Contour clustering as an exploratory tool for intonation analysis

It is a challenge to understand the form-meaning relationship in f0 contours. Traditional (auto-segmental metrical) analyses are not seldomly applied to highly stylized laboratory speech, which leaves a gap in explaining f0 movements found in more naturalistic speech and often requires (learning) an annotation system. This also provides a threshold to do intonation analysis on underdescribed languages. The present demonstration proposes an exploratory workflow with the same aim of finding prototypical - i.e. phonologically underlying - contours. The approach is suitable for field recordings of any type and essentially requires only segmentation (additional annotation is helpful, though). The tool performs cluster analysis on time-series f0 data and provides the user with multiple ways of finding the optimal number of clusters. There are no principled restrictions to the language, the amounts of data, or the speech unit under investigation to obtain sensible results. Although more analysis is needed to come to phonological descriptions, contour clustering offers a fully data-driven and reproducible basis for further hypothesizing and testing in production and perception tasks. The demonstration shows how this method can be applied to freshly obtained (field) data. Participants are more than welcome to bring data (.wav and .textgrid) for analysis using contour clustering. The tool is freely available, so are the documentation and example datasets:

Web: https://constantijnkaland.github.io/contourclustering/

Article: https://doi.org/10.1017/S0025100321000049