Spreading the Word: the Challenge of the Corpus as an Agent of Change

Hilary Nesi*

Department of English and Languages, Coventry University, Priory Street, Coventry CV1 5LW, UK

Abstract

Analysis of the British Academic Written English (BAWE) corpus has revealed similarities and differences in student writing across disciplines, levels of study and ‘families’ of genres that have implications for assessment practices across the academy. So far, however, findings from the corpus have only been published in papers addressing fellow corpus linguists, and have been expressed in ways incomprehensible to many teaching staff and students outside our discipline. Now a follow-on project has led to the creation of corpus-based academic writing activities for the British Council LearnEnglish website. This paper describes the process of selecting and presenting corpus evidence for a new readership, linking traditional CALL-style exercises to genre- and discipline-specific Sketch Engine displays and our ‘Wordtree’ corpus interface. It is hoped that the materials will help to bring about changes to the way academic writing is taught, as well as stimulating further corpus-driven research.

Keywords: academic writing; BAWE corpus; CALL; genre analysis; EAP; materials development

1. Introduction

The project 'An Investigation of Genres of Assessed Writing in British Higher Education', funded by the UK Economic and Social Research Council (ESRC), aimed to identify and describe the kinds of texts university students actually write. It involved collecting proficient student assignments from different disciplines and levels of study, and analysing these, alongside course documentation and interview responses from academic staff and students, in order to identify ‘genre families’ of university student work. The resulting British Academic Written English (BAWE) corpus contains just under 3000 of these assignments, marked up for genre family, discipline and level, and for part of speech and structural features (see Alsop & Nesi, 2009; Nesi & Gardner, 2012).

* Corresponding author. Tel.: +44 (0)2476 888984; fax: +44 (0)2476 888679
E-mail address: h.nesi@coventry.ac.uk
The BAWE corpus has since been analyzed from many perspectives, revealing similarities and differences in student writing across disciplines, levels of study and genres that have implications for university assessment practices. These findings were not predictable – as lecturers, writing tutors, textbook writers, academics and/or students we all know something about some kinds of student writing, and it is easy for us to jump to the conclusion that the types of writing we are familiar with are the ones that others also know and use. It is this lack of shared understanding about the varied purposes and linguistic features of university assignments that leads to problems with fair assessment: writing tutors often overlook important minority genres, lecturers do not make explicit the requirements of the assignment tasks they set, and students coming from different social, cultural or disciplinary backgrounds sometimes fail to recognize the conventions that their lecturers require.

Despite the need for wider dissemination, prior to the Writing for a Purpose project findings from analysis of the BAWE corpus had been published almost exclusively in papers addressing fellow linguists, and had been expressed in ways incomprehensible to many teaching staff and students outside our discipline. Thus, in order to present what has been discovered about British university student writing in ways that could be reached by a much broader readership of teachers and learners, members of the original BAWE project team approached the British Council, an organization ‘specializing in international educational and cultural opportunities’. With their support, and follow-on funding from the ESRC, the Writing for a Purpose materials were developed for their LearnEnglish website (http://learnenglish.britishcouncil.org/en/). This site is used by many English language teachers and learners around the world.

1.1. The genre families and their purposes

The Writing for a Purpose materials focus on the genre families in the BAWE corpus. Thirteen genre families have been identified, according to their social purposes, their generic stages, and the networks they form with other professional and academic genres (Nesi & Gardner 2012; Gardner & Nesi, 2013). These are:

