

Gradient acceptability between naïve and expert linguistic intuitions

Jacee Cho¹, Yafei Li², Rebecca Shields²

¹Department of English and Language Sciences, University of Wisconsin-Madison, US

²Language Sciences, University of Wisconsin-Madison, US

<https://doi.org/10.36505/ExLing-2021/12/0013/000486>

Abstract

The current study examines non-linguists' judgments about one structure which was reported in Li et al. (2012): interaction of adverbs in *wh*-movement which involves rare sentence types. Acceptability judgment data collected from 199 nonlinguists differ from the expert intuitions reported in Li et al. (2012). We argue that sentence acceptability judgments are necessarily influenced by extraneous factors; therefore, studying sentences that are hard to process or rare requires introspective judgments from linguists who can carefully contrast sentences under investigation by ignoring irrelevant information.

Keywords: gradient acceptability, introspective judgments

Introduction

Some recent empirical studies have reported that nonlinguists' judgments may differ from those of linguists or those reported in the syntactic literature (Dąbrowska 2010; Gibson & Fedorenko 2013). In particular, Gibson and Fedorenko (2013) showed cases where linguists' intuitions reported in the literature are not shared by the general population. Gibson and Fedorenko (2013) proposed a number of possible reasons why linguists' judgments might be faulty and argued that quantitative data collected from a large number of naïve participants should be used to study syntax. Following this line of inquiry, the current study examines non-linguists' judgments about one structure which was reported in Li et al. (2012): interaction of adverbs in *wh*-movement (Section 2). Acceptability judgment data collected from 199 nonlinguists differ from the expert intuitions reported in Li et al. (2012). However, we do not think that these findings necessarily indicate that linguists' judgments reported in Li et al. (2012) are faulty or biased. Rather, we argue that linguists are better at ignoring irrelevant factors in judging grammaticality, in particular implausible or rare but grammatical sentences. Therefore, we argue that experts' introspective judgments are necessary to study certain types of sentences.

Linguistic phenomenon

A well-known phenomenon in the study of syntax is Relativized Minimality (Rizzi 1990), which states roughly that a syntactic constituent cannot move across another constituent of the same type. This principle was claimed to affect adverb topicalization in a variety of languages in Rizzi (2001), based on the fact that the topicalization of an adverb is blocked by the presence of an intervening adverb, as in (1).

- (1) Topicalization: A' movement blocked by an intervening A' modifier
- a. Noisily, Sam (*frequently) mows his lawn.
 - b. Noisily, Sam (*apparently) mows his lawn.

Crucially, this principle is “Relativized” according to syntactic type, so that elements in argument (A) position block movement to another argument position, and elements in non-argument (A') position block movement to another non-argument position. Li et al. (2012) reported a more complex pattern of blocking with *wh*-movement of adverbs (1). Specifically, they noted that adverbs which are not themselves capable of undergoing *wh*-movement do NOT block *wh*-movement across them, despite being located in an A' position (1a). This pattern is unexpected assuming traditional Relativized Minimality, which would have predicted both (2a) and (2b) to be unacceptable.

- (2) Wh-movement: A' movement blocked by some intervening A' modifiers, but not others
- a. How noisily does Sam (*frequently) mow his lawn?
 - b. How noisily does Sam (apparently) mow his lawn?

This study was based largely on the judgements of four native speakers of English, who were either PhD students or faculty members in linguistics.

Methodology

We subsequently tried to replicate the empirical findings of this earlier study on a larger scale ($n = 199$) using an acceptability judgment task with a 7-point Likert scale. To test the claim about the interaction between specific adverb combinations and *wh*-movement, we designed an experiment by manipulating two factors, sentence type (*wh*-question vs. declarative) and adverb combination (wh-wh vs. nonwh-wh), as illustrated in Table 1. The target condition is shaded.

Table 1. 2x2 factorial design.

