

Hagen Hirschmann: *Zur Korpusarchitektur der Falko-Korpora und der jüngeren Entwicklung hinsichtlich der Korpusauswertung*

Die Korpora der Falko-Familie bilden eine Reihe von Einzelkorpora mit beginnenden bis fortgeschrittenen Deutschlernenden aus unterschiedlichen Muttersprachenkontexten sowie deutschen MuttersprachlerInnen, die als Vergleichsdaten zu den Lernerdaten herangezogen werden können. Sämtliche durch die Korpora gespeicherte Sprachdaten bestehen aus Aufsätzen zu verschiedenen Themen. Die Korpora sind im Wesentlichen gleich aufgebaut, enthalten aber unterschiedliche bzw. unterschiedlich tiefe Annotationen (linguistische Analysen, die mit den ursprünglichen Lernertexten gespeichert wurden). Alle Korpora wurden in ein einheitliches Zielformat überführt, so dass sie gleichzeitig bzw. nebeneinander durchsucht und ausgewertet werden können. Im Vortrag werden Gemeinsamkeiten und Unterschiede zwischen den einzelnen Korpora der Falko-Familie herausgearbeitet; dabei wird auf den Erstellungsprozess und auf jüngere Studien, die diese Korpusdaten als Grundlage verwenden, eingegangen.