

# Neue Dienste und Tools des Bayerischen Archivs für Sprachsignale

Christoph Draxler

draxler@phonetik.uni-muenchen.de

Florian Schiel, Thomas Kisler, Julian Pömp

24.04.2018

16-18 Uhr, DOR 24, 3.308

Das Bayerische Archiv für Sprachsignale (BAS) ist eine Einrichtung des Instituts für Phonetik und Sprachverarbeitung der LMU München. Das BAS ist ein Zentrum in der deutschen CLARIN-D Forschungsinfrastruktur. Es stellt Forscherinnen und Forschern Dienste, Werkzeuge und andere Sprachressourcen zur Verfügung.

Im Vortrag gebe ich einen Überblick über die aktuellen Arbeiten an den Werkzeugen des BAS:

**Octra** ist ein neuartiger Editor für die effiziente orthographische Transkription

**ASR** in Zusammenarbeit mit akademischen und industriellen Anbietern ist nun auch automatische Spracherkennung verfügbar

**Chunker** ermöglicht die automatische Segmentation mit WebMAUS auch für sehr lange Audiodateien

**Pipeline** bündelt mehrere Verarbeitungsschritte zu vorkonfigurierten Abläufen

Ursprünglich wurden diese Tools entwickelt, um die Bedürfnisse der Phonetik, Dialektologie und Sprachtechnologie abzudecken. Im Laufe der Zeit sind jedoch neue Forschungsgebiete wie Ethnologie, Soziologie oder auch Geschichtswissenschaften als Anwender dazugekommen. Diese Gebiete interessieren sich zunehmend auch für die Verarbeitung gesprochener Sprache, und sie haben ganz eigene Anforderungen, die die Tools erfüllen sollen. Im Gegenzug haben sie sehr große Datenmengen, die auch für die Phonetik und Sprachverarbeitung relevant und interessant sind, z. B. Zeitzeugeninterviews in der Oral History.

[clarin.phonetik.uni-muenchen.de/BASWebServices/](http://clarin.phonetik.uni-muenchen.de/BASWebServices/)