

## Magnitude Estimation and the (non-)Linearity of Acceptability Judgments

Keywords: magnitude estimation, experimental syntax

The term *experimental syntax* generally describes the use of psycholinguistic methodologies for the collection of acceptability judgments, and can cover any number of designs, tasks, and statistical analyses (Cwart 1997, Schütze 1996). Over the past decade, one task in particular, the magnitude estimation task, has received significant attention (Bard et al. 1996, Keller 2000, and Keller 2003). While magnitude estimation has been a staple task of psychophysics for over 50 years (Stevens 1957), it only became part of linguistic methodology when Bard et al. (1996) demonstrated how it could be profitably adapted for the collection of acceptability judgments. And while magnitude estimation has been applied to a number of areas of syntactic research in the ensuing decade, there has been little research into the task itself since that seminal paper. Because magnitude estimation was originally developed to measure the perception of physical stimuli, there are certain assumptions built into the magnitude estimation task that may not be compatible with the perception of linguistic acceptability. In this talk, I investigate two of these assumptions – that participants can report acceptability judgments in a linear system, and that acceptability is a continuous quantity without any categorical distinctions – through a meta-analysis of the roughly 20 magnitude estimation experiments originally presented in Author 2007. These experiments suggest that both assumptions are incompatible with acceptability, or in other words, that acceptability is a non-linear system. These findings force a reconsideration of some (but not all) of the methodological reasons offered for the widespread adoption of magnitude estimation, as well as some of the analyses that have been built around the continuous (i.e., non-categorical) appearance of acceptability under magnitude estimation tasks (e.g., Keller 2000, 2003, Sorace and Keller 2004).

*Assumption 1: Participants have the ability to estimate the magnitude of acceptability*

The assumption that participants can precisely estimate the magnitude of acceptability is rooted in the successful estimation of perceptual stimuli in psychophysics. However, there is a crucial difference between the two types of stimuli: perceptual stimuli have a physical zero point that corresponds with the absence of the stimulus, while acceptability has no obvious zero point. While this fact has been acknowledged in every linguistic magnitude estimation study beginning with Bard et al. 1996, the linguistic version of the task always instructs participants to give ratio responses (x is twice as acceptable as y), which necessitates a zero point. Given the lack of a true zero point, participants must create their own (personal) zero point. In other words, the quantified distances between stimuli reported by participants are, at most, stable for that participant and for that modulus sentence only. This predicts that, unlike in psychophysics, linguistic magnitude estimation experiments cannot predict the values of replications with different modulus sentences. Figure 1 reports the results of just such a prediction between two replications, confirming that there is no prediction when the modulus is changed.

*Assumption 2: Acceptability is continuous (i.e., non-categorical)*

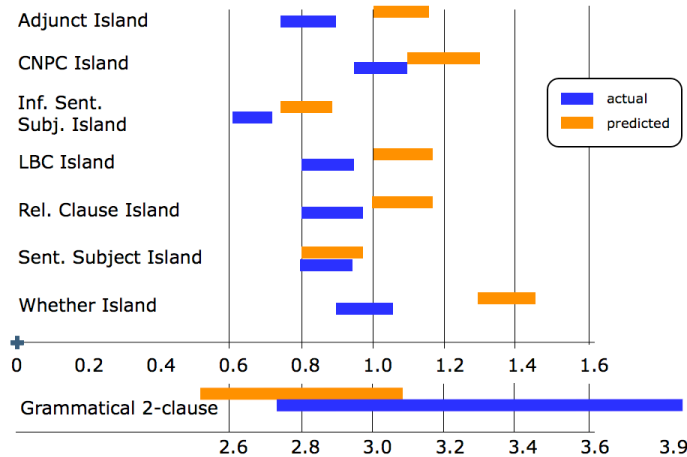
Magnitude estimation assumes that there is a single spectrum on which the stimuli are measured without any (psychological) categorical differences. For most physical stimuli this is an obvious truth. However, many syntactic theories suggest that grammatical and ungrammatical sentences are categorically different. While the continuous nature of the response scale in magnitude estimation gives the appearance of continuous acceptability, the categorical difference in grammaticality manifests in two ways. First, Levene tests for homogeneity of variance on all of the matched conditions in the experiments analyzed reveal that conditions that are (by hypothesis) grammatical have significantly different (and higher) variances than minimally different conditions that are ungrammatical. Second, the matched experiments reported above for assumption 1 suggest a tendency for ungrammatical sentences to be rated below the modulus even when the modulus is less acceptable than the items being tested (Figure 2).

These facts about acceptability directly affect the interpretation of magnitude estimation data, from suggesting that the calculation of weighted constraints will not hold across experiments (e.g., Keller 2003), to suggesting that continuous acceptability does not necessarily preclude categorical grammaticality (e.g., Keller 2000). They also suggest that many of the mathematical reasons for adopting magnitude estimation do not hold, although additional reasons that do hold will be offered.

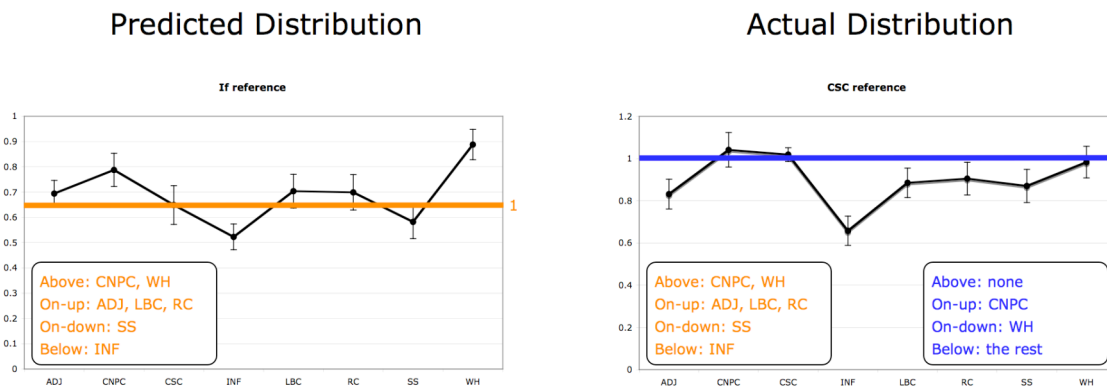
## Magnitude Estimation and the (non-)Linearity of Acceptability Judgments

Keywords: magnitude estimation, experimental syntax

**Figure 1:** Predicted range of values (orange) based on experiment 1, actual range of values (blue) based on experiment 2



**Figure 2:** Predicted distribution around the modulus (orange) based on experiment 1, actual distribution around the modulus (blue) based on experiment 2



## References

- Bard, Ellen Gurman, Dan Robertson, and Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72:32–68.
- Cowart, Wayne. 1997. *Experimental syntax: Applying objective methods to sentence judgments*. Sage.
- Featherston, Sam. 2005a. Magnitude estimation and what it can do for your syntax. *Lingua* 115.
- Keller, Frank. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Doctoral Dissertation, University of Edinburgh.
- Keller, Frank. 2003. A psychophysical law for linguistic judgments. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society*.
- Sorace, Antonella, and Frank Keller. 2005. Gradience in linguistic data. *Lingua* 115:1497–1524.
- Schütze, Carson. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. The University of Chicago Press.
- Stevens, Stanley Smith. 1957. On the psychophysical law. *Psychological Review* 64:153–181.