Contribution ID: **30**   Type: **Oral presentation**

Challenges of heterogeneous data for building Linguistic Theory

Tuesday 1 October 2024 16:05 (20 minutes)

Linguistics thrives on data, whether it stems from small highly controlled laboratory studies or from large heterogeneous datasets. Speech technology is increasingly providing new and varied tools to test linguistic theories (from sound change to second language learning) on large scale data. This, however, does not come without its challenges. In this presentation, we address one of the key challenges for testing linguistic theory (such as diverging voicing systems within language families), posed by heterogeneous data: the frequent absence of linguistically formatted metadata.

Contribution length

Middle

Primary author: POPESCU, Anisia (LISN)

Presenter: POPESCU, Anisia (LISN)