

Akustische Analysen bairischer Dialektmerkmale: Zur Validität automatisch segmentierter Daten

Thomas Kisler, Felicitas Kleber

Institut für Phonetik und Sprachverarbeitung, Ludwig-Maximilians-Universität München
[[kisler|kleber](mailto:kisler@phonetik.uni-muenchen.de)]@phonetik.uni-muenchen.de

Das IDS-Korpus „Deutsch heute“ (Brinkmann et al., 2008) enthält orthographische Verschriftungen und Sprachsignalen von je vier Sprechern aus 160 Orten in Deutschland, Österreich und der Schweiz und stellt somit eine wertvolle Datensammlung zum gesprochenen Gegenwartsdeutsch dar. Aufgrund des hohen Zeitaufwands, der mit einer manuellen Aufbereitung und Auswertung von Korpora für phonetische Analysen verbunden ist, rückt die Frage nach der Einbindung automatisierter Abläufe und transkriptionsunabhängiger Analysen akustischer Merkmale in den Vordergrund. Ziel der vorliegenden Studie ist es, anhand der „Deutsch heute“-MapTask-Daten aus den bayrischen und österreichischen Aufnahmeorten diverse bairische Dialektmerkmale auf der Grundlage eines automatisch segmentierten Korpus akustisch zu untersuchen und die Vorteile und Grenzen dieser Methode aufzuzeigen.

Hierfür wurden zunächst die MapTask-Daten aller Erhebungsorte in Bayern und Österreich auf der Grundlage von orthographischen Verschriftungen mit WebMAUS (Kisler et al., 2012) automatisch segmentiert. Anschließend wurden aus den am häufigsten vorkommenden Wörtern diejenigen ausgewählt, die sich für die Analyse der folgenden Dialektmerkmale eignen: (1) komplementäre Länge von Vokal und darauffolgendem Plosiv (Moosmüller & Brandstätter, 2014), (2) Monophthongierung und „/E/-Verwirrung“ in österreichischen Varietäten (Moosmüller & Scheutz, 2013) und (3) Vokalqualität und -quantität in ungespannten, hohen Vorderzungenvokalen (Cunha et al., 2015). Phrasen mit diesen Wörtern wurden extrahiert und in eine emuR-Datenbank (Winkelmann, 2015) konvertiert. Alle akustischen Analysen erfolgten ohne weitere manuelle Bearbeitung der Daten direkt in R (<https://www.r-project.org/>).

Die Ergebnisse aus den akustischen Analysen sind vergleichbar mit Beschreibungen dialektaler Merkmale in der Literatur. Auch wenn man für die Untersuchung feiner phonetischer Details nicht auf eine manuelle Überprüfung der Segmentgrenzen verzichten sollte, so beschleunigt die automatische Segmentierung die Datenauswertung erheblich und schafft eine solide Basis für die Untersuchung akustischer Merkmale, die wiederum eine transkriptionsunabhängige Analyse dialektaler Merkmale ermöglicht.

Literatur

- Brinckmann, C., Kleiner, S., Knöbl, R., & Berend, N. (2008). German Today: an areally extensive corpus of spoken Standard German. *Proc. 6th LREC*, Marokko.
- Cunha, C., Harrington, J., Moosmüller, S. & Brandstätter, J., P. (2015). The influence of consonantal context on the tense-lax contrast in two standard varieties of German. In: A. Leemann et al. (eds.), *Trends in Phonetics and Phonology*. Lang, 65-77.
- Kisler, T., Schiel, F., & Sloetjes, H. (2012). Signal processing via web services: the use case WebMAUS. *Proc. Digital Humanities 2012*, Hamburg, Germany, pp. 30-34.

- Moosmüller, S. & Brandstätter, J. (2014). Phonotactic information in the temporal organization of Standard Austrian German and the Viennese dialect. *Language Sciences*, 46, 84–95.
- Moosmüller, S. & Scheutz, H. (2013). *Chain shifts revisited: The case of Monophthongisation and E-confusion in the city dialects of Salzburg and Vienna*. In P. Auer et al. (eds.), *Language variation European Perspectives IV*. Benjamins, 173-186.
- Winkelmann, R. (2015). Managing Speech Databases with emuR and the EMU-webApp. *Proc. 16th Conf. of the International Speech Communication Association*.