1. **Case Studies**, which demonstrate and develop understanding of professional practice through the analysis of an exemplar, for example a company in need of business advice or a patient in need of treatment
2. **Critiques**, which demonstrate and develop understanding of an object of study, such as a literary text, and the ability to evaluate it and assess its significance
3. **Design Specifications**, which demonstrate and develop the ability to design a product or procedure that can be manufactured or implemented
4. **Empathy writing**, which demonstrates and develops understanding and appreciation of the relevance of academic ideas by translating them into a non-academic register, for example through the media or a public information leaflet
5. **Essays**, which demonstrate and develop the ability to construct a coherent argument and employ critical thinking skills
6. **Exercises**, which provide practice in key skills (e.g. the ability to interrogate a database, perform complex calculations, or explain technical terms or procedures), and consolidate knowledge of key concepts
7. **Explanations**, which demonstrate and develop understanding and the ability to describe and/or account for the significance of an object of study, for example in the manner of a textbook or an encyclopedia entry
8. **Literature Surveys**, which demonstrate and develop familiarity with literature relevant to the focus of study
9. **Methodology Recounts**, which demonstrate and develop familiarity with disciplinary procedures, methods, and conventions for recording experimental findings, for example as in a laboratory report
10. **Narrative Recounts**, which demonstrate and develop awareness of motives and/or behaviour in individuals or organizations, for example through a reflective piece or a short story
11. **Problem Questions**, which provide practice in applying specific methods in response to simulated professional problems, most often in the field of law
12. **Proposals**, which demonstrate and develop the ability to make a case for future action, for example in a research proposal or business plan
13. **Research Reports**, which demonstrate and develop the ability to undertake a complete piece of research, including its design and an appreciation of its significance in the field.
These genre family groupings aim to include every type of assignment submitted to the BAWE corpus. They were arrived at by reading and rereading all the assignments, noting their statements of purpose and structural features, and referring to information extracted from lecturers regarding degree programme writing requirements. The names of the genre families do not reflect the readers’ and writers’ names for their assignments, however, as academics and students generally make far fewer distinctions between text-types (Nesi and Gardner 2006). In fact, in the departments we investigated the terms ‘essay’ and ‘report’ were often used almost interchangeably to cover a range of different genres. This of course contributes to the confusion felt by many students when faced with a new type of writing task, not clearly distinguished from previous tasks but seeming to require a different perspective and organizational approach.

All 13 genre families are represented in the Life, Physical and Social Science components of the BAWE corpus – we did not receive any Case Studies or Problem Questions from Arts and Humanities. The most widely used genre families, Essays and Critiques, occur in all 24 of the disciplines represented by 50 or more assignments. The least widely used genre families (Design Specifications and Problem Questions) occur in seven different disciplines.

1.2. Writing for a Purpose

The assignment types can also be grouped into a smaller number of categories, to reflect their broad educational and social purposes. Five groupings were identified in this way; they shed light on reasons for writing at university level:

1. to demonstrate knowledge and understanding (Explanations and Exercises)
2. to develop powers of informed and independent reasoning (Essays and Critiques)
3. to develop research skills (Literature Surveys, Methodology Recounts and Research Reports)
4. to prepare for professional practice (Case Studies, Design Specifications, Problem Questions and Proposals)
5. to write for oneself and others (Narrative Recounts and Empathy Writing).

The Writing for a Purpose materials created for the British Council website begin by inviting teachers and students to consider their own writing in terms of these five broad functions. Those who progress through the exercises in a linear way are led to examine relevant genres in ever-increasing depth.

2. The website and the authoring tool

When the Writing for a Purpose project started, the LearnEnglish site offered almost no material for academic purposes apart from IELTS preparation work. Most sections of the site favoured ‘fun’ activities that fitted easily into traditional language teaching syllabuses.

LearnEnglish uses a bespoke authoring tool which generates CALL style exercises of the following types:

- True or false
- Multiple choice
- Multiple selection
- Gap fill drag and drop
- Gap fill typing
- Reordering vertical
- Reordering horizontal
- Matching
- Grouping
- Error correction

All of these exercises are essentially matching activities, and the present version of the authoring tool does not allow for the creation of more open-ended exercises, or ‘noticing’ activities which might, for example, require learners to highlight sections of text that contain certain features or realize certain functions. The authoring tool only permits generic feedback, and there is no facility to provide comments on possible alternatives, or explain why particular answers are incorrect. A forum facility enables users to post comments, but the site does not yet offer any other interactive social media features.
Of particular relevance to this project is the fact that the site is intended to be used with screens of all different sizes, from tablets to whiteboards. Materials for use with mobile devices can maximize the number of users and the frequency of use, but they work best if the texts they refer to are short, and if rubrics and commentary are also kept to the minimum. In fact, the Writing for a Purpose exercise template discourages activities at discourse level (i.e. beyond the paragraph) and is most appropriate as a means of focusing on lexical items and very short pieces of writing.

