

Speaker-individual rhythmic features in both L1 and L2 speech: implications for forensic voice comparison

Marie-José Kolly, Adrian Leemann, and Volker Dellwo

*Phonetics Laboratory, Department of Comparative Linguistics, University of Zurich,
Switzerland*

{marie-jose.kolly|adrian.leemann}@pholab.uzh.ch,
volker.dellwo@uzh.ch

It is known that speakers often transfer speech rhythmical patterns from their L1 to their L2, which may affect their intelligibility (Adams, 1979; Wenk, 1985). In the present contribution we address how these L1-interference phenomena could be leveraged for forensic phonetic purposes: Do certain (speaker-individual) rhythmic characteristics remain unchanged when a speaker talks in different languages?

A number of speech rhythmic features, e.g. the percentage of voiced portions in the speech signal (Dellwo, Fourcin & Abberton, 2007), were shown to have potential for forensic voice comparison as they strongly vary between speakers but remain largely unaffected by within-speaker variability in speaking style (spontaneous vs. read) and transmission channel (hifi vs. telephone) (Leemann, Kolly & Dellwo, 2014), and by within-speaker variability when speakers disguise their voice by obstructing their articulators (Leemann, Hove, Kolly & Dellwo, submitted). The overall objective of the present contribution is to examine speech for speaker-individual rhythmic features that are independent of the language being spoken.

Our research is based on the TEVOID corpus (Dellwo, Leemann & Kolly, 2012; Leemann, Kolly & Dellwo, 2014) that contains Zurich German (L1) speech of 16 speakers and French and English (L2) speech of the same 16 speakers. Results based on 16 sentences per speaker and language showed that selected, automatically extracted rhythmic measures, e.g. the percentage of voiced portions in the speech signal, varied between speakers but remained largely unaffected by within-speaker variability in the language spoken (Kolly, Dellwo & Leemann, 2013). We have now collected more material per speaker and are currently segment-labeling this material, which will allow us to calculate a wider variety of rhythmic measures.

The present contribution reports on between- and within-speaker variability of a number of rhythmic measures, using 32 Zurich German (L1), 32 French (L2) and 32 English (L2) sentences per speaker. Based on preliminary results (cf. Kolly, Dellwo & Leemann, 2013) we expect high between- and low within-speaker variability in selected measures of speech rhythm.

In forensic voice comparison, cases occur where there is a mismatch in language between acoustic trace and comparison material. In a considerable number of forensic phonetic casework, practitioners have to make decisions about speaker identity based on speech samples where the trace material is in one language – e.g. the speaker’s L1–, and the suspect material is in another language – e.g. the speaker’s L2 (Herbert R. Masthoff, personal communication). This may happen, for example, when a suspect uses an L2 in order to disguise his/her voice. However, the impact of L2 speech on speaker-individual characteristics is largely unknown – this is why forensic phoneticians “should exercise particular caution if the samples for comparison are in different languages” (IAFPA Code of Practice). The present contribution is thus expected to have implications for forensic voice comparison.

References

- Adams, C. (1979). *English speech rhythm and the foreign learner*. The Hague: Mouton.
- Bohn, O.-S. (1998). Wahrnehmung fremdsprachlicher Laute. Wo ist das Problem? In H. Wegener (Ed.), *Eine zweite Sprache lernen. Empirische Untersuchungen zum Zweitspracherwerb*, 1–20. Tübingen: Narr.
- IAFPA Code of Practice. <http://www.iafpa.net/code.htm> (accessed 28.04.2014).
- Dellwo, V., A. Fourcin and E. Abberton. (2007). Rhythmical classification based on voice parameters. *Proceedings of the ICPHS 2007*, Saarbrücken: 1129–1132.
- Dellwo, V., A. Leemann and M.-J. Kolly (2012). Speaker idiosyncratic rhythmic features in the speech signal. *Proceedings of Interspeech 2012*, Portland, OR (USA).
- Leemann, A., M.-J. Kolly and V. Dellwo. (2014). Speaker-individuality in suprasegmental temporal features: Implications for forensic voice comparison. *Forensic Science International*, **238**, 59–67.
- Leemann, A., I. Hove, M.-J. Kolly and V. Dellwo (submitted). Testing the effect of voice disguise on suprasegmental temporal features: Corpus creation and preliminary results. Submitted to the Interspeech 2014 conference.
- Lloyd James, A. (1929). *Historical introduction to French phonetics*. London: ULP.
- Kolly, M.-J., V. Dellwo and A. Leemann. (2013). Speaker-idiosyncratic temporal patterns in L2 speech. Abstract presented at the annual IAFPA 2013 meeting, Tampa, FL (USA).
- Taylor, D. S. (1981). Nonnative speakers and the rhythm of English. *International Review of Applied Linguistics in Language Teaching*, **19**, 221–226.
- Wenk, B. (1985). Speech rhythms in second language acquisition. *Language and Speech*, **28**, 157–175.