

Pragmatic inferences with numeral modifiers: novel experimental data

We present results from two experiments that investigate the nature of pragmatic inferences triggered by numeral modifiers (namely *at least* and *more than* when embedded under universal quantifiers). We have found that while both types of modifiers give rise to implicatures, they differ in the extent to which they trigger such inferences, contra current theoretical analyses (e.g. Büring 2008, Schwarz 2013). We also found that the precise nature of the inferences in question is one that favours a theory positing a rather modest set of alternatives for modified numerals (as in Büring 2008, contra Schwarz 2013).

Background — Geurts and Nouwen (2007) argued that superlative modified numerals (SMNs, e.g. "at least 3") differ from comparative ones (CMNs, e.g. "more than 3") in giving rise to epistemic inferences, which may disappear in certain embedded contexts (cf. Buring 2007). While it is (relatively) uncontroversial that these two kinds of quantifiers differ with respect to the ignorance inferences they trigger, it is an open question how this difference extends to other pragmatic inferences. Two inferences are of relevance here: variability effects akin to free choice and scalar implicatures. Consider the sentence *Every student read at least 3 / more than 2 books*. Assuming the scale of numbers as the relevant scale, such a sentence would give rise to implicatures of the kind *not every student read more than 3 books*, suggesting in turn that at least some student read 3 books. Variability effects are (only slightly) different: they say that not every student read the same number of books.

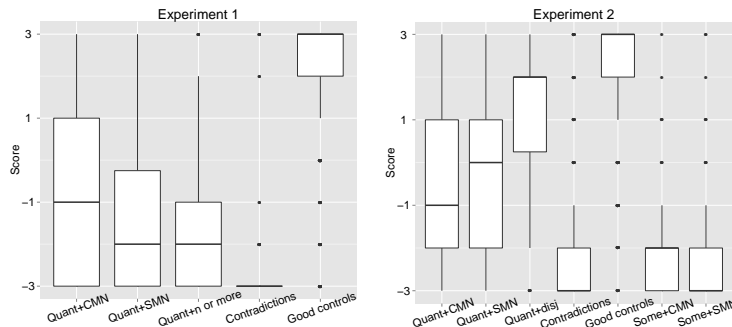
Experiments — In 2 experiments, we investigated both kinds of inferences, whilst also comparing CMNs to SMNs. Both experiments consisted of short dialogues in Dutch between a researcher and an interviewer; the researcher makes a claim and the interviewer poses a question about the claim. Participants had to read these dialogues and rate how well the interviewer has understood the researcher's claim on a Likert scale from -3 (*the claim is not understood*) to 3 (*the claim is understood*).

Both experiments used the same target items, varying only in the interviewer's question. The following box gives a translation of such an item.

Researcher: During the event *every* street was guarded by *more than 6/at least 7* policemen.
Exp1. Interviewer: Were they all guarded by the same number of policemen?
Exp2. Interviewer: How did you find out that there were streets guarded by 7 policemen?

The interviewer's question in Exp1 targets a variability inference connected to the researcher's claim (i.e., not every street has the same number of guards) and is inconsistent with this inference. The question in Exp2 targets and requires the inference triggered by scalar implicature (i.e., not every street was guarded by more than 7/at least 8 policemen). Aside from the modified numerals, Exp1 also tested the variability effects in the *n or more* disjunctions and Exp2 tested scalar implicatures under three determiners (*every, all, some*) and the implicature of disjunction under *all*. (The items with *every* and *all* showed almost identical behavior, so they are merged from now on under the heading *Quant*.) All experimental items (6 in Exp1, 14 in Exp2) were rotated through lists in a standard way, so that each participant only saw one condition per item. Contradictory controls (6 in Exp1, 10 in Exp2) as well as semantically and pragmatically well-formed items (13 in Exp1, 10 in Exp2), and fillers were added to every list. There were 68 participants in Exp1, and 43 in Exp2. The data, summarized in the boxplots below, were analysed with mixed-effects ordered probit regression models, including the maximal converging random effect structure for subjects and items. In both experiments Quant+CMNs (and SMNs) were rated significantly lower than the semantically and pragmatically well-formed items ($p < .0001$) and significantly higher than the self-

contradictory items ($p < .05$ in Exp1, $p < .0001$ in Exp2). However, there was also a significant difference between the two kinds of MNs, with Quant+CMN items scoring higher than Quant+SMN in Exp1 ($p = .023$) and lower in Exp2 ($p = .019$). Finally, in Exp2, MNs with *some* received a significantly lower rating than Quant+MNs ($p < .0001$).



The difference found between the Quant+MNs and the well-formed controls in Exp1 reveals the availability of variability effects. In turn, the fact that MNs are considered as better than contradictions shows that those variability effects are not semantically encoded (since that would make the MNs contradictory), rather, they are pragmatic inferences. However, the significant difference between the two kinds of MNs indicates that the inference with CMNs is pragmatically weaker than that with SMNs. Exp2 shows that scalar implicatures are more available in the context of *every* and *all* than in the context of *some*. The fact that Quant+MNs are judged as worse than good controls is evidence that the implicature is by no means obligatory. Importantly, Exp2 also shows that the scalar implicature of Quant+CMNs is less acceptable than the implicature of Quant+SMNs. That is, Exp2 replicates the result of Exp1 that an inference is pragmatically weaker with CMNs than with SMNs. But Exp1 found this for the variability effect, while Exp2 reveals the same for scalar implicatures.

Consequences — The results of this experiment lend support to approaches that draw a parallel between the inferences of modified numerals and the implicatures of disjunctions (Büring 2007, Cummins and Katsos 2010, *inter alia*). Under such accounts, SMNs come with symmetric alternatives: $ALT(\text{at least } 3) = \{\text{more than } 3, \text{ exactly } 3\}$. Symmetric alternatives yield epistemic (i.e. primary) implicatures (Sauerland 2004). In the scope of a universal quantifier, the alternatives are no longer symmetric. Thus for such sentences strong implicatures arise, triggering a variability inference and the scalar implicature of Exp2. This is compatible with the results of our experiments.

Schwarz (2013) notes that there is no motivation for these alternatives. Following Mayr 2012, he explores richer sets of alternatives, e.g. based on having independent scale mates for both modifier and numeral. Such a setup predicts a variability inference for modified numerals in the scope of universal quantifiers, but not the implicature we targeted in Exp2. As such, Exp2 provides evidence in favour of theories positing $ALT(\text{at least } 3) = \{\text{more than } 3, \text{ exactly } 3\}$ over theories positing richer sets of alternatives.

Strikingly, this result is not limited to superlative modified numerals, but extends to CMNs: $ALT(\text{more than } 3) = \{\text{more than } 4, \text{ exactly } 4\}$. That is, just like for SMNs, we found that CMNs has the scalar/variability effect. This is a surprising conclusion, since it would predict that, to some extent at least, CMNs ought to give rise to ignorance implicatures in unembedded environments. Perhaps the second result from our experiments will help here: the inferences triggered by CMNs are generally weaker than those of SMNs. Most of the literature remains silent on the pragmatic effects of CMNs (the exception is Mayr 2012). Our experiments allow us to draw a first conclusion: CMNs and SMNs do not differ with respect to the pragmatic inferences they trigger; they rather differ w.r.t. the strength of such inferences.