Comparison of similarity and dissimilarity indices between speech samples in filtered and non-filtered conditions for the speakers of the Croatian language

Gordana Varošanec-Škarić1, Iva Pavić2, and Gabrijela Kišiček2

1,2,3Department of Phonetics, University of Zagreb, Croatia
gvarosan@ffzg.hr, ipavic2@ffzg.hr, gkisicek@ffzg.hr

Acoustic-statistical measurements of similarity index (R) and dissimilarity index (SDDD) on the basis of long term average spectra (LTASS) can be used as a support measurement in forensic phonetic cases (Harmegnies, 1995). In this research similarity and dissimilarity indices were compared for speech samples in filtered and non-filtered conditions. The data consisted of 86 speakers originating from 8 largest Croatian cities representing three dialects of Croatian language. All speakers were recorded under the controlled conditions reading standardized text and during the spontaneous speech. Recordings were edited in Cool Edit program and speech samples (duration 60 s) were filtered. Filtered and non-filtered speech samples were then compared on the basis of LTASS (non-filtered conditions (0 - 10 kHz) and filtered (0.8 – 4 kHz)). Using index R and index SDDD intraspeaker variations and interspeaker variations were compared respectively for male and female speakers. Results of intraspeaker variations showed that average values of similarity index (R) in non-filtered conditions were between 0.94 for male speakers in reading texts to 0.98 for female speakers in reading text and spontaneous speech. Results of interspeaker variations showed lower values of index R in the non-filtered conditions: from 0.86 in spontaneous speech to 0.94 in reading text for female speakers. Average values of R in filtered conditions for intraspeaker variations were between 0.83 for both female and male speakers in spontaneous speech to 0.95 in reading texts. Average values of R index in filtered conditions for interspeaker variations were significantly lower; from 0.57 for male spontaneous speech to 0.9 for female reading texts. Average values of index SDDD in non-filtered conditions for intraspeaker variations were generally lower – from 2.27 for female speakers to 3 for male speakers in reading. SDDD index showed higher values in non-filtered conditions for interspeaker variations; from 4.75 in female reading speech and male spontaneous speech to 5.12 for male reading speech. In filtered conditions intraspeaker variations resulted with SDDD index between 2.14 for male reading speech to 3.01 for female spontaneous speech. As expected, results in filtered conditions for interspeaker variations showed higher values of SDDD index, from 3.06 for female to 4.71 for male reading speech. The differences between similarity index (R) in intraspeaker variations were statistically significant for female speakers (p<0.0001) and for male speakers (p<0.05) in both spontaneous speech and reading. Results of interspeaker variations showed statistically significant differences in similarity index (R) for male speakers (p<0.0001 in reading and p<0.0001 in spontaneous speech) and female speakers (p<0.0001 in reading and p<0.0001 in spontaneous speech) and statistically significant dissimilarity index (SDDD) differences for male speakers (p<0.0001 in reading and p<0.0001 in spontaneous speech) and female speakers (p<0.0001 in reading and p<0.0001 in spontaneous speech). Overall results of this research show that acoustic-statistical measurement of similarity and dissimilarity indices are a useful method in speaker recognition in forensic phonetic expertise. Further on, results show that speaking conditions should not be neglected in forensic phonetic cases.
References


Harrison, Philip, Peter French (2010). Assessing the suitability of BATVOX for UK Casework or Evaluation of the BATVOX automatic speaker recognition system for use in UK based forensic speaker comparison casework Part II. U Abstracts for the 19th Annual Conference of the International Association for Forensic Phonetics and Acoustics, Trier, Germany, Department of Phonetics, University of Trier, str. 13.