In the early stages of the project the team invited key stakeholders - mostly EAP teachers and materials writers – to a meeting to talk about what they expected from the materials. Most of their suggestions assumed that the interface would be more sophisticated, with the capability to open comment boxes by moving the cursor across the screen, for example, or to enable users to create their own individualised glossaries. Some of their ideas, however, offered ways to circumvent the restrictions of the authoring tool, for example by incorporating films of lecturer-tutor interactions, or lecturers’ audio commentaries on genres in their fields.

Multimedia files offer variety and interest, and we also found them to be an excellent way of presenting additional text without disadvantaging small screen users. Some extracts from interviews with lecturers have been included in the materials, but as has been noted in previous studies (Haggis, 2006; Elton, 2010; Turner 2011) lecturers’ knowledge of generic conventions is often tacit and automatic, making it difficult for them to explain their thought processes and expectations. Therefore in order to provide more explicit accounts of specific genres we sometimes drew on student discussions of writing tasks, and also included audio-recorded dialogues presented as if they were spoken by a lecturer and a student, but actually delivered by members of the team.

A further attempt to circumvent the restrictions of the authoring tool involved creating links to the open Sketch Engine site (https://the.sketchengine.co.uk/open) so that users can interrogate relevant parts of the corpus for themselves. They need not leave the LearnEnglish site to do this, as the Sketch Engine interface is embedded on the site, within a frame.

We used pictures to represent genre families and writing purposes. These improve the appearance of the pages and may also make key concepts more memorable for some users. Undergraduate students in the School of Art and Design at Coventry University are required to produce logos and symbols in response to customer briefs, so we wrote a design brief for lecturers to use in class, and eventually commissioned one of the best students to produce the icons for our site. Fig.1 shows from left to right our icons for the five social purposes of university student writing: demonstrating knowledge and understanding, developing powers of informed and independent reasoning, developing research skills, preparing for professional practice and writing for oneself and others.

Fig. 1. Icons illustrating the purposes of student writing

3. The materials

The materials are currently organized in four main sections, following on from a one-page introduction:

- Primary Purposes
- Genre Families
- Your Writing Task
- Disciplines
A menu at the side of the screen allows users to move between sections, and to click on a section to view its contents in more detail. An additional page provides information about the materials writing team, and there is also a section for teachers containing lesson plans and word lists for every genre family and discipline.

In the ‘Primary Purposes’ section each of the main writing purposes is introduced using a short dialogue accompanied by gap-fill exercises. These are followed by matching exercises. The ‘Genre Families’ section explains each of the genre families in more detail, with further dialogues and exercises to match excerpts from assignments to their appropriate genre families.

The most innovative exercise types are to be found in the section ‘Your Writing Task’. It is here that we provide examples of assignments in each genre family, information regarding their basic structure, and tasks relating to their linguistic features. In the ‘vocabulary’ sections for each genre family learners are introduced to useful phrases, and can see how these phrases are used in genres associated with specific disciplines. For example a criterial feature of Case Studies is that they contain some sort of recommendation for future action and management. The n-gram function in Wordsmith Tools was used to identify the phrase frame ‘It is X that’, commonly used to introduce recommendations within Case Studies in the BAWE corpus, and this phrase frame is linked to 40 concordance lines in Sketch Engine, specified using the CQL query function to include only Case Studies and exclude examples from bibliographies and footnotes. The relevant Sketch Engine page is embedded within the Writing for a Purpose page, and is configured to show the discipline and genre family of the text represented in each concordance line. Table 1 shows the Sketch Engine output for ‘it is important/recommended/ suggested/vital that’ in business Case Studies.

