
Textlinguistik

Eine Datenstruktur zur Umschrift handschriftlicher Lernertexte¹

Sören Ohlhus

Abstract

This paper provides a data structure for the transcription of handwritten texts. The data structure is modelled as an xml markup language which allows to code additional information about the graphical arrangement of a text and thus to avoid the loss of those information in the course of governing corpora of manuscripts with the computer. In particular, the presented markup language focuses on phenomena in learner texts (e.g. revisions and embedded pictures). Furthermore it allows annotating information about the writing process and that accounts for varying authors of an original text.

1 Das Problem und Umriss seiner Lösung

Im Umgang mit handschriftlichen Lernertexten, insbesondere von Schreibern im Grundschulalter, hat man es mit einer Fülle von Phänomenen zu tun, die es erforderlich machen, über inhaltliche und im engeren Sinne linguistische Kriterien hinaus ein besonderes Augenmerk auf die grafische Gestalt eines Textes. Neben dem grafischen Aufbau des Textes kann es hierbei etwa um die Rolle von Bildern und ihrer Beziehung zum Text, um Revisionen oder schriftliche Anmerkungen anderer Autoren gehen, die auch Hinweise auf den zu Grunde liegenden Schreibprozess geben können.

Die vorliegende Arbeit stellt den Entwurf einer Beschreibungssprache zur Umschrift handschriftlicher Texte vor, in der solche Phänomene berücksichtigt und einer weiteren computergestützten Verarbeitung zugänglich gemacht werden können.

¹ Für anregende Diskussionen verschiedener Vorversionen dieses Textes danke ich Thomas Schmidt.