

One Speaker, Multiple Dialects: Insights from the Language Familiarity Effect for Forensic Speaker Recognition

Sophie Möller¹, Radek Skarnitzl¹, Volker Dellwo², Elisa Pellegrino², Tyler Perrachione³

¹*Institute of Phonetics, Charles University, Czech Republic*
{sophie.moeller|radek.skarnitzl}@ff.cuni.cz

²*Department of Computational Linguistics, University of Zurich, Switzerland;*
{elisa.pellegrino|volker.dellwo}@uzh.ch

³*Department of Speech, Language and Hearing Sciences, Boston University, United States*
tkp@bu.edu

Introduction

In forensic speaker comparison, practitioners may encounter language mismatch conditions involving standard and dialect speech across recordings. During auditory assessments by native-language experts, similarities in vocal characteristics may suggest that recordings originate from the same speaker, whereas language- or dialect-related differences may instead indicate different speakers.

Background

The present study investigates how linguistic and voice quality characteristics influence native (naïve) listeners' judgments of speaker identity under dialect mismatch conditions. This question is particularly relevant in sociolinguistic contexts such as German-speaking Switzerland, where dialects play a central role in social identity and perception, and where unexpected shifts in salient dialectal features may interfere with voice recognition by native listeners.

In such contexts, it is therefore plausible that familiarity with a language or one of its dialects may bias native listeners toward perceiving different speakers under mismatch conditions, whereas not knowing about the dialect differences (as a non-native speaker) may reduce such biasing effects by encouraging greater reliance on more invariant acoustic cues to talker identity. This possibility is particularly interesting in relation to the Language Familiarity Effect (LFE) – the well-documented advantage listeners show when recognizing voices speaking their native language compared to a non-native one (McLaughlin et al. 2019; Perrachione 2017). Importantly, this does not necessarily imply an overall disadvantage for native listeners but rather predicts a stronger mismatch-related decrement between matched and mismatched conditions. While previous studies have investigated the impact of dialectal variation on talker identification, less is known about whether native and non-native listeners are affected differently by dialect mismatch conditions within a single language. We refer to this predicted pattern as a “native familiarity disadvantage”.

Methodology

Using an online training-identification paradigm, the study investigates how mismatch between Swiss Standard German and regional Swiss German dialects influences perceptual speaker recognition by native (Swiss) and non-native (Czech) listeners. During training, listeners learn six target speakers, half presented exclusively in Swiss Standard German and half in their regional dialect, through repeated exposure to different full-sentence recordings. They repeatedly associate each voice with a speaker identity and complete immediate speaker-identification trials by selecting the learned target speaker from two foil speakers. During the test phase, listeners complete the same identification task using new sentence material, with target speakers presented either in matched (e.g., dialect–dialect) or mismatched speaking-mode (Standard–dialect) conditions relative to training. Stimuli are drawn from the DACH corpus (Schmid et al. 2019).

Hypothesis

In line with our expectations about a “native familiarity disadvantage”, we hypothesize that dialect mismatch between training and test recordings will reduce speaker recognition accuracy by native Swiss German listeners. In contrast, we expect non-native Czech listeners to show comparable performance across conditions, indicating reduced susceptibility to dialect mismatch.

The findings may provide insights into whether dialect mismatch leads native listeners to underestimate speaker similarity across speaking varieties, with potential implications for the interpretation of auditory speaker comparisons in forensic contexts. More broadly, the study may contribute to our understanding of how top-down linguistic expectations and bottom-up acoustic cues interact during speaker identity judgements.

References

- McLaughlin, D. E., Carter, Y. D., Cheng, C. C., & Perrachione, T. K. (2019). Hierarchical contributions of linguistic knowledge to talker identification: Phonological versus lexical familiarity. *Attention, Perception, & Psychophysics*, 81(4), 1088–1107. <https://doi.org/10.3758/s13414-019-01778-5>
- Perrachione, T. K. (2019). Recognizing speakers across languages. In S. Frühholz & P. Belin (Eds), *The Oxford Handbook of Voice Perception* (pp. 515–538). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198743187.013.23>
- Schmid, S., Kleber, F., & Moosmüller, S. (2019). *Typology of vowel and consonant quantity in Southern German varieties: Acoustic, perception, and articulatory analyses of adult and child speakers* (DACH corpus). Swiss National Science Foundation project 164377. Accessible at: <https://data.snf.ch/grants/grant/164377>