Focusing on learning outcomes: using corpora at a university of technology

Andreas Eriksson

Chalmers University of Technology, Centre for language and communication - Sweden

The present paper describes the objectives behind a work-in-progress project on the implementation of corpus material in language courses at Chalmers university of technology, Göteborg, Sweden. Chalmers offers a challenging environment with students who are not language students primarily. The courses given by the Centre for language and communication are both elective courses and obligatory courses on a wide range of engineering programmes. The project covers three courses with quite different objectives and involves both written and spoken language. A central argument in the paper is that the implementation of corpora for teaching purposes in ESP environments can be facilitated if it is anchored in identified learning outcomes. The learning outcomes of the three courses and how corpora might contribute to the fulfilment of these outcomes are discussed in the paper.

Keywords: corpora, learning, ESP, learning outcomes

Introduction

The value of corpora for language teaching purposes as well as the limitations of corpora for such purposes are issues that have been discussed and commented on by many scholars (e.g. Chambers 2007:6-7, Gabrielatos 2005:25, Johansson 2007:26, Lee and Swales 2006:57, O'Keefe, McCarthy and Carter 2007:246-247). One point that has been made explicitly by several authors is that corpora are no magic wands that inevitably generate language learning (see e.g. Conrad 2000:548, Gaskell and Cobb 2004:315, Mauranen 2004:103). There is thus generally great awareness of the problems and limits of corpora for teaching purposes among teachers and researchers. Still, since one of the major strengths of corpora is that they lend themselves to providing students with new and innovative types of language input, there is always a

risk that emphasis is put on instructional input rather than learning outcomes. There is obviously no inherent contradiction between improved input and learning outcomes, but without basing the use of specific material in particular learning outcomes, the value of the material may not be as strong as it could have been.

The aim of the present work-in-progress report is to account for the early stages of a project where particular learning outcomes have been identified in three courses and to describe in what way corpora are believed to help students reach these outcomes. The idea is thus to base the use of corpora in learning outcomes and make the use of corpora outcome driven.

The study is carried out at the Centre for language and communication (CLC) at Chalmers university of technology and involves three different courses and EFL/ESP students from several engineering programmes. CLC has long experience of providing English proficiency courses, academic and technical writing courses as well as of integrating communication practice into engineering education (see e.g. Börjeson et al. 2007, Carlsson and Wranne 2008, Evertsson et al. 2007). However, the centre has comparatively little experience of using corpora or corpus-related material in their teaching, and the implementation and adaptation of such material into various courses thus partly means breaking new ground. However, it should be emphasised that the decision to use corpora stems from the identification of particular learning outcomes. The three courses described below are courses where certain learning outcomes were identified and where it is hypothesized that corpora could enhance learning.

Since students at Chalmers are students of engineering, language is not their major subject at university. It is likely that many of these students differ from language students in terms of motivation, objectives and familiarity with terminology. Consequently, the students that participate in the present investigation are different from the students in most other studies which have dealt with the use of corpora for teaching purposes. Students in such investigations have typically been university language students and the studies have been carried out as action research projects at departments of language at universities (Chambers 2007:7-8, Mauranen 2004:90-91).

Description of courses and desired learning outcomes

The first course to be investigated is an academic writing course open to all doctoral students at Chalmers. The course has been given for several years and covers the writing of texts common in academic writing, for example abstracts, research articles and conference papers. Overall the course has been very successful but one learning outcome that could be strengthened is students' knowledge about writing in their own disciplines, i.e. what is often referred to as genre knowledge. Genre knowledge is diverse in nature and involves several features. Hyland (2004:84), for instance, recognizes eight major aspects of genre knowledge. One of these is knowledge of grammar and phraseology. Corpora should be good sources for enhancing this particular aspect of genre knowledge, as they can give information about how and where particular words and phrases are used. The strengthening of this aspect would have to involve active participation from students, and therefore the collection of both individual and discipline specific corpora, in line with the study carried out by Lee and Swales (2006), are seen as potentially useful activities. As a result, the work will have to include the use of text analysis software, such as WordSmith Tools in order to facilitate comparisons of for instance collocations and clusters (Scott and Tribble 2006).

The second course is an elective proficiency course primarily aimed at students at basic or lower intermediate level. It involves both spoken and written proficiency but there is a focus on spoken proficiency as the course is supposed to help international Master's students during their first year at Chalmers. Master's courses at Chalmers are often project-based and thus involve a great deal of both formal and informal spoken English (all Master's courses are taught in English). The fairly general learning outcome of the spoken part of the course is to prepare students for future studies at Chalmers, particularly in terms of improving their ability to participate in discussions and give oral presentations. The material that corpora can provide, and which has not been used at Chalmers previously, consists of patterns common in spoken English. The effect of teaching such patterns, often referred to as formulaic sequences (Wray 2000, Meunier and Gouverneur 2004), is not fully clear but the arguments that are usually put forward are that the use of formulaic sequences can free processing capacity and that they can be

beneficial for students in handling speech events (Ellis 1996, Wray 2000, Wray and Perkins 2000, Mauranen 2004). The corpus material in the present course will most likely be corpus-based and taken from sources like O'Keefe, McCarthy and Carter (2007), who list a number of formulaic sequences (referred to as 'chunks' by O'Keefe, McCarthy and Carter 2007:65-67) common in spoken language. The advantage of using this type of material is that it consists of common and naturally occurring sequences of spoken language. In other words, students are shown examples of language as it is actually spoken. The learners' use of chunks in both formal and informal university settings can be compared with material from a corpus like the Michigan Corpus of Academic Spoken English (MICASE) (Simpson et al. 2002).

