Identifying a speaker’s dialect: the role of segments and prosody

Adrian Leemann, Francis Nolan
Department of Theoretical and Applied Linguistics, University of Cambridge
{al764|fjn1}@cam.ac.uk

Rationale
Dialectal information is typically exploited in the analysis of criminal and suspect recordings (cf. Jessen, 2007; Köster et al., 2012). However, not only experts evaluate dialects but also earwitnesses: when asked to describe the voice they heard, dialect forms a central part in that description (cf. Hollien, 2002). In crime scenes, the voices that are heard by earwitnesses are often mixed with background noise, which can make speech highly unintelligible (Olaf Köster (BKA), personal communication). As an example, Figure 1 demonstrates the effect of multi-talker babble on the spectrum of a vowel.

![Figure 1](image)

**Figure 1** (A) spectrum of the vowel /ɛ/, (C) spectrum of multi-talker babble, (D) combination of (A) + (C) at SNR 0 dB (adapted from Assmann & Summerfield 2004, 244).

Compared with panel (A), panel (D) shows a drop in the degree of spectral contrasts introduced by the competing voices. How does noise affect earwitness and expert performance in evaluating a suspect’s dialect? Is the contribution of segmental and suprasegmental cues in the identification of a speaker’s dialect the same in adverse and ideal listening conditions? Recent research has revealed that background noise can trigger a re-ordering of acoustic cues to linguistic categories: ‘secondary’ cues in laboratory settings can become the only accessible, and thus ‘primary’, cues in noisy environments (cf. Summerfield & Haggard, 1977; Parikh & Loizou, 2005; Jiang et al., 2006; Van Engen & Bradlow, 2007:519). In this contribution we present the methodological framework and first results of an ongoing project that examines the role of segmental and suprasegmental cues in identifying a speaker’s dialect in ideal and adverse listening conditions.

Method
In two experiments we separate segmental and suprasegmental features and play them off against each other. This can be achieved by manipulating the speech signal so that suprasegmental features of one dialect (Bern Swiss German) will be grafted onto the segments of another dialect (Zurich Swiss German) and vice versa (cf. Vaissière & Boula de Mareüil, 2004; Boula de Mareüil & Vieru-Dimulescu, 2006). These manipulated stimuli are
presented to naïve Bern and Zurich Swiss German listeners (both for ideal and adverse listening conditions) who will be asked to judge whether the stimulus they heard stems from a speaker of Bern or Zurich Swiss German.

Implications
Knowing more about which cues are important for dialect identification in laboratory and realistic environments will make claims of both expert and naïve listeners forensically more conclusive. The insights gained from the current study will further be relevant to research on LADO. An issue currently being debated is whether native-speakers of the language in question should be involved in the assessment of the speech of an asylum seeker, given that the expertise of a linguist is not the same as that of a native-speaker (cf. Nolan, 2012). Results of the current study will shed light on naïve listeners’ dialect identification performance.

References


