

Automatische phonetische Segmentierung von Dialektdaten: eine Evaluation

Markus Jochim¹, Felicitas Kleber¹

¹Institut für Phonetik und Sprachverarbeitung,
Ludwig-Maximilians-Universität München
{markusjochim|kleber}@phonetik.uni-muenchen.de

In vielen phonetischen Forschungsprojekten stellt sich die methodologische Frage, wie die erhobenen Sprachdaten akustisch segmentiert werden sollen. Wie in vielen anderen Fachbereichen auch geht diese Frage verstärkt mit der Forderung nach einer Automatisierung des Arbeitsprozesses einher. In dieser Studie vergleichen wir automatisch vs. manuell segmentierte Dialektdaten (Westmittelbairisch); wir zeigen, dass Dialektdaten selbst dann zufriedenstellend mit automatischen Werkzeugen segmentiert werden können, wenn das Sprachmodell der überdachenden Standardsprache Verwendung finden muss – in Ermangelung eines spezifisch für den Dialekt trainierten Sprachmodells. Wir zeigen außerdem, dass die Qualität automatischer Segmentierung auch für dauerbasierte Forschungsfragen ausreichen kann.

Als automatisches Segmentierungswerkzeug dient MAUS [1]. Unser Testkorpus besteht aus 450 gelesenen Äußerungen der Form *Er muss noch Butter kaufen* (45 Sätze * 2 Sprechgeschwindigkeiten * 5 Wiederholungen). Die Sätze wurden 16 Sprecher*innen des Westmittelbairischen in nicht normierter, lautorientierter Dialektschreibung präsentiert (z. B. *Ea muas no Butter kaffa*) und von ihnen gelesen. Damit ergeben sich 7.200 Äußerungen, für die wir der automatischen Segmentierung eine manuell durchgeführte Segmentierung gegenüberstellen können. Es zeigt sich sowohl für die Dauer der akzentuierten Wörter als auch für die Dauer der betonten Vokal-Konsonant-Sequenzen (/ʊt/ in *Butter*) eine äußerst geringe Diskrepanz zwischen automatischer und manueller Methode.

Auf diesen Erkenntnissen aufbauend sollen in einem späteren Schritt generalisierbare Zahlen abgeleitet werden, wie groß der zu erwartende Fehler einer automatischen Segmentierung ist, und zwar in Abhängigkeit von den jeweiligen Lautklassen.

Solche Zahlen sollen eine valide und reliable a priori Einschätzung erlauben, ob eine rein automatisch erstellte Segmentierung für eine gegebene Forschungsfrage ausreicht. Erste Ergebnisse deuten darauf hin, dass sogar dauerbasierte Forschungsfragen mittels einer rein automatischen Segmentierung bearbeitet werden können.

Literatur

- [1] Thomas Kisler, Uwe Reichel, and Florian Schiel. Multilingual processing of speech via web services. *Computer Speech & Language*, 45:326 – 347, 2017.