

## The likelihood of upper-bound construals among different modified numerals

Stavroula Alexandropoulou, Yaron McNabb, Sophia Bimpikou, Dominique Blok, and Rick Nouwen  
Utrecht University

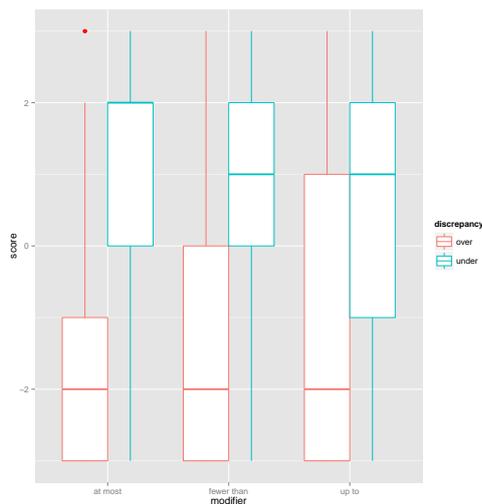
Numeral modifiers have received a lot of attention lately. Geurts & Nouwen (2007), for example, show that superlative modifiers like *at most* give rise to ignorance inferences. Nouwen (2010) argues that the absence or occurrence of such inferences is the hallmark of an essential distinction between two classes of scalar quantifiers. In this paper we show that modified numerals do not just differ with respect to such epistemic effects, but also to the nature of the lower or upper bound they express. We studied expressions like *fewer than*, *at most* and *up to*, and investigated to what extent such modifiers impose an upper bound. We did so in an experiment as well as in a pilot follow-up. We found that *at most* and *fewer than* are more likely than *up to* to give rise to an upper-bound construal and explored additional contextual factors that affect the strength of such inferences.

**Motivation** — Schwarz *et al.* (2012) observe that negative polarity items are licensed in the scope of *at most*, but not in that of *up to*. Based on this and other such observations, they conclude that there must be a fundamental semantic difference between these two modifiers. Based on armchair data, Blok (2015) argues that this crucial difference is (in part at least) due to the fact that the upper bound expressed by *up to* (and other so-called directional numeral modifiers) is implicated rather than entailed. Roughly, Blok’s proposal is that while *at most*  $n$  denies the existence of occurrences of values higher than  $n$ , *up to*  $n$  asserts the existence of values between some implicit lower bound and  $n$ . Higher values are only excluded by implicature.

**Experiment 1** — In our first experiment we used target items like (1) and asked subjects to rate the naturalness of such pairs of sentences occurring together, on a  $-3$  to  $+3$  Likert scale.

(1) Every student read { fewer than / at most / up to }  $n$  papers. One of them read  $m$  papers.

The experiment was conducted in Greek. Conditions differed with respect to choice of modifier and with respect to whether  $m < n$  (the ‘under’ conditions) or  $m > n$  (the ‘over’ conditions). All target items ( $N=12$ ) were rotated through 6 lists, so that each participant only saw one condition per item. Fillers were added too. Both Schwarz *et al.* and Blok assume in their analyses that *at most* provides an upper bound entailment. This predicts that subjects are more likely to approve of an ‘over’ item when the modifier is *up to* than when it is *at most*.



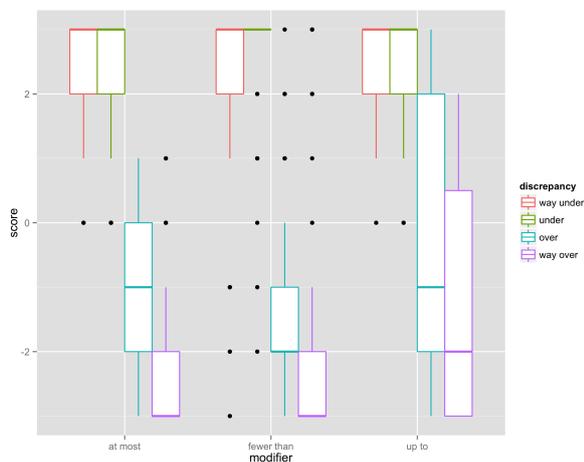
Indeed this is what we found. This boxplot summarises the data by 77 participants, analysed with mixed-effects ordered probit regression models with random effects for subjects and items. ‘Up to-Over’ items were rated significantly higher than ‘At most-Over’ items ( $\beta=-.284$ ,  $SE=.124$ ,  $p < .05$ ). ‘Fewer than-Over’ items received scores significantly lower than ‘Up to-Over’ items ( $\beta=.309$ ,  $SE=.124$ ,  $p < .05$ ). On the natural assumption that the upper bound provided by *fewer than* is entailed, this result can be interpreted to indicate that the upper bound for *up to* is derived via pragmatics. This confirms Blok’s (2015) relevant prediction.

