Mapping regional variation of speech rate using automatic measurements of amplitude envelope peaks

We investigate the region-specific interweaving of speech rate and phonetic reduction on the basis of standard-intended reading pronunciation in the German-speaking countries of central Europe (Hahn & Siebenhaar, 2019). The data base for this research are 1652 recordings of men and women in two age groups (17-20 and 50-60 years respectively) available (Kleiner, 2015), all of which have already been automatically aligned with WebMAUS (Kisler et al., 2017). For more reliable evaluations of sound durations, the segmentations at sound level were also corrected manually for approximately 600 of these recordings (the group of young men).
On the basis of these segmentations and annotations specific reduction phenomena are investigated, correlated to speech rate and displayed in geolinguistic space. For the speech rate various measurements are used such as articulation and speaking rates in phones or syllables/second.
Due to the amount of workload involved in manual correction, not all data can be included in the study, which is why it is not yet possible to compare speech rates between the sexes and age groups in terms of geolinguistic space.
We therefore examine in a first step whether automatic procedures such as the calculation of the peaks of the amplitude envelopes (He & Dellwo, 2016, 2017) can provide sufficiently reliable data to be able to include the missing recordings in the analysis and to rise the local density of the data points. Envelopes were calculated by low-pass filtering (cut-off: 10Hz) a full-wave rectified speech signal. A peak point was defined as the sample that is preceded and followed by lower amplitude samples and that has a minimum amplitude of one tenth of the maximum amplitude sample in the signal. Inter-peak intervals were defined as the intervals between two peak points starting and ending with the first and last peak in an utterance respectively (c.f. fig 1). First tests are promising and achieve high correlations between the manually segmented data and the automatically measured data.
In a second step, the measured speech rate values are transferred to geolinguistic maps using arcGIS (ESRI, 2015) and REDE SprachGIS (Schmidt et al., 2008ff.) to test and compare regional patterns of the groups in question.
The aim of these comparisons is to show that the regional patterns already worked out for young men do not appear to be exclusive to this group of speakers, but also show similar patterns for women and older speakers. In this way it can be substantiated that regionality is an important factor for the variation of speaking speed. Moreover, it can be justifiably assumed that the patterns of reduction phenomena and the temporal microstructures observed for the group of the young speakers can be transferred to the other speaker groups.

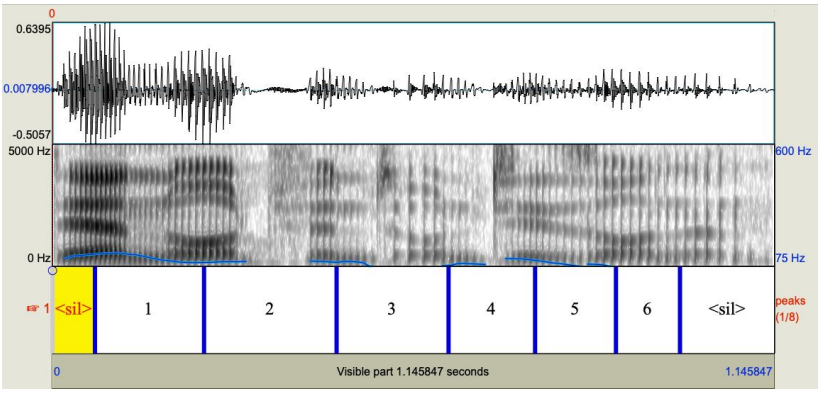

Figure 1: Waveform and spectrogram of the utterance “Der Nordwind und die Sonne” produced by a male speaker with peak-to-peak text TextGrid containing 7 peaks (boundaries), hence 6 peak-to-peak intervals.

References

ESRI. (2015). ArcGIS Desktop (Version 10.3). ESRI (Environmental Systems Research). Redlands, CA.

Hahn, M., & Siebenhaar, B. (2019). Spatial Variation of Articulation Rate and Phonetic Reduction in Standard-Intended German. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019 (pp. 2695–2699). Australasian Speech Science and Technology Association Inc.

He, L., & Dellwo, V. (2016). A Praat-Based Algorithm to Extract the Amplitude Envelope and Temporal Fine Structure Using the Hilbert Transform. In Interspeech 2016 (pp. 530–534). ISCA.

He, L., & Dellwo, V. (2017). Between-speaker variability in temporal organizations of intensity contours. The Journal of the Acoustical Society of America, 141(5), 488–494.

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. Computer Speech & Language, 45, 326–347.

Kleiner, S. (2015). 20. „Deutsch heute“ und der Atlas zur Aussprache des deutschen Gebrauchsstandards. In R. Kehrein, A. Lameli, & S. Rabanus (Eds.), Regionale Variation des Deutschen: Projekte und Perspektiven (pp. 489–518). De Gruyter.

Schmidt, J. E., Herrgen, J. & Kehrein, R. (2008ff.). Regionalsprache.de (REDE). Forschungsplattform zu den modernen Regionalsprachen des Deutschen. Teil 6: REDE SprachGIS – Das forschungszentrierte sprachgeographische Informationssystem von Regionalsprache.de. Forschungszentrum Deutscher Sprachatlas.