

*Einladung zum Vortrag im Rahmen des statistischen Praktikums
im Wintersemester 2009/10*

Bestimmung der regionalen Herkunft anhand gesprochener Ziffern

Abstract: Unter der Leitung von PD Dr. Christoph Draxler führte das Institut für Phonetik und Sprachverarbeitung der Universität München im gesamten Bundesgebiet Sprachaufnahmen mit Schülern ab der 8. Klasse durch. Diese Ph@tSessionz genannte Sammlung von Aufnahmen jugendlicher Sprecher soll in Zukunft etwa dabei helfen, sprachverarbeitende Technologien besser an die Stimmen sowie das Vokabular Jugendlicher anzupassen.

Doch können die Daten auch genutzt werden, um die Sprecher ihren Herkunftsregionen zuzuordnen? Zur Beantwortung dieser Frage steht ein Datensatz mit gesprochenen Einzelziffern zur Verfügung, der die jeweiligen Formantwerte und Segmentdauern enthält. Die deskriptive Analyse erfolgt mithilfe von Streudiagrammen und so genannter eplots. Anschließend wird mittels Random Forest eine Klassifikation versucht, wobei nicht nur das Finden geeigneter Unterscheidungsmerkmale, sondern vor allem auch die Adäquation der regionalen Herkunft eine Herausforderung darstellt.

Ort	Institut für Statistik, Ludwigstraße 33, Seminarraum 1. Stock
Zeit	Montag, 8.2.2010 um 13.00 Uhr
Projektpartner	PD Dr. Christoph Draxler, Institut für Phonetik und Sprachverarbeitung der LMU
Betreuer	Prof. Dr. Friedrich Leisch
Referentinnen	Julia Pläß, Felicitas Mittereder, Sabine Felbinger