The third course is called *Safety communication* and is given at the programme of nautical science. In the course, students learn to use a set of phrases and terminology published by the International Maritime Organization (IMO) as the IMO Standard Marine Communication Phrases (IMO 2002).[1] The phrases constitute one aspect of Maritime English and have been developed in order to make communication at sea as simple as possible so that also seamen with limited knowledge of English can communicate effectively in commonly occurring situations at sea and in harbour. The number of phrases is limited and all of them have been compiled and published in one single document. As a consequence, it might at first seem fairly clear what the students need to learn. Not surprisingly, however, it has turned out that knowing the phrases is seldom enough in real life situations. In these situations, the user often needs to apply language structures and vocabulary that extend far beyond the limits of the IMO phrases. One particular aspect of the need for a general knowledge of English concerns the cotext surrounding the fixed IMO phrases. Minor investigations have shown that the cotext of IMO phrases often contains formulaic sequences, but that both students and teachers are unaware of these sequences. The hypothesis is therefore that knowing more about the cotext of the IMO phrases would help students use these phrases more correctly and more effectively. It seems as if not only the phrases but also the words co-occurring with these phrases are fairly fixed, and being able to clarify the relationships

---

[1] Examples of IMO phrases are: "You are proceeding at dangerous speed" and "Stand by for assistance" (IMO 2002).

with phrases and their cotext is believed to help students in their use of the IMO phrases. This part of the project is more extensive than the first two parts described, since it involves the identification of formulaic sequences as well as decisions about how the use of sequences can be learnt.

Critical factors

All three courses come with a number of critical factors that may hinder learning and which therefore have to be addressed. *Time* is one such factor. Both the doctoral student course and the safety communication course are intensive and run over seven weeks only. Several studies have emphasised the need for extensive training periods if students are to use corpora independently (e.g. Gaskell and Cobb 2004, Johansson 2007:25). Considering such findings, it is likely that the three courses in the present project will be corpus-based or corpus-informed rather than corpus-driven.

Another critical factor is students' level of English, particularly in the proficiency course. There is some doubt as to whether corpus material is useful for low-level learners or not, particularly when it comes to learner consultation of corpora (Chambers 2007; see also O'Keefe, McCarthy and Carter 2007:24). This is thus yet another factor which supports the use of selected corpus material rather than data driven learning in order to reach the learning outcome identified.

The third critical factor is *comparability* and primarily concerns the doctoral student course. If the students are to collect a mini-corpus of their own work and compare that with other material from their field, it is important that the material is comparable. It may not be possible to monitor this process fully and it is therefore a factor that needs to be considered when evaluating the learning outcome.

Evaluating the learning outcome is in itself perhaps the most difficult part of the whole process and a well-known problem in research on learning (Barr and Tagg 1995). The use of spoken language is particularly difficult to evaluate for a number of reasons. First of all, it involves recording, which in itself may affect participants. Secondly, a great

deal of material is needed to evaluate the use of particular phraseology, and thirdly, transcribing recorded material is tedious and time-consuming work.

Another difficult aspect of the effects of corpus material and corpus-related teaching is to determine what generates learning. However, even if it is not always possible to show that a particular type of methodology results in a particular type of learning, it might be possible to show that a course where corpora have been used has resulted in particular learning outcomes. This might be a reasonable first step to indicate that corpora facilitate learning. The results can then be used to change certain parameters in order to be able to gain more knowledge about learning effects (cf. Gaskell and Cobb 2004:315-317).

Conclusion

Meunier and Gouverneur (2007:132) aptly show how general learning processes can be linked to exercises for the learning of formulaic sequences. This type of linking and the emphasis on learning outcomes made in the present paper may not be revolutionary approaches to corpora, but grounding the use of corpora in learning outcomes and general learning processes may strengthen teachers' and researchers' awareness of how corpora can be employed for teaching purposes.

Another insight from more general approaches to learning that seem to be worth remembering when dealing with corpora is Gardener's (1993:24) and Biggs'(2003:46) emphasis on the risks of trying to cover too much. According them, coverage is one of the greatest enemies of deep learning, as it prevents students from analysing the material carefully enough. This is certainly one of the risks of corpora as students can quickly be presented with a great deal of unanalysed material. Another aspect of this problem is addressed by Bowden and Marton (1998:24), who argue that "[f]or each phenomenon there is a limited number of critical aspects that can be discerned and focused on simultaneously. So differences in how this phenomenon is experienced reflect differences in what critical aspects are discerned and focused on simultaneously". Guiding inexperienced corpus users to focusing on critical aspects and

to noticing use in various situations seems to be essential if corpus material is to foster learning in many ESP environments. Corpora in ESP contexts often require a great deal of contextualisation, but if such contextualisation can be based in explicitly stated learning outcomes, corpora can be used for a variety of purposes and in many different ESP contexts.

