient of -0.85 on the 2001 dummy variable means that you would score a proportion of 0.85% less than if you were a 2002 student. Finally the coefficient of -8.51 on the 2000 dummy variable means that you would score a proportion of 8.51% less than if you were a 2002 student. These are shown in equations 3 to 6.

---

1 As reported in question 1, the correlation coefficient was 0.67.

2 Refer to “Correlation matrices for these quantitative variable” in the Appendix.

3 Refer to the Appendix for the Multivariate equation results table.
Outline

1. Background

2. Corpus data and methods

3. Findings
   - from corpus linguistic keywords and counts
   - from comparison of pairs of assignments
   - from interview data

4. Implications
Interviews with lecturers

Importance of visuals

• Diagrams and formulae are ‘the spine of the essay’ (Economics)
• The ‘challenge’ is ‘to marry the diagrams with the text’ (Economics)
• Including visuals helps students gain better marks as it avoids having to describe and introducing errors (Biology)
• ‘there is no existing document out there which explains how to interpret their data’ (Biology)
• Marks for presentation may include the assessment of diagrams, tables and overall layout (Engineering).

Being concise

• Lecturers value writing which is ‘clear and concise’, and ‘succinct’ and dislike ‘verbosity’ (Engineering)
• Preference for ‘precision, incision, concision’ (Economics)
• ‘there’s never been a penalty for an essay that’s too short’ (Biology)
Summary

• Chinese students make significantly greater use of visuals and lists than British students
• All BAWE assignments have been judged proficient
• \(=>\) different, yet equally valued, ways of writing

Possible reasons

• Are visuals and lists used as a strategy to meet the challenge of producing extended pieces of writing in unfamiliar genres in L2?
• Perhaps some students are more visually-oriented? – Chinese languages are more visual…
• Use of visuals is highly valued in particular disciplines e.g. Engineering, Economics, Biology. Since Chinese students tend to study hard-applied disciplines more than soft-pure disciplines, maybe the use of visuals and lists crosses over??
Outline

1. Background

2. Corpus data and methods

3. Findings
   - from corpus linguistic keywords and counts
   - from comparison of pairs of assignments
   - from interview data

4. Implications
Embracing different semiotic modes

• Yet…. ‘graphic literacy’ is seldom taught in EAP classes – *why?*

• Most applied linguists are ‘trained in the humanities, where words are central to disciplinary values and argumentation’
  Johns (1998:183)

• There’s often a concentration on ‘linear text’ (Johns, 1988: 183) rather than on the interaction of visuals with text.

• Tutors may ‘find themselves relying on disciplinary norms they are familiar with’ (Gardner and Holmes, 2009: 251)

• ‘for students who face the challenge of writing extended, factual, evidence-based, and disciplinarily specific texts, there is still relatively little on the market’.
  Tribble (2009, p. 416) in a review of EAP textbooks
Implications for practice: EAP tutors

- remain open-minded as to what might be acceptable within unfamiliar disciplines and genres;
- include ‘graphic literacy’ in academic writing classes;
- search corpora (e.g. BAWE, MICUSP) for particular discipline features to identify ways in which these vary;
- collect exemplars of the kinds of writing their students are asked to produce;
- move beyond lexicogrammatical considerations such as the acceptability of *I* or the choice of passive or active voice to considering assignments holistically (e.g. Is it ok to use a table to display results or should these be given in prose? Can the conclusion be presented as a bulleted list? If images are given, can a lengthy caption be included?);
- encourage students to question their discipline tutors;
- work with discipline tutors to undertake the following strategies ….
Implications for practice: discipline tutors

• provide undergraduate students and EAP tutors with explicit guidance as to what is required in assignments, particularly within unfamiliar genres;

• give exemplars and accompanying commentary to illustrate possible assignment responses - providing several examples gives a sense of the range of acceptability permitted;

• attempt a new assignment themselves in order to see where the rubric fall short;

• allow dedicated time within lectures for students to interview them about the assignment;

• avoid frequent misunderstandings of rubric by adding clarification in subsequent years.
Implications for practice: students

• research writing in their discipline (cf. Johns', 1997, plea for students to become researchers of their disciplines' practices);

• ask questions of both EAP and discipline tutors as to what is, or might be, considered proficient writing;

• seek out exemplars of writing of the type they are being asked to produce (e.g. ask tutor to provide previous cohort’s assignments);

• talk about writing with peers – what is expected? What will this assignment look like? (layout, use of tables/graphs/images/lists).
References