

Mehrsprachige Korpora am SFB 538: Daten, Tools und Standards

Thomas Schmidt
SFB 538 Mehrsprachigkeit
Universität Hamburg

01.12.2010

12:00 s.t.

Bibliothek Institut für Phonetik
Schellingstr. 3 VG Raum 226

Abstract

Der Vortrag beschäftigt sich mit den Korpora gesprochener Sprache, die am SFB 538 erhoben wurden und werden, um verschiedene Aspekte der Mehrsprachigkeit beim mehrsprachigen Individuum und in der mehrsprachigen Gesellschaft zu untersuchen. Dazu gehören beispielsweise Longitudinalstudien zum Spracherwerb bei zweisprachig aufwachsenden Kindern, Korpora zum professionellen und nicht-professionellen ("community") Dolmetschen oder Korpora zur rezeptiven Mehrsprachigkeit im skandinavischen Raum.

Neben einer Vorstellung der Daten selbst wird der Vortrag auch auf die Methoden, die zur ihrer Erhebung, Aufbereitung und Publikation verwendet wurden, eingehen. Dazu gehören insbesondere die Werkzeuge des EXMARaLDA-Systems (www.exmaralda.org) zum Erstellen, Verwalten und Auswerten von Korpora gesprochener Sprache. Abschließend werden einige Konzepte und Ideen zur Standardisierung von Datemodellen und -formaten für Daten gesprochener Sprache zur Diskussion gestellt.