Sentence type Adv combination	Declarative	Wh-question
Wh-wh	Lucas frequently sprayed the paint thickly.	*How thickly did Lucas frequently spray the paint?
Nonwh-wh	Lucas apparently sprayed the paint thickly.	How thickly did Lucas apparently spray the paint?

Results

All scores including fillers were transformed into z-scores for statistical analysis. The z-score mean for each of the four conditions is presented in Figure 1.

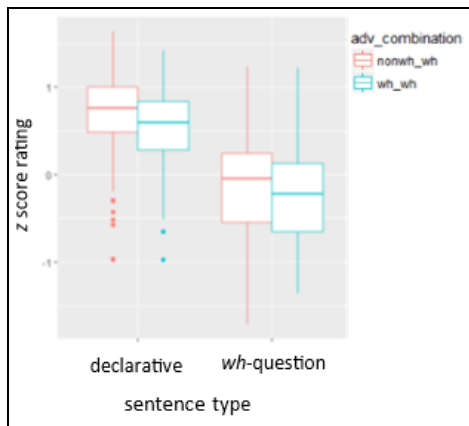


Figure 1. Mean z-scores for each condition.

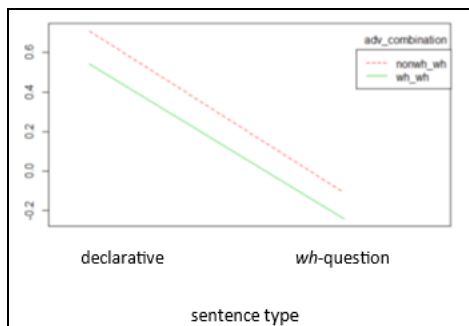


Figure 2. Interaction plot.

A linear mixed-effects regression analysis was conducted on z-transformed rating scores with Sentence type (*Wh*-question vs. Declarative) and Adv. Combination (Frequency-Manner (wh&wh) vs. Evidential-Manner (nonwh & wh)) as fixed factors. Subject and Item were included as random factors. Results indicated that there were significant main effects of Sentence type (Estimate=-.7, SE=.09, df=68, t=-8.59, p <.0001) and Adv. Combination (Estimate=-.2, SE=.07, df=207, t=-2.9, p <.003). Declarative sentences were rated higher than *wh*-questions and sentences containing evidential & manner adverbs were rated higher than sentences with frequency (wh-adverb) & manner (wh-adverb) adverbs. However, there was no significant interaction effect of these two factors, as shown in the interaction plot and the differences-in-differences (DD) score in Figure 2.

Discussion and conclusion

Gradient acceptability judgments between linguists and nonlinguists are largely due to the fact that they attend to different types of information in judging the acceptability/grammaticality of sentences (Francis, in press). Since sentence acceptability judgments are necessarily influenced by extraneous factors such as sentence complexity and frequency of occurrence “we cannot rely on non-expert judgements in the absence of theory-based reasoning and carefully-controlled stimuli” (Juzek et al., 2020, p.6). Therefore, it is necessary to have both quantitative data from non-linguists and judgments from linguists who can carefully contrast sentences under investigation by ignoring irrelevant factors.

References

- Dąbrowska, E. 2010. Naïve v. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review* 27, 1-23.
- Francis, E. (in press). *Gradient Acceptability and Linguistic Theory*. Oxford University Press.
- Gibson, E., Fedorenko, E. 2013. The need for quantitative methods in syntax and semantics research. *Language and Cognitive Processes* 28(1/2), 88-124.
- Juzek, T., Francis, E., Häußler, J. 2020. Syntax-semantics interferences: making a case for expert reasoning and multi-method approaches. *The Limits of Experimentation: Current Challenges in Experimental Linguistic Practice*.
- Li, Y., Shields, R., Lin, V. 2012. Adverb classes and the nature of minimality. *Natural Language & Linguistic Theory* 30(1), 217-260.
- Rizzi, L. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- Rizzi, L. 2001. Relativized minimality effects, in Mark Baltin and Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell. 89-110.