
Markus Bader, Jana Häussler
University of Konstanz, University of Potsdam
markus.bader@uni-konstanz.de, jana.haeussler@uni-potsdam.de

From graded ratings to binary decisions: A case study on argument alternations in German

The growing use of experimental methods in linguistics has led to a widespread recognition that grammaticality is a gradient phenomenon. With regard to language performance, this raises several important questions: (i) How does graded grammaticality relate to language comprehension? In particular, is graded grammaticality a consequence of the fact that the perception of grammaticality is mediated by the language comprehension mechanisms? (ii) How does graded grammaticality relate to language production? In particular, do production frequencies determine degree of grammaticality? (iii) What are the processes leading to traditional binary grammaticality judgments?

We present a formal model addressing all three questions. The major claims of our model are as follows: (1) Graded grammaticality is a property of the grammar itself (e.g., Pater 2009) and not an epiphenomenon caused by performance mechanisms (e.g., Hawkins 2006). (2) During language production, the probability of a sentence is determined *inter alia* by its continuous grammaticality score. That is, in contrast to many usage-based models (see Manning 2003), we claim that graded grammaticality determines usage and not the other way round. (3) During language comprehension, sentences are automatically assigned continuous grammaticality values. In contrast to models denying a direct relationship between graded and binary judgments (Featherston 2005), we hypothesize that, when required by the task, continuous values are mapped directly onto binary grammaticality judgments in the way of Signal Detection Theory (Green & Swets 1966).

The evidence for this model comes from a series of experiments and an associated corpus study. The experiments investigated identical sentence materials with different judgment procedures: magnitude estimation (ME) and binary grammaticality judgments (BGJ). The sentence materials covers 120 ditransitive verbs in six syntactic contexts according to the factors *Voice* (active, regular passive and *bekommen* ('get') passive) and *Number of arguments* (2 or 3 arguments). In the following, we concentrate on the *bekommen* passive which is known to be subject to lexical variation (see (1)).

- (1) *Ich glaube, dass Peter ein Buch geschenkt bekam / ??gestohlen bekam.*
I believe that Peter a book presented got stolen got.
'I believe that Peter was presented/stolen a book.'

The corpus study is based on the *deWaC* corpus described in Baroni et al. (2009), a huge corpus of German built by web crawling. With about 1,278,177,539 tokens, the *deWaC* corpus is of a size sufficient for present purposes. At the same time, this corpus

avoids the disadvantages brought about by querying the web directly by means of one of the popular search engines.

The experiments and the corpus study confirm this lexical variation for the 120 verbs in the *bekommen* passive:

- ME (z-transformed acceptability scores): range = -1.63 – 2.15
- BGJ (percentages of grammatical judgments): range = 19% - 100%
- Corpus frequencies: range = 0-2500

The direct mapping of graded grammaticality scores to binary grammaticality judgments is confirmed by a logistic regression analysis predicting BGJ results from ME scores which revealed a reasonably good fit ($C = .79$). The claim that gradient grammaticality is not caused by performance mechanisms is based on the finding of gradient judgments for sentences which are closely matched with respect to syntactic complexity, length and construction frequencies.

Several pieces of evidence support the claim that grammaticality is among the factors determining frequency and not the other way round. First, high grammaticality was found to be independent of frequency whereas low grammaticality was consistently associated with low frequency. Second, a poisson regression analysis predicting syntactic frequencies from verb frequencies and ME scores showed that both predictor variables lead to significant and substantial reductions in deviance. Third, we run a preference experiment along the lines of Bresnan & Ford (2010). Participants read a context question establishing a topic and then had to choose between two syntactic realizations of the answer (active voice versus *bekommen* passive). When the underlying dative object was established as topic (see (2)), the preferred answer was the *bekommen* passive variant which allows the topic to be realized as a subject.

- (2) Was gibt es Neues von eurem Opa? ('What's about your grandpa?')
- Topic = subject (*bekommen* passive)
Unser Opa hat von Robert einen Rasenmäher überreicht bekommen.
our grandpa has by R. a lawnmower handed-over got
'Our grandpa was handed over a lawnmower by Robert.'
 - Topic = dative object (active)
Unserem Opa hat Robert einen Rasenmäher überreicht.
our grandpa has R. a lawnmower handed-over
'To our grandpa, Robert handed over a lawnmower.'

Importantly, the choice of *bekommen* passive answers was modulated by the verb-specific grammaticality scores estimated from the prior judgment experiments. For verbs that had received degraded grammaticality scores in the *bekommen* passive, the *bekommen* passive answer was chosen less often. Frequency information also had an influence, although a smaller one. For verbs that were judged as fully compatible with the *bekommen* passive, the probability of choosing the *bekommen* passive variant was higher for verbs with higher corpus frequency for the *bekommen* passive.

References

- Baroni, Marco, Bernardini, Silvia, Ferraresi, Adriano and Zanchetta, Eros 2009. The WaCky Wide Web: A collection of very large linguistically processed web-crawled corpora. *Language Resources and Evaluation Journal* 23, 209-226.
- Bresnan, Joan and Ford, Marilyn 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86, 168-213.
- Featherston, Sam 2005. The decathlon model of empirical syntax. In *Linguistic evidence*, Reis, Marga and Kepser, Stephan (eds.), 187-208. Berlin: de Gruyter.
- Green, David M. and Swets, John A. 1966. *Signal Detection Theory and Psychophysics*. New York: Wiley.
- Hawkins, John 2006. Gradedness as relative efficiency in the processing of syntax and semantics. In *Gradience in grammar: Generative perspectives*, Fanselow, Gisbert, Féry, Caroline, Vogel, Ralf and Schlesewsky, Matthias (eds.), 207-226. New York: Oxford University Press.
- Manning, Christopher D. 2003. Probabilistic syntax. In *Probabilistic linguistics*, Bod, Rens, Hay, Jennifer and Jannedy, Stefanie (eds.), 289-341. Cambridge, MA: MIT Press.
- Pater, Joe 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33, 999-1035.
-