The extent of upper-bound construals among different modified numerals

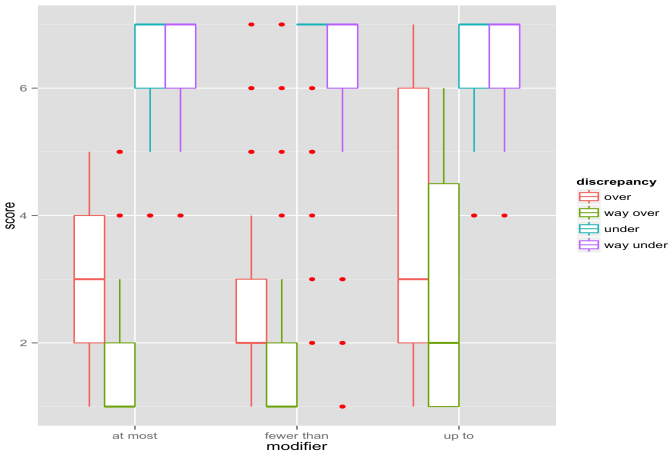
Numeral modifiers have proven to be a gainful case study of central questions in semantic and pragmatic theory, such as the calculation of ignorance inferences and scalar implicature (Geurts & Nouwen, Nouwen 2010, *inter alia*). In this paper we show that modified numerals differ with respect to the nature of the lower or upper bound they express. We investigate to what extent the modifiers *fewer than*, *at most* and *up to* impose an upper bound. We do so in two experiments that differ in the language under study (English and Greek) and the experimental task (evaluating the semantic-pragmatic compatibility of two independent statements vs. evaluating a the discourse coherence of two utterances). 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. 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* and asserts no lower bound, *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 Experiment 1, 45 native speakers of English were asked to rate to what extent a CLAIM (1-a) was compatible with a subsequently provided FACT (1-b) on a –3 to +3 Likert scale.

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

Conditions differed with respect to the choice of modifier (*less/fewer than, at most, up to*), and the discrepancy between the number in the claim (n_{claim}) and the number in the fact (m_{fact}). There were two 'under' (i.e., $m_{fact} < n_{claim}$) and 'over' (i.e., $m_{fact} > n_{claim}$) 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' ($m_{fact} = n_{claim} * 1.75$) (1-b). The idea behind the two levels of each number discrepancy between the CLAIM and FACT was 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.



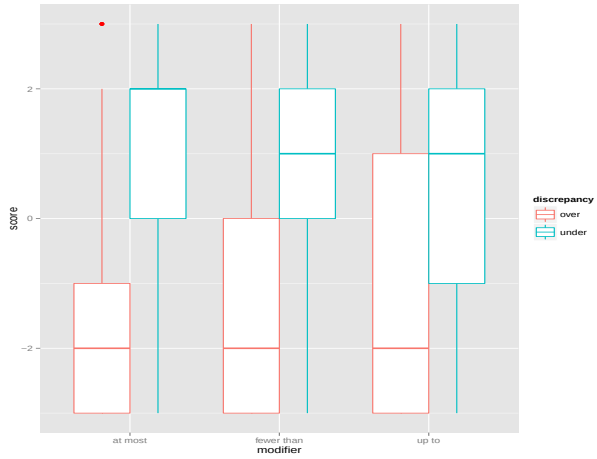
Both Schwarz et al. and Blok assume in their analyses that *at most* provides an upper bound entailment. This predicts that subjects are more likely to approve of an 'over' item when the modifier is *up to* than when it is *at most*, and this is indeed what we found. The boxplot above summarises data from 45 participants, analysed with mixed-effects ordered probit regression models

with random effects for subjects and items. Participants were more likely to rate higher CLAIM-FACT compatibility in an 'over' or 'way over' condition when the modifier was *up to* than when the modifier was *at most* or *fewer than* (β=.783, *SE*=.415, $p < .05$; β=-.937, *SE*=.413,$p < .05$, respectively). The ratings were lower in the 'way over' conditions than in the 'over' both for *up to* (β=-.987, *SE*=.278, $p < .05$) and *at most* (β=-.199, *SE*=.285, $p < .05$); that is, the scalar distance between the numerals in the CLAIM and FACT was a significant factor.

Experiment 2. We examined the difference between the Greek equivalents of *fewer than*/*at most* and *up to*, but with a different task that tested discourse coherence of pairs of sentences rather than two independent statements, as in Experiment 1. In this experiment, subjects were asked to rate the naturalness of pairs of sentences occurring together on a –3 to +3 Likert scale. Conditions differed with respect to choice of modifier (*lighotero apo* 'less than'/ *to poli* 'at most'/ *mehri* 'up to') and with respect to whether m < n ('under' condition) or m > n ('over' condition), as in (2).

(2) Every student read {*less than / at most / up to*} *n* papers. One of them read *m* papers.

All target items (N=12) were rotated through 6 lists, so that each participant only saw one condition per item. This boxplot summarises the data by 143 participants, analysed with mixed-effects ordered probit regression models with random effects for subjects and items. The results of Exp. 1 are replicated in Experiment 2: Participants were more likely to give higher coherence rates to statements in 'over' condition when the modifier was *up to* than when it was *fewer than* (β = -.188, *SE* = .089, $p < .05$) or *at most* (β = -.277, *SE* = .09, $p < .01$). In the 'under condition', *up to* items received significantly lower scores than *at most* and *fewer than* items (β = .266, *SE* .088, $p < .001$, and β = .215, *SE* = .088, $p < .05$, respectively), a difference we didn't find in Experiment 1. This difference can be explained in terms of directivity (Nouwen 2010b, Blok 2015): whereas *up to* points to the elements of which the sentence holds, *at most* and *fewer than* point to the elements of which the sentence doesn't. In discourse, a speaker using a numeral modifier with upward directivity like *up to* and then proceeding to highlight what a subset did using a smaller value would come across as uninformative, in contrast with *at most* and *fewer than*.



Consequences. 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. On the natural assumption that the upper bound provided by *fewer than* is entailed, those results can be interpreted to indicate that the upper bound for *up to* is derived via pragmatics, whereas in *at most* it is derived from the lexical semantics. In Experiment 1 we also show that the upper-bound implicature is sensitive to additional contextual factors, 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 that the stronger alternatives on that scale are negated; i.e., the likelihood of upper-bound construal (Horn 1972, Beltrama & Xiang 2012, van Tiel et al. 2015). REFERENCES: **Geurts, B. & R. Nouwen**. 2007. At least et al. **Blok, D.** 2015. Directional numeral modifiers. **Nouwen, R.** 2010a. Two kinds of modified numerals.