Grammaticality Judgments as Linguistic Evidence
Theoretical Implications

Brian Murphy

Language, Interaction and Computation Lab
Centre for Mind/Brain Sciences (CIMeC)
University of Trento

ESSLLI 2009, Bordeaux
Course Outline

- Notions of Grammaticality, and Current Practice in Linguistics
- Others Sources of Linguistic Evidence
- Scales for Measuring Grammaticality
- Methodology for Eliciting Judgements of Grammaticality
- Theoretical Implications
Safe Assumptions?

- Informants will always have to perform some kind of processing to convert their intuitive response to an utterance into some response modality.
- If the inner cognitive scale and the response modality have many common characteristics, you would expect intuitive responses to be recorded relatively faithfully.
- If the inner cognitive scale and response modality have few common characteristics, additional transformation/processing would be required, and you would expect more noisy data.
- Noisy data = spurious variance = weak statistics.
The Paradox

- Binary, 7-point categorical Likert, and Magnitude Estimation are all equally noisy.
- Is the real cognitive scale one of these, ‘between’ them, or something else?
Steps towards a solution

- [Armstrong et al., 1983] demonstrates that the nature of cognitive representations can be obscured by task-dependent functions.
- In their case, the internal function used might have looked something like:
  - response = f(parity, 1/magnitude)
- This is easy, because we already know what even numbers are (because we invented them).
- So, to solve our problem, we have to find the function(s) that apply to intuitions of grammaticality.
- ... general agreement?
A binary scale

- Grammars in the generative tradition hold that grammaticality is binary, due to syntax alone.
- Performance factors cloud this competence:
  - response = f(grammaticality, lexical frequency/familiarity, sentence/phrase length, sentence/phrase complexity/depth, plausibility, ...)
- This account is hard to test, since little work has gone into discovering and quantifying the effects of performance factors.
A ranked dichotomous scale

- [Sorace and Keller, 2005] suggest that there are hard and soft constraints

- All hard constraints are more severe (i.e., cause a bigger loss in grammaticality if violated) than soft constraints
  - Soft constraints are sensitive to context, hard constraints are not
  - [hinted at: hard constraints are syntactic; soft constraints are other stuff]

- Constraints are also ranked within each class

- Violations are cumulative
  - response = f(grammaticality) = f(sum({violated constraint weights}))

- Difficult to test, because we would need to understand the grammar, quantifying the weight of all constraints
A continuous scale I

- [Bader and Häussler, ] assume intuitive grammaticality is continuous
- Most sensations and opinions are (agreement?)
- When asked for binary judgements, informants apply an arbitrary cut-off threshold
- The appeal to this feature of Signal Detection Theory
A continuous scale II

- Evaluation:
  - find what ME threshold partitions the experimental sentences into the same proportions of un/grammatical as the (speeded) binary judgements (across all informants)
  - see if these thresholds work for individual experimental conditions

---

**Figure 7**
Plot of observed means (SGJ part of Experiment 2) and means predicted by the SDT model described in the text. The sixteen conditions of the three subexperiments are numbered in accordance with the result tables in Appendix A.
Frameworks?

- Anything possible!
  - Variations on OT [Boersma, 2004, Keller, 2006]
  - Harmonic Theory [Legendre et al., 1990]
  - Linear/weighted theories
    [Bresnan and Nikitina, 2003, Blache et al., 2006, Bresnan et al., 2007]

- But most of them involve ditching the foundational assumptions of Chomskyan grammars
Summary I

- Range of views on grammaticality, and how it arises
- Traditional practice is particularly dangerous because linguists (often) generate and evaluate data for theories that they generate and evaluate
- Linguists disagree surprisingly often
- Cognitive science is starting to ignore linguistics
- Experimental psycholinguistics and neurologistics have very powerful methods, but they are not sensitive enough for large-scale measurement of grammaticality
- So intuitions remain the best methodology, particularly since they demand less interpretation (we can ask informants what they did)
- The choice of measurement scale used is not very important, so use the one that suits your informants
- So for practical purposes, it is independent of your theory
To improve your experiments: use realistic materials, use more materials, ask more people, give clear instructions with training/calibration/practice phase, ask experts and non-experts, use statistics, report methods

We really have no (empirically supported) idea what the internal scale of grammaticality looks like, nor how it relates to grammar

... which means plenty for young ambitious linguists to do, and publish!
Thanks

... to everyone for the enthusiastic participation, lively discussion, and new ideas

If you have ideas for work in this area, please come and talk to me about it, e.g.:

- building a more comprehensive software infrastructure for creating, managing and analysing grammaticality experiments
- more stress-testing of alternative scales
Acknowledgments

- Doctoral supervisor at Trinity College Dublin, Carl Vogel
- Fellow group members Roberto Zamparelli, Marco Baroni, Massimo Poesio and the participants of our Linguistic Evidence seminar
- Current cabal of Linguistic Methodologists, Sam Featherston, Tom Weskott, Gisbert Fanselow, Markus Bader, Jana Häussler, Jon Sprouse, and probably some others I’ve forgotten
- All the anonymous informants who have taken part in our experiments
References


Bader, M. and Häussler, J. Toward a model of grammaticality judgments. under review.


Boersma, P. (2004). A stochastic OT account of paralinguistic tasks such as grammaticality and prototypicality judgments. Unpublished manuscript.


Linear optimality theory as a model of gradience in grammar. 

Harmonic grammar – a formal multi-level connectionist theory of linguistic well-formedness. 

Gradience in linguistic data. 