Linguistic and Computational Methods in Forensic Linguistics

László Hunyadi, Kálmán Abari, Enikő Tóth

University of Debrecen
Department of General and Applied Linguistics
4010, Debrecen, Pf. 24.

hunyadi@llab2.arts.klte.hu
abarik@pmail.arts.klte.hu
teniko@pmail.arts.klte.hu

Keywords: forensic linguistics, experimental phonetics, digital hanganyag, statistical methods, computational soundanalysis, semantic analysis

The paper is a report on a case in forensic linguistics in which linguistic and computational approaches are combined to answer the question whether it can be proved if a digital recording has been tampered with. With the growing use of digital applications the chances of digital forgery are significantly increasing, accordingly, the detection of tampering with audio recordings is also becoming an important task for forensic linguists. In the given case, we assumed that the most straightforward way of tampering with the given digital audio recording might have been the removal of some material and so our aim was to identify the location of this kind of tampering in the file. Due to the complexity of the given task the approach presented is interdisciplinary: first, it uses a traditional semantic analysis to identify possible discontinuous segments of the recorded text, second, it introduces an experimental phonetic approach to identify cues of the digital cutting of the audio signal, third, it applies statistical calculations to specify the bit-level characteristics of audio recordings. The combination of these measurements proved to be quite helpful in answering the initial question, and the proposed new methodologies can be used in further areas of linguistics and computation.