Table 1. Recommendation n-grams for business Case Studies

<table>
<thead>
<tr>
<th>However, since</th>
<th>it is important that</th>
<th>the firm meets it needs they should be kept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Therefore,</td>
<td>it is important that</td>
<td>an accurate estimation of the duration of each</td>
</tr>
<tr>
<td>Therefore,</td>
<td>it is important that</td>
<td>the key project stakeholders are identified and classified</td>
</tr>
<tr>
<td>Having decided on the most profitable product mix,</td>
<td>it is important that</td>
<td>the firm now develops a clear production schedule</td>
</tr>
<tr>
<td>Therefore,</td>
<td>it is important that</td>
<td>Yahoo focuses its attention not only on like-for-like</td>
</tr>
<tr>
<td>Hence,</td>
<td>it is recommended that</td>
<td>the project pace should be of a medium one. This</td>
</tr>
<tr>
<td>However,</td>
<td>it is recommended that</td>
<td>other methods, such as Payback, ARR are used at</td>
</tr>
<tr>
<td>Besides that,</td>
<td>it is recommended that</td>
<td>the review of forecast by top management should be</td>
</tr>
<tr>
<td>Hence,</td>
<td>it is recommended that</td>
<td>sales and manufacturing units are treated as follows</td>
</tr>
<tr>
<td>Therefore,</td>
<td>it is suggested that</td>
<td>manufacturing unit is treated as a pseudo-profit</td>
</tr>
<tr>
<td>However in this case</td>
<td>it is vital that</td>
<td>they are well-informed of the changes that will be</td>
</tr>
</tbody>
</table>

The materials thus offer a form of data driven learning (Johns 1991), under conditions where concordance output is pre-selected, yet learners also have the freedom to pursue their own corpus investigations. The information in Sketch Engine is used as the basis for exercises created with the authoring tool. Learners can also access further information about assignments represented by concordance lines (for example the title, module, and level of study), and, most usefully, click on the keywords in the central column to view additional context. The expanded context for the last line in Table 1, for example, is shown in Fig. 2. Usually a few lines of text are sufficient to indicate how a word or expression might be used in the performance of a given generic move.
We imagine that the majority of users will be unfamiliar with corpus query tools, so we have created some help pages to show them how to make the most of the lines we have selected (see Fig. 3).

A further innovation is the use of the Wordtree interface (see Fig. 4) a kind of interactive visualization which groups concordance lines into tree-like branching structures which users can ‘grow’ by clicking on word nodes. Tree structures provide the same amount of information on screen as concordance lines, but paradigmatic patterns are more immediately apparent, the amount of text is reduced, and users may be more tempted to interact with the corpus output.

Fig. 2. Expanded context for a business Case Study in Sketch Engine

Fig. 3. Help page for a business Case Study in Sketch Engine
4. Conclusion

At the time of writing this article the *Writing for a Purpose* materials had not yet been made live on the British Council *LearnEnglish* site. Adjustments will almost certainly have to be made to the format and the content in the coming months, in response to user feedback. It will be interesting to see how users respond to our data-driven approach to the study of academic written genres. The corpus is a very rich source of information that is relevant to learner writers, but much work remains to be done to extract this information and convey it to learners in an interesting, memorable, and comprehensible way.

Acknowledgements

The project 'An Investigation of Genres of Assessed Writing in British Higher Education' was funded by the Economic and Social Research Council (ESRC) between 2004 and 2007 (project number RES-000-23-0800). The ESRC granted follow-on funding from March 2012 to April 2013 to draw on the original findings and develop academic writing materials (project number ES/J010995/1). More details of these projects can be found at www.coventry.ac.uk/bawe. The team members for the *Writing for a Purpose* project are Sheena Gardner and Hilary Nesi (genres and corpus), Tim Kelly (video and multimedia), Alex Woolner (Coventry Serious Games Institute), Elly Hutchins (art work), Andy Gillett (materials developer) and Adam Kightley (British Council).

References


