# Exploring language variation and change through forced alignment of pronunciation variants

## Ioana Vasilescu & Ioana Chitoran

Large scale corpora are increasingly used to explore patterns of synchronic variation. A recently developed method for processing speech corpora (Adda-Decker & Lamel 1999 ; Adda-Decker & Snoeren 2011 ; Ernestus 2011) proposes to exploit lexicons built for speech recognition systems as testing ground for linguistic hypotheses. The method relies on the process of forced alignment of the manual transcription with the audio signal. This process can be used to address questions about the manifestation of synchronic variation, in particular reduction phenomena, and potentially related sound change. In our talk, we first describe the approach, and we then present two case studies from Romanian. The first one concerns the vowel alternation [e] - [ʌ] encountered in Romanian historically, and synchronically as a phenomenon observed in continuous speech (Chitoran et al., 2018). The specific question addressed is whether the synchronic phenomenon in Romanian continuous speech is simply the result of general vowel reduction, or it preserves some of the characteristics of the sound change, such as its apparent sensitivity to a labial consonantal context. The second case study concerns the deletion of the enclitic definite article -l in continuous speech (Vasilescu et al., to appear). We examine the role of speaking style and of surrounding segmental context in L-dropping and L-retention, and the potential for the reanalysis of an inflectional marker. We propose that the results obtained allow us to better understand the conditioning environment of the sound changes, and at the same time reveal language-specific articulatory and co-articulatory settings and dynamic patterns as they emerge and can be observed in continuous speech.