Abstract

Learner corpus research is a growing field of linguistic inquiry, which holds great potential for investigations into second language acquisition as well as foreign language teaching, and provides new possibilities for the development of teaching materials. In recent years, many projects involving the collection of computer learner corpora have been launched, and thus a diversity of database programs, annotation schemes, and text retrieval software have evolved.

After a general introduction to the field, this study focuses on English and German learner corpora. On the basis of a survey of English and German learner corpora, the impact of those corpora on language teaching is evaluated. Furthermore, the differences between English and German learner corpus projects are established and analyzed, and future prospects for the collection and annotation of learner corpora are discussed.

The two principal challenges facing learner corpus research revealed by the analysis of the survey are the need for more standardization in learner corpus collection and annotation, and consequently the need for more collaboration among individual projects in a networked environment.