**Experiment 2** — In a follow-up pilot study in English, we chose not just to have the type of modifier as a factor, but also what we would like to refer to as *scalar distance*. That is, as in

experiment 1, we compared conditions with *at most* to conditions with *up to*, but we had the distance between  $n$  and  $m$  as an additional factor. The task was also different. Experiment 2 asked subjects to rate to what extent a given statement was compatible with a subsequently provided FACT. For instance, subjects had to rate whether the fact in (2-b) is compatible with the CLAIM in (2-a) on a  $-3$  to  $+3$  Likert scale.

- (2) a. CLAIM: Clarendon High School used its smart classrooms 50 times last year, with {  $\emptyset$  / at most / up to } 39 students participating in this virtual classroom environment.  
 b. FACT: On one occasion the smart classroom was used at Clarendon High School last year, { 10 / 37 / 41 / 68 } students participated.

As in Experiment 1, conditions differed with respect to the choice of modifier (no modifier, *up to*, *at most*), but here we added two ‘under’ and two ‘over’ conditions, dubbed as ‘under’ ( $m_{fact} = n_{claim} * 0.95$ ) and ‘way under’ ( $m_{fact} = n_{claim} * 0.25$ ), and ‘over’ ( $m_{fact} = n_{claim} * 1.05$ ) and ‘way over’ ( $m_{fact} = n_{claim} * 1.75$ ), (see (2-b)). The idea behind this additional level of number discrepancy between the CLAIM and FACT is to investigate whether this contextual factor would affect the likelihood of the upper bound inference. All target items (N=30) were rotated through 15 lists, so that each participant only saw one condition per item. Fillers were added too.



Data from 45 participants (for *up to* and *at most* only) was analysed with mixed-effects ordered probit regression models with random effects for subjects and items. Similarly to Experiment 1, participants were more likely to rate higher CLAIM-FACT compatibility in the ‘over’ or ‘way over’ condition with *up to* than with *at most* ( $\beta = -.778$ ,  $SE = .286$ ,  $p < .05$ ;  $\beta = -.198$ ,  $SE = .440$ ,  $p < .05$ , respectively), though curiously the ratings for *at most* were higher in Experiment 2 than in Experiment 1.

The ratings were lower in the ‘way over’ conditions than in the ‘over condition’ both for *up to* ( $\beta = -0.987$ ,  $SE = .278$ ,  $p < .05$ ) and *at most* ( $\beta = -0.199$ ,  $SE = .285$ ,  $p < .05$ ); that is, the scalar distance between the numerals in the CLAIM and FACT was a significant factor.

Taken together, the results from the two experiments show that upper-bound construals are more likely in *at most* and *fewer than* than they are for *up to*, suggesting that this difference is due to the difference in how the upper bound is derived: in the *at most* and *fewer than* it is derived from the lexical semantics, whereas in *up to* it is derived as a pragmatic inference. In the second experiment we additionally show that the upper-bound implicature is sensitive to an additional contextual factor, namely the scalar distance between possible alternatives and the number modified and asserted. This ties in with previous theoretical and experimental studies that show that the distance of alternatives on an entailment-based scale affects the likelihood of an upper-bound construal (Horn 1972, Beltrama & Xiang 2012, van Tiel et al. 2015).

REFERENCES: Geurts, B. & R. Nouwen. 2007. At least et al. Beltrama, A. & M. Xiang. 2012. Is good better than excellent? Blok, D. 2015. The semantics and pragmatics of directional numeral modifiers. Horn, L. 1972. On the semantic properties of logical operators in English. Nouwen, R. 2010. Two kinds of modified numerals. van Tiel, B., et al. 2015. Scalar Diversity.