Perceptual speaker discrimination based on German consonants

Carola Schindler, Eva Reinisch, Jonathan Harrington
Institute of Phonetics and Speech Processing, Ludwig Maximilians University, Munich, Germany
{carola.schindler|evarei|jmh}@phonetik.uni-muenchen.de

Nasal and fricative consonants appear to contain high amounts of speaker-specific information. Their acoustic properties tend to differ between speakers to a larger extent than within one speaker's different productions. Mook and Draxler (2012) showed this for German by analysing spectral moments of different types of consonants and vowels (see also Schindler & Draxler, 2013). Speaker-specific information in nasals and fricatives is also evident in perception. In a speaker discrimination experiment participants were more accurate when the words they heard contained /m/ or /s/ than /l/ or /t/ (Andics, 2013). The goal of the present study was to explore in more detail and with a larger set of consonants whether listeners' ability to perceptually discriminate between speakers depends on the types of consonants they hear and whether this pattern would match the acoustic analyses.

Participants performed a speaker discrimination task. Stimuli were nonsense words that consisted of a consonant in a bilateral /a/-context. Consonants were nasals, fricatives, and stops, each in labial and alveolar place of articulation (i.e., /m/, /n/, /f/, /s/, /p/, /t/). Four different tokens from nine Bavarian speakers were recorded and paired to same-speaker and different-speaker pairs. Participants performed a same-different discrimination task. Each stimulus was flanked by 500 ms pink noise to make the task harder. Trials were blocked by consonant and presented in randomised order.

Since listeners were very good at discriminating speakers for all consonants (mean accuracy in a pilot study was 0.95) and in order to reach a better degree of separation between consonants, the stimuli were manipulated to make the task more difficult. The pitch contour was flattened and normalised, and the vowels shortened to 50 ms on each side. This caused the overall accuracy to drop to 0.83. In a second experiment the consonants were spliced into an identical vowel context for all speakers, so that the listeners could not use the vowel information to discriminate between the speakers (mean accuracy 0.62). Both experiments showed differences between the consonants, with larger effects for the more difficult task. Also the place and manner of articulation modulated listeners' speaker-discrimination abilities. Comparing accuracy rates for the different types of manipulated stimuli (with or without speaker information in the vowel) will also help to pinpoint what kinds of information contribute to speaker discriminability.

References

Andics, A. (2013). Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. PhD Thesis, Radboud University Nijmegen, Nijmegen.

Mook, C. and Draxler, C. (2012) The speaker discriminating power of nasals, fricatives and vowels, Poster presented at the IAFPA 2012, Santander, Spain.

Schindler, C. and Draxler, C. (2013) The influence of the place of articulation on the speaker specificity of German phonemes, Poster presented at the 4th International Summer School on "Speech Production and Perception: Speaker-Specific Behavior" 2013, Aix-en-Provence, France.