Phonetic characteristics of hesitation vowels in Swiss German and their use for forensic phonetic speaker identification.

Dario Brander
Department of Comparative Linguistics, University of Zurich
dario_brander@gmx.ch

Hesitation vowels (commonly transcribed as äh / ähm in German-speaking regions) are one-syllable verbal utterances used to fill a speaking pause between two words or other linguistic elements. While the actual function of these sounds is still disputed (Clark & Fox Tree 2002, O’Connell & Kowal 2005), interest in using this type of verbal utterance for forensic phonetics has grown recently. An early study showed that interpersonal and interdialectal variations can be observed in f0 and F1/F2 of hesitation vowels (Alaoui in Jessen 2005, 273f.). Meanwhile, another study conducted under the instruction of the German Federal Police Office (Trouvain & Bauer 2005) confirmed these results and showed variability in the factors of usage (äh vs ähm and positioning in verbal utterances), temporal features (articulation rate hes/min) and intra-speaker variation. One further study (Klug & König 2012) additionally considered speakers’ spread of data to be a speaker-specific factor as well.

The objectives of this contribution are to confirm the aforementioned results in a Swiss German speaker setting and to analyze further factors of potential inter-speaker-variation in hesitation vowels. The study features first results in the analysis of f0-ratio (the comparison of the initial and final 25% and central 50% of the vowel), F1-3 of the bilabial nasal in ähm and the temporal features general duration, vowel / nasal duration and vowel-nasal-ratio in ähm. Additionally, this study presents and discusses the possibility to optimize inter-speaker-variation by grouping the data material of speakers according to matters of usage (äh vs ähm and positioning in a verbal utterance) and separate analysis under these conditions. In a small-scale comparative analysis, the data stability of one speaker’s hesitation vowels is compared to the stability of [a] of the same speaker in a read condition to determine if the analysis of hesitations alone yields clearer results than a general f0 and F1-3 analysis of a commonly used sound in Swiss German.

Methodologically, I proceeded as follows: 20 speakers of Zurich German (students, age 20-35) were recorded at the University of Zurich as part of Dellwo et al. 2012. The data was recorded in a sound treated booth. The participants partook in a 20-35 minutes long interview in which they were instructed to answer the questions freely. From this corpus, the 4 speakers of both gender groups with the highest number of hesitations were selected. Their hesitation vowels were extracted and analyzed. For the comparative analysis, a recording of a later phase of the recording sessions was chosen. One of the chosen speakers read 256 transcribed sentences of the former spontaneous recording session in Swiss German. His [a]-vowels were extracted and their f0 and F1-3 analyzed. The standard deviations of those factors were then compared to the standard deviations of f0 and F1-3 of the hesitation vowels in the spontaneous condition.

References


Abstract: [http://p3.snf.ch/project-135287](http://p3.snf.ch/project-135287) [13.05.2013]


