To be grammatical or not to be grammatical - Is that the question?

EVIDENCE FOR GRADIENCE

"Grammar sets up a sharp division between a class G of grammatical sentences and a class G' of ungrammatical sentences." (Chomsky, 1955/1975)

Introduction and Background

Most contemporary grammar theories are based on a sharp distinction between grammatical and ungrammatical sentences. In fact, almost all (94%) acceptability judgments published in Linguistic Inquiry 2001-2010 are either starred or unmarked, i.e. indicated as completely unacceptable or fully acceptable, even though 81% of the articles include intermediate levels of acceptability.

Approach

To check whether non-linguists perceive the divide between G and G' as equally sharp, we randomly sampled 50 starred sentences (“* items”, members of G’) and 50 unmarked sentences (“OK- items”, members of G), all of them from LI-papers that employed a scale with three levels or more. We then had 80 non-linguists rate the 100 sentences, using a 7-point acceptability scale (1 “fully unnatural” to 7 “fully natural”). See Sprouse et al. (2013) for related work.

Analysis and Results

Participants’ ratings correlate reasonably well with the LI judgments (p_pb = .66, R^2 = .43), rarely flatly contradicting them. Only 3 *-items received a mean rating > 5, and only 4 OK- items received a mean rating < 3. Yet, there are important discrepancies between the LI judgments and the experimental ratings.

Crucially, experimental ratings cover the whole space on the experimental 7-point scale rather than clustering at the two extremes. The intermediate range is densely populated, both for individual ratings and for mean ratings. In fact, 43 out of 100 items received mean ratings between 3 and 5.

A sharp division between G and G’ predicts an S-curve as a noisy approximation of the corresponding step function. Instead, we find a steady linear increase from 1 to 7 and no sign of an S-curve:

![Figure 1: Mean ratings by the non-linguists (y-axis) for the 100 LI items (x-axis; in ascending order). Items with a mean rating of ≥ 3 and ≤ 5 are in red (*-item) or blue (OK-item).](image)

Discussion and Conclusion

The observed gradience is not due to aggregation, as evidenced by the fact that individual ratings also cover the entire scale, including the intermediate range. The observed gradience is also not due to performance, at least not entirely as suggested inter alia by Newmeyer (2003). The fact that most items with intermediate ratings are *-items (N=31) would force us to assume and explain grammatical illusions in the sense of Phillips et al. (2011).

We conclude that gradience is not an epiphenomenon or artefact and should be built into theories of grammar.


Many thanks to Mary Dalrymple and Greg Kuchanski for their valuable feedback on this project.