References

Barr, R. and **Tagg, J.** 1995. "From teaching to learning — a new paradigm for undergraduate education." *Change* 27/6: 12–25.

Biggs, J. 2003. *Teaching for Quality Learning at University.* London: The Society for Research into Higher Education & Open University Press.

Börjeson, F., Eriksson, A-M., Erlandsson, J., Hillman, K., Molander, S. and **Rex, E.** 2007. "Ger koppling av kunskapsinlärning och färdighetsövning ökad djupinlärning? – Utveckling av kursmoment i "Miljö- och resursanalys för hållbar utveckling V2"." In *ESA rapport 2007:12*, Göteborg: Chalmers University of Technology.

Bowden, J. and **Marton, F.** 1998. *The university of learning. Beyond quality and competence in higher education.* London: Kogan Page.

Carlsson, C-J. and Wranne, O. 2008. "Using writing as a cognitive tool in a car design project". In *Proceedings of the 4th International CDIO Conference, Hoogeschool Gent, Gent Belgium, June 16-19, 2008.* Gent, Belgium

Chambers, A. 2007. "Popularising corpus consultation by language learners and teachers." In Hidalgo, E, Quereda and Santana, J. *Corpora in the foreign language classroom. Selected papers from the sixth international conference on teaching and language corpora (TALC 6). University of Granada, Spain, 4-7 July, 2004.* Amsterdam: Rodopi, 3-16.

Conrad, S. 2000. "Will corpus linguistics revolutionize grammar teaching in the 21st century? " *TESOL Quarterly* 34/3: 548-560.

Ellis, R. 1996. "Sequencing in SLA. Phonological memory, chunking, and points of order." *Studies in Second Language Acquisition* 18:91-126.

Evertsson, M., Bankel, J., Enelund, M., Eriksson, A., Lindstedt, P. and **Räisänen, C.** 2007. "Design-Implement Experience from the 2nd Year Capstone Course "Integrated Design and Manufacturing"." In *Proceedings, Third International CDIO Conference*, *June 11-14, 2007*. MIT, USA.

Gabrielatos, Costas. 2005. "Corpora and language teaching: Just a fling or wedding bells?" *Teaching English as a Second or Foreign Language (TESL-EJ)* 8/4:A-1

Gardener, H.W. 1993. "Educating for understanding". *The American School Board Journal*, 180/7: 20-24.

Gaskell, D. and **Cobb, T.** 2004. "Can learners use concordance feedback for writing errors?" *System* 32: 301-319.

Hyland, Ken. 2004. *Genre and second language writing*. Ann Arbor: University of Michigan Press.

IMO. 2002. *IMO Standard Marine Communication Phrases*. London: International Maritime Organization.

Johansson, S. 2007. "Using corpora: from learning to research." In Hidalgo, E, Quereda and Santana, J. *Corpora in the foreign language classroom. Selected papers from the sixth international conference on teaching and language corpora (TALC 6). University of Granada, Spain, 4-7 July, 2004.* Amsterdam: Rodopi, 17-28.

Lee, D. and **Swales, J. M.** 2006. "A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora." *English for Specific Purposes*, 25: 56-75.

Mauranen, A. 2004. "Spoken corpus for an ordinary learner." In Sinclair, J.M. *How to use corpora in second language teaching.* Philadelphia, PA: John Benjamins, 89-105.

Meunier, F. and **Gouverneur, F.** 2007. "The treatment of phraseology in ELT textbooks." In Hidalgo, E, Quereda and Santana, J. *Corpora in the foreign language classroom. Selected papers from the sixth international conference on teaching and language corpora (TALC 6). University of Granada, Spain, 4-7 July, 2004.* Amsterdam: Rodopi, 3-16.

O'Keefe, A., McCarthy, M. and **Carter, R.** 2007. *From corpus to classroom*. Cambridge: Cambridge University Press.

Simpson, R.C., Briggs, S. L., Ovens, J. and **Swales, J. M.** 2002. *The Michigan Corpus of Academic Spoken English*. Ann Arbor, MI: The Regents of the University of Michigan.

Scott, M. and **Tribble, C.** 2006. *Textual patterns. Key words and corpus analysis in language education*. Amsterdam: John Benjamins.

Wray, A. 2000. "Formulaic sequences in second language teaching: principle and practice." *Applied Linguistics* 21/4: 463-489.

Wray, A. and **Perkins, M. R.** 2000. "The functions of formulaic language: an integrated model". *Language and communication* 20/1: 1-28.

Andreas Eriksson *obtained his PhD of linguistics at Göteborg university in 2004 with the thesis* Tense and Aspect in Learner Writing. Advanced Swedish learners' use of tense and aspect in English argumentative text*. He has been working as a senior lecturer at the Centre for language and communication at Chalmers university of technology since 2007. His research interests are second language acquisition (SLA), genre studies, corpus linguistics, tense, and aspect.*