

# Comparing Real and Ideal Language Learner Input: The Use of an EFL Textbook Corpus in Corpus Linguistics and Language Teaching

Ute Römer (Cologne, Germany)

It is widely accepted that both native speaker corpora and non-native speaker or learner corpora can be of extremely high utility in foreign language learning and teaching. But beside analysing learner output and ideal learner input, i.e. the language actually produced by competent speakers of English, it might also be helpful to look at the input pupils really get in their English lessons.

In the present paper I would like to discuss how an electronic corpus of EFL textbook texts can be used to compare so-called "school English" (Mindt 1996: 232) with real English as used by native speakers, and in which way such a corpus may help us answer two crucial questions: "Do we teach our pupils authentic English, i.e. do we confront them with the same type of English they are likely to be confronted with in natural communicative situations?" and "What can we do to improve EFL teaching materials?" It will be shown that a textbook corpus (TC) can prove a useful resource if one wants to find out more about the status of authenticity in ELT. The importance of using authentic language in the EFL classroom has recently been stressed by Beaugrande (2001) and Granger who states that "... textbooks are more useful when they are based on authentic native English." (1998: 7)

The corpus used in this study consists of texts taken from volumes 1-6 of two textbook series most widely used in EFL teaching in Germany, *Green Line New* and *English G 2000*. The texts chosen are all supposed to represent spoken language. Entirely written material, such as narrative passages, letters or excerpts from novels, is not included in the corpus. Spoken texts were selected to enable a comparison with the spoken component of the British National Corpus (BNC) which is used as the main source of authentic data.

In the analyses of TC and BNC the focus is on aspects related to grammar teaching, and especially on some of the grammatical structures (German) language learners misuse or have constant problems with. Some very interesting findings result from sample queries on items that count among the grammatical problem areas in EFL teaching, such as modal verbs or conditionals. So far the rather significant mismatches that were found between BNC data and TC data rendered clear that the language used in textbooks does not mirror authentic language use. Thus further studies on lexicogrammatical phenomena in TCs are likely to show how an improvement of EFL textbooks (and teaching materials in general) can be brought about on the basis of native-speaker corpus data.

## References

- Beaugrande, R. de. 2001. 'If I were you...': Language Standards and Corpus Data in EFL. <http://beaugrande.bizland.com/Ifiwereyou.htm> (submitted to *TESOL Quarterly*).
- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.
- Granger, S. (ed.). 1998. *Learner English on Computer*. London: Longman.
- Mindt, D. 1996. English Corpus Linguistics and the Foreign Language Teaching Syllabus. In: Thomas, J. and M. Short. (eds.). *Using Corpora for Language Research*. London: Longman.
- Sinclair, J. 2001. Preface. In: Ghadessy, M., A. Henry and R. L. Roseberry. (eds.). *Small Corpus Studies and ELT. Theory and Practice*. Amsterdam: Benjamins.