Experiencers from different perspectives: Comparing scalar and binary judgments

Anne Temme & Elisabeth Verhoeven

Humboldt-Universität zu Berlin

`anne.temme(at)posteo.de; verhoeve(at)cms.hu-berlin.de`

1   Preliminaries

A central question in the discussion about the reliability and necessity of empirical data is the representational power of linguistic data measurement. As pointed out in Featherston (2005) and replied by following studies, different data types relate to different kinds of underlying cognitive processes [1], [2], [3], [4], [5]. In short, corpus data represent the result of a selection process of competing structures whereas scalar (acceptability) judgments provide evidence for cognitive workload. The distribution of these two data types usually follows a typical pattern: the most acceptable structure is the most frequent and the less acceptable structure is the least frequent. At the same time, corpus-based frequencies are known to be determined by the power law which results in high frequencies of the optimal structures and compressed contrasts between the lower-ranked structures; In contrast, judgments lead to finer-grained distinctions between the corresponding data points with less extreme values for different conditions.

The aim of our study is to test the assumptions associated with *output-selection* (characteristic for corpus data) vs. scalar judgments under maximally controlled conditions. Instead of using corpus data, we rely on frequency data generated through a forced-choice task, i.e. a forced selection between two competing structures. We assume that both types of data, i.e. corpus data and forced-choice data are similar in constituting the selection of competing structures.

The comparison of the methods generally reflects underlying mechanisms of different output types but also the nature of constraints for the relevant linguistic phenomenon to be introduced for our case in the next section.

2   Phenomenon and expectations

In languages with flexible word order, argument linearization may be contextually determined. Furthermore, in languages like German, certain verb types are associated with non-canonical argument orders, as e.g. experiencer-object verbs, which license object-before-subject order (O>S) without further contextual licensing, e.g. [6]. As for other predicates, e.g. non-experiential action verbs, the change of the argument order needs to be contextually licensed.

In order to test these assumptions and taking into account the methodological question outlined above, we conducted two experimental studies, a relative judgment task and a two-alternatives forced choice test. Both studies examine the influence of the factors CONTEXT (licensing vs. non-licensing) and VERBTYPE (object-experiencer

vs. canonical verbs) on the dependent variable ARGUMENT ORDER (S>O, O>S). For both studies, we assume that the factors CONTEXT and VERBTYPE and their interaction significantly affect the order of arguments. The four relevant conditions are: non-licensed experiencer, licensed experiencer, non-licensed non-experiencer, licensed non-experiencer structures.

Regarding the two methods, we expect the typical distribution of the data types described above: the ranking of the conditions should align for the two alternative measures but we should also find the typical patterns representing the different underlying processes, i.e. the intuition of well-formedness on the one hand and output-selection on the other hand.

3 Material and methods

We examine the conditions that license OS orders with German experiencer-object verbs. For the examination of scalar intuitions, we collected relative (instead of absolute) judgments in order to observe speakers' intuitions in comparing structures, which is the assumed step before the output-selection process. For the examination of the output-selection process we conducted a two-alternatives forced choice test.

As licensing context, we implemented a set-member relationship between the subject of the context sentence and the object of the target sentence and a contrastive reading between the predicates of the two structures [7]. The non-licensing context was represented by an 'all-new' context "Was gibt es Neues?" ('What's new?'). We tested 16 verbs for each class. Both experiments have identical material and factorial structure. We conducted a web-based study[1] with 32 subjects, respectively. For the forced choice task, subjects were presented two alternative target sentences and were instructed to choose the most coherent one in a given context. For the relative judgment, subjects were presented the same two targets within the same context and were instructed to award points to both alternatives (e.g. 50/50, 0/100, 80/20, etc.). See (1) for an example of the presented structures.

4 Results

Regarding the linguistic hypotheses, we found significant VERBTYPE and CONTEXT effects in both studies[2]. The occurrence of the effects in both cases confirms the O>S licensing power of the relevant factors. We found no interaction of the factors, which indicates a (partly) different nature of the kind of O>S licensing of context and the O>S licensing of predicates.

Regarding the methodological hypotheses, the basic distribution is as expected. The condition that leads to the highest frequency of O>S also represents the most acceptable O>S structure. See Table 1. The lowest frequency aligns with the lowest rating. In addition, the comparison of the results of both methods reveals a difference in effect size for at least two conditions. Due to the parallel material and method, these differences cannot be associated with non-controlled factors, but rather should be related to the underlying cognitive processes. On the other hand, we find a

---

[1] OnExp Version 1.2; http://onexp.textstrukturen.uni-goettingen.de
[2] We analyzed the data with mixed effect regression analyses, assuming Subject and Item as random factors.

high occurrence probability for more than one condition and no condition with highest or zero frequency, a pattern that differs from the general characteristics of corpus data. Thus, forced choice frequencies reveal properties of judgment data reflecting gradient well-formedness as well as strong preferences similar to corpus data, reflecting an output-selection process, cf. [3].

