WAS SOLL(T)EN WIR MIT SPRACHDATENBANKEN ANFANGEN?

Bernd Pompino-Marschall

Institut für deutsche Sprache und Linguistik der Humboldt-Universität zu Berlin
bernd.pompino-marschall@rz.hu-berlin.de

Abstract: Der Beitrag greift die auf der Eurospeech '93 erstmals vorgestellte Idee einer datenbasierten kompletten phonetischen Theorie ('databased complete phonetic theory', CPT) [5] auf. Während in den letzten Jahrzehnten eine Vielzahl von Sprachdatenbanken vor allem für die Zwecke der sprachverarbeitenden Industrie (z.B. als Trainingsmaterial für die automatische Spracherkennung) – unter großem Personaleinsatz auch seitens der Phonetikinstitute (s. PHONDAT / VERBMOBIL) ganz Deutschlands – erstellt wurden, blieb deren Weiternutzung im Rahmen der phonetischen Forschung eher auf einzelne Untersuchungen, meist zu Spezialfragen, beschränkt [1-4]. Es sollte heute jedoch möglich sein, dieses annotierte Material gezielt dazu zu nutzen, ein phonetisches Modell zu entwickeln, das die Veränderung der kanonischen Aussprache eines Wortes einer Einzelsprache (hier: Deutsch) in Isolation (wie sie in den Aussprachewörterbüchern angegeben ist) unter unterschiedlichen Bedingungen in fließender Rede vorhersagt. Das Ziel eines solchen Unterfangens wäre die Antwort auf die Frage, wie sich ein Wort aufgrund der vom Modell vorhergesagten prosodischen und segmentalen Modifikation der im mentalen Lexikon gespeicherten kanonischen Form in Abhängigkeit von syntaktischen, semantischen, pragmatischen und anderen Faktoren wie z.B. der individuellen Sprechercharakteristik und der Stilebene in fließender Rede phonetisch darstellt. Es ist zu erwarten, dass die in diesem Szenario zu gewinnenden Resultate einen entscheidenden Beitrag auch zum besseren Verständnis der bei der Sprachproduktion real ablaufenden Prozesse liefern.

Literatur

[1] Kohler, K.J.: Glottal stops and glottalization in German. Data and theory of connected speech processes. In: Phonetica 51 (1994), pp. 38-51.

[2] Kohler, K.: Plosive-related glottalization phenomena in read and spontaneous speech. A stød in German. In: Grønnum, N./Rischel, J. (Eds.): To Honour Eli Fischer-Jørgensen. Kopenhagen 2001 (= Travaux du Cercle Linguistique de Copenhague 31), pp. 174-211.

[3] Pfitzinger, H.R.: Local Speech Rate Perception in German Speech. In: Proc. 14th ICPhS 1999. San Francisco 1999, vol. 2, pp. 893-896.

[4] Simpson, A.P. : Phonetische Datenbanken des Deutschen in der empirischen Sprachforschung und der phonologischen Theoriebildung. Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK) 33 (1998).

[5] Tillmann, H.G. / Pompino-Marschall, B.: Theoretical principles concerning segmentation, labelling strategies and levels of categorical annotation for spoken language database systems. In: Eurospeech '93. Berlin 1933, pp. 1691-1694.