Review of: "Trust is the best policy. Game theoretical analysis of bias in elicitation procedures in linguistics"

Prakash Mondal[1]

1 Indian Institute of Technology, Hyderabad

This article offers game-theoretic and decision-theoretic analyses of linguistic intuitions in linguistic theorizing and shows that the much-debated confirmation bias of linguistic intuitions is not as much a problem as it is assumed to be. The topic of the paper is definitely interesting and so are the analyses. There are some issues the paper has not addressed and I believe the discussion can be improved by appropriately tackling these issues.

First of all, I don't wish to push for a criticism of the game-theoretic and decision-theoretic analyses of linguistic intuitions in terms of binary decisions, naïve probabilities, and discrete values of payoffs, etc. These are trivial issues, as any kind of idealization brings with it some sort of simplification in the modeling involved. After all, this is a model, not reality. It needs to be seen how well the model fares in capturing the real-world scenario. That's something which can be empirically tested for refinements. What is more important is the question of why or how linguistic intuitions tend to be reliable, stable, or fairly consistent across subjects. The author hasn't engaged with the matter of psycholinguistic or cognitive mechanisms or evidence that can support the game-theoretic and decision-theoretic analyses of linguistic intuitions. One plausible candidate is the paper of Steven Gross (2020) entitled 'Linguistic intuitions: Error signals and the Voice of Competence' in the book 'Linguistic Intuition: Evidence and Method' edited by Samuel Schindler, Anna Drożdżowicz and Karen Brøcker. In this paper, Gross has argued that error signals generated during language processing as part of the psychological monitoring mechanisms help discern unacceptability in subjects. This is one way of securing judgments of unacceptability or even ungrammaticality. I'm sure there can be many such mechanisms that are stable enough over populations of language communities in such a manner that they ensure a fair degree of consistency. Any differences in judgements over complex linguistic examples (usually marked with ? or ?? or #, etc.) may arise from variation in these psycholinguistic mechanisms.

Second, regardless of whether or not the payoffs for different possibilities in the case of the confirmation bias of linguistic intuitions tend to converge on 0, the game-theoretic and decision-theoretic analyses themselves don't tell us why this should be so. Is there any fundamental principle of rationality or central processes of cognition or something socio-cognitive that is empirically motivated and actually governs the game-theoretic and decision-theoretic analyses and also underscores the said trust?

Third, I think the data collection of procedures of field linguists with the help of elicitation does not solely handicap linguists. It extends to field anthropologists or ethnographers too. Even here, one needs an explanation of how reliability in methodological practices is secured by somewhat universal or at least recurrent patterns of human behaviour and

principles of rationality. This warrants some discussion.

Finally, the paper needs to be thoroughly edited in order for certain typos to be eliminated. The rest is fine. Overall, this has been a good read, though.