

Acceptability Judgement Tasks and Grammatical Theory

Abstract :: Tom S Juzek

The thesis considers various questions about acceptability judgement tasks (AJTs).

In Chapter 1, we compare the prevalent informal method of syntactic enquiry, researcher introspection, to formal judgement tasks. We randomly sample 200 sentences from Linguistic Inquiry and then compare the original author judgements to online AJT ratings. Sprouse et al., 2013, provided a similar comparison, but they limited their analysis to the comparison of sentence pairs and to extreme cases. We think a comparison at large, i.e. involving all items, is more sensible. We find only a moderate match between informal author judgements and formal online ratings and argue that the formal judgements are more reliable than the informal judgements. Further, the fact that many syntactic theories rely on questionable informal data calls the adequacy of those theories into question.

In Chapter 2, we test whether ratings for constructions from spoken language and constructions from written language differ if presented as speech vs as text and if presented informally vs formally. We analyse the results with an LME model and find that neither mode of presentation nor formality are significant factors. Our results suggest that a speaker's grammatical intuition is fairly robust.

In Chapter 3, we empirically compare regular AJT data to their Z-scores and ranked data. For our analysis, we test resampled data for significant differences in statistical power. We find that Z-scores and ranked data are more powerful than raw data across most common measurement methods.

Chapter 4 examines issues surrounding a common similarity test, the TOST. It has long been unclear how to set its controlling parameter δ . Based on data simulations, we outline a way to objectively set δ . Further results suggest that our guidelines hold for any kind of data.

The thesis concludes with an appendix on non-cooperative participants in AJTs.