Assessing the potential of crowdsourced 'Dialäkt Äpp' speech data for forensic phonetics

Adrian Leemann and Marie-José Kolly
Department of Comparative Linguistics, University of Zurich
`{adrian.leemann|marie-jose.kolly}@pholab.uzh.ch`

The free of charge iPhone application *Dialäkt Äpp* (Leemann & Kolly, 2013; Kolly & Leemann, in press) features the following two core functionalities: (1) users click on the pronunciation variants of 16 words and the application predicts their local dialect, (2) users record their pronunciation of the same 16 words, which are then uploaded on a server and displayed on an interactive map. The goal of the application is science communication to a broad public. The app has been downloaded by >59'000 users.

As speech scientists we are now in the position to analyze the data gathered through *Dialäkt Äpp*. With the users' consent, we retrieve acoustic pronunciation data of 16 words for thousands of dialect speakers originating from all over German-speaking Switzerland (cf. function (2) above). Until recently, traditional methods for empirical linguistic research based their analyses mostly on small sets of speakers. The use of smartphone app technology for crowdsourcing linguistic data is relatively new: smartphone applications have hitherto been used to collect speech to train acoustic models (de Vries, Davel, Badenhorst, Basson, de Wet et al., 2014) or to document endangered languages (Iwaidja Inyman Team, 2012).

The crowdscourced speech data from *Dialäkt Äpp* allows for the collection and analysis of a number of speech signal parameters in order to create large-scale population statistics. In the field of forensic phonetics, such population statistics only exist for certain languages and typically feature <150 speakers (Künzel, Masthoff & Köster, 1995; Jessen, 2007). A preliminary analysis of *Dialäkt Äpp* recordings of 115 users from Bern (city) and 205 users from Zurich (city) revealed that Bern SwG speakers speak significantly slower than Zurich SwG speakers. For 6 disyllabic words per speaker we measured the temporal duration between the two vowel onsets. We call this vowel-onset-to-vowel-onset measure *durVonVon* (see Figure 1). In theory, this measure is motivated by Allen's (1972) findings that vowel onsets represent perceptually prominent centers of a syllable.
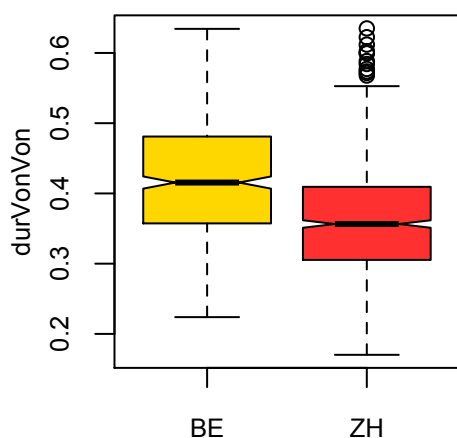

Figure 1 shows the boxplots of the two dialects' *durVonVon* values. The longer the temporal duration between the two vowel onsets, the slower the articulation rate. The values between the two dialects are significantly different: the durational information contained in a few words alone discriminates the two dialects (cf. Leemann, Kolly & Dellwo, 2014). In the present contribution, we will use this example of articulation rate differences to illustrate the potential of *Dialäkt Äpp* speech data for forensic phonetic purposes.

Figure 1: *Boxplots of the dialects' durVonVon values.*

References

Allen, G. D. (1972). The location of rhythmic stress beats in English: An experimental study I. *Language and Speech*, **15**, 72–100.

Iwaidja Inyman Team. (2012). Ma! Iwaidja. https://itunes.apple.com/au/app/ma-iwaidja/id557824618?mt=8 (accessed 06.05.2014).

Kolly, M.-J. and A. Leemann. (in press). 'Dialäkt Äpp': Communicating dialectology to the public – crowdsourcing dialects from the public. To appear in A. Leemann, M.-J. Kolly, S. Schmid and V. Dellwo (Eds.), *Trends in Phonetics in German-speaking Europe*. Bern/Frankfurt: Peter Lang.

Leemann, A. and M.-J. Kolly. (2013). Dialäkt Äpp. https://itunes.apple.com/ch/app/dialakt-app/id606559705?mt=8 (accessed 06.05.2014).

Leemann, A., M.-J Kolly, and V. Dellwo. (2014). Crowdsourcing regional variation in speaking rate through an iOS app. To appear in: *Proceedings of Speech Prosody 2014*.

Künzel, H., H. Masthoff and Köster, J. (1995). The relation between speech tempo, loudness, and fundamental frequency: an important issue in forensic speaker recognition. *Science and Justice*, **35**, 291–295.

de Vries, N., M. H. Davel, J. Badenhorst, W. D. Basson, F. de Wet, E. Barnard and A. de Waal. (2014). A smartphone-based ASR data collection tool for under-resourced languages. *Speech Communication*, **56**, 119–131.