From the linguistic perspective, we provide evidence that the OS order with experiencer-object verbs does not require a contextual licensor. Since there is an alignment of effects on well-formedness and choice probabilities, it reinforces the existence as well as the ranking of the effects of the underlying factors of the phenomenon.
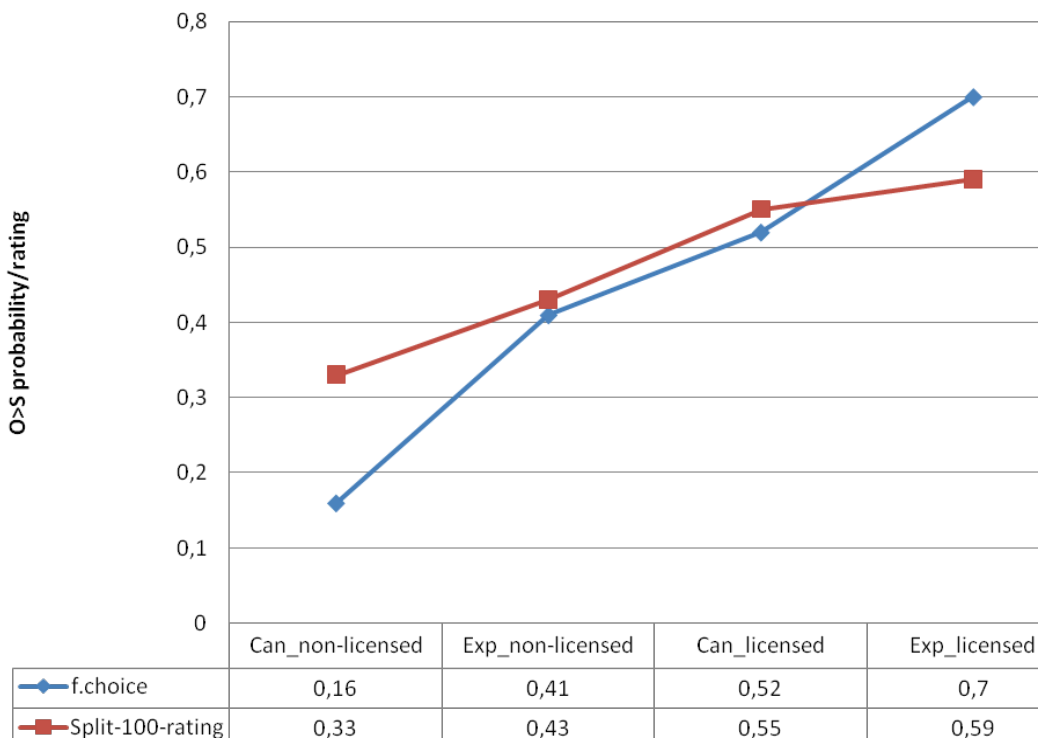
Appendix

(1)   a. *Alternative I*
      [Die meisten Bürger] hatten keine Probleme mit dem Bahnübergang.
      S>O: Die Schranke hat [den Pfarrer] aufgeregt$_{EXP}$/ aufgehalten$_{CAN}$.
    b. *Alternative II*
      [Die meisten Bürger] hatten keine Probleme mit dem Bahnübergang.
      O>S: [Den Pfarrer] hat die Schranke aufgeregt$_{EXP}$/ aufgehalten$_{CAN}$.
      'Most of the citizens had problems with the railway crossing.'
      'The pastor was upset/ delayed by the barrier.'

Figure 1: O>S probability/split-rating in comparison

	Can_non-licensed	Exp_non-licensed	Can_licensed	Exp_licensed
f.choice	0,16	0,41	0,52	0,7
Split-100-rating	0,33	0,43	0,55	0,59

Method: Logistic Mixed Effect Regression fit by AIC and Maximum likelihood; Number of obs: 512; groups: item, 32; subject, 32; Formula: word.order ~ verbtype + context+ (1|subject) + (1| item)

Factors	Estimate	Standard error	p-value
Fixed effects			
Intercept	-1.2148	0.2853	*p<.001*
Verb type(non-experiencer)	0.9907	0.2262	*p<.001*
Context(non-licensing)	1.7998	0.2193	*p<.001*
Random effects	**Standard deviation**		
Items(intercept)	0.2078		
Subjects(intercept)	1.1619		

Table 1: Results of the Forced choice test

Method: Linear Mixed effect regression; Number of obs: 371[3]; groups: item, 32; subject, 32; Formula: word.order ~ verbtype + context+ (1|subject) + (1| item)

Factors	Estimate	Standard error	t-value
Fixed effects			
Intercept	54.545	3.337	16.348
Verb type(non-experiencer)	-7.991	2.562	-3.119
Context(non-licensing)	-25.733	2.463	-10.446
Random effects	**Standard deviation**		
Items(intercept)	2.169		
Subjects(intercept)	14.522		

Table 2: Results of Split-100 test[4]

References

[1] Arppe, A. & J. Järvikivi (2007). Every method counts: Combining corpus-based and experimental evidence in the study of synonymy.

[2] Bresnan, J. (2007). Is syntactic knowledge probabilistic? Experiments with the English dative alternation.

[3] Featherston, S. (2005). The Decathlon Model of empirical syntax.

[4] Jónsson, Jóhannes Gísli. 2009. Covert nominative and dative subjects in Faroese. Nordlyd 36, 142-164.

[5] Kempen, G. & K. Harbusch (2005). The relationship between grammaticality ratings and corpus frequencies: A case study into word order variability in the midfield of German clauses.

[6] Lenerz, J. (1977). Zum Einfluß von "Agens" auf die Wortstellung des Deutschen.

[7] Weskott T., Hörnig, R., Fanselow G. & R. Kliegl (2011). Contextual Licensing of Marked OVS Word Order in German.

---

[3] Due to statistical analysis, the 50/50 values have been deleted. 72,5% of the data is left. Note that the values in figure 1 contain the 50/50 values.

[4] The model was fitted via ML-method. VERBTYPE and CONTEXT are significant predictors for the results, but not their interaction.