

USE – The Uppsala Student English Corpus: An instrument for needs analysis

*Margareta Westergren Axelsson
Uppsala University*

At the Department of English, Uppsala University, a corpus of learner English is currently being compiled. The project has been running for a year – the two academic terms of 1999 – and is planned to proceed for another two years. The corpus consists of essays written by the department's students of English, the majority in their first term of full-time studies of the subject. The project has attracted a large number of participants, who hand in their 700-word essays on diskettes or online to project leaders Margareta Westergren Axelsson and Ylva Berglund. The present size of the corpus is about 600,000 words. At the current rate of student participation, we can look forward to a collection of texts comprising close to two million words by the end of the project period.

Unlike other corpora of learner English, which mainly serve purposes of research into the nature of learner English, the USE Corpus has a double aim. One aim, perhaps the most important one for the department, is for the corpus to serve as an instrument for diagnosing the language difficulties our students experience at different levels. To have access to such an instrument is of great value in times of shrinking resources for language studies. Swedish university students of English normally reach a high level of spoken proficiency and good communicative competence even in upper secondary school. What is lacking is a good command of the written medium, with its demands on greater precision, variation and stylistic awareness. This observation, however, does not entail knowledge of the exact extent and nature of the problems experienced by students in their writing. The USE Corpus can provide us with a more accurate picture, revealing the details that are important for instruction focused on the most pressing needs.

Of course, a needs analysis cannot be carried out without some systematic research. Therefore, another aim is to carry out didactic research on the corpus material. So far, a half-time position is allocated to corpus collection and

research. In addition, students in their third and fourth term will be engaged. At the time of writing—early 2000—the project leaders have presented a few papers related to the corpus, one on corpus annotation and tag frequencies (Berglund and Prütz, presentation at the ASLA conference KORFU in Växjö, November 11–12, 1999) and another on subject-verb concord (Axelsson, forthcoming).

The essays contained in the corpus are essays written as part of the regular curriculum. They cover set topics of different genres. The following list shows the topic areas of the essays constituting the most comprehensive part of the corpus, viz the essays handed in by first-term students:

1. 'English. My English'. A personal essay in which students are asked to describe their relationship with the English language and evaluate their strengths and weaknesses in the four language skills of reading, writing, speaking, and listening. Format: About 700 words.
2. 'Argumentation'. Students argue for or against a statement concerning a topical issue. Format: About 700 words.
3. 'Reflections'. Against the background of an article about filtering out violent TV programmes, students reflect on aspects of violence on TV or in films, on censorship, parental responsibility or other related issues of their choice. They also produce a summary of at most 100 words in addition to the customary 700-word essay.
4. Literature course assignment. Students choose between a discussion of theme, character, narrator and analysis of the use of a set passage in a novel. Detailed instructions are given about using accepted conventions for formal pieces of writing. Format: About 800 words.
5. Culture course assignment. Students are asked to study certain topics in a set of secondary sources and compose an essay using that material. The topics concern, for example, the education of women in the 19th century, the industrial revolution, slavery, and utopias. Format: About 1,000 words.

A special questionnaire coded in a separate data base provides background data for each student on age, sex, first language, educational background, and time spent in English-speaking environments.

This brief note on the USE Corpus has outlined its present size and contents, its aims and foreseeable future development. A more detailed description is given in Axelsson and Berglund (forthcoming). It is our hope that, with time, the USE Corpus will be an asset not only to teachers and students at the Department

of English at Uppsala University but also to the international research community interested in learner English.

References

- Axelsson, Margareta Westergren. 1999. Project USE (Uppsala Student English). *ASLA Information* 25:2, 25-26.
- Axelsson, Margareta Westergren (forthcoming). The use of a corpus of students' written production in university English teaching. Paper presented at the ASLA conference KORFU, Växjö, November 11–12, 1999.
- Axelsson, Margareta Westergren and Ylva Berglund (forthcoming). The Uppsala Student English Corpus (USE): A multi-faceted resource for research and course development.
- Berglund, Ylva and Klas Prütz. Tagging a learner corpus – a starting-point for quantitative comparative analyses. Paper presented at the ASLA conference KORFU, Växjö, November 11–12, 1999.