

## **Talking data, making choices: Analytical decisions in the study of spoken interaction**

**Robbie Love**

Spoken corpus linguistics has made enormous strides in documenting and analysing everyday interaction, giving researchers unprecedented access to spoken discourse as it is used in a range of contexts. Yet the apparent ‘messiness’ of spoken interaction means that researchers face an array of methodological choices – some obvious, some less so – that may fundamentally influence the conclusions we draw. In this talk, I argue that the real analytical work of analysing spoken corpora lies not only in the choices that are universal to corpus research, but in the often underdiscussed choices, about what we count, how we annotate, and what units of analysis we choose, that are brought to the fore in the context of spoken interaction.

Drawing on examples from my own research on variation and change in spoken British English, I explore some of these critical choices. In the study of grammatical variation and change, what counts as a meaningful baseline for frequency analysis in spontaneous conversation? How (if at all) do we account for individual speaker variation? When it comes to pragmatic annotation, how much context do we need to be certain enough of our interpretations? And when we are comparing interactional genres, to what extent do externally defined register categories encourage us to assume a homogeneity in our data that may ignore important signals of variation?

These examples reveal broader questions about corpus design, representativeness, and annotation reliability, especially in relation to variation, change, and discourse-level phenomena. I will reflect on how spoken interaction data invites – and demands – a methodological sensitivity that is responsive to the ‘messiness’ of the data we are working with.