Pragmatic inferences with numeral modifiers: Novel experimental data *

Stavroula Alexandropoulou
Utrecht University

Jakub Dotlačil
University of Groningen

Yaron McNabb
Utrecht University

Rick Nouwen
Utrecht University

Abstract We present two experiments that probe so-called variation effects of modified numerals that appear in the scope of a universal quantifier (Geurts & Nouwen 2007; Büring 2008). For example, *Every street was guarded by at least three policemen* suggests that not every street was guarded by the same number of policemen. This kind of variation is similar to inferences observed with epistemic indefinites. We show, however, that indefinites and modified numerals must differ with respect to the underlying mechanism, or, more specifically, with respect to the structure of the set of alternatives that determines the pragmatic inferences. Results from our experiments indicate that the variation effects of modified numerals include the inference of a witness for the lowest number compatible with the modified numeral. We found the same effects for *at least* as for *more than*, but the inferences are weaker for the latter.

Keywords: modified numerals, variation effects, alternatives, implicature

1 Introduction

Modified numerals provide a rich area of research for a great variety of implicature-like inferences, ranging from standard scalar implicatures and their obviation (Fox & Hackl 2006; Mayr 2013; Schwarz 2013) to ignorance implicatures (Geurts & Nouwen 2007; Büring 2008; Nouwen 2010; McNabb & Penka 2015) and free choice effects (Geurts & Nouwen 2007; Büring 2008; Nouwen 2015). The fact that modified numerals give rise to inferences beyond the standard scalar ones has been formulated in a number of analyses that, at some level of description, offer a parallel between numeral modifiers and disjunction (Büring 2008; Cummins & Katsos 2010; Schwarz

---

2013; Coppock & Brochhagen 2013; Nouwen 2015; Kennedy 2015). The leading intuition behind such a comparison comes from pairs like (1)-(2) and (3)-(4). In the same way that a simple unembedded disjunction like (1) gives rise to an epistemic reading, (2) is likewise interpreted as conveying that the speaker does not have knowledge of the exact number of pieces of fruit John ate.

(1)   John ate an apple or a banana.   IGNORANCE

(2)   John ate at least two pieces of fruit.   IGNORANCE

This ignorance effect can disappear in embedded constructions: (3) has a reading that lacks an ignorance implicature, as it can be used even when the speaker knows which fruit each one of the boys ate, as long as some boys ate an apple and some boys ate a banana, and nothing else was eaten by the boys. Similarly, (4) has an interpretation that lacks an ignorance implicature, where the speaker knows precisely how many pieces of fruit each boy ate, as long as no boy ate fewer than two pieces of fruit.

(3)   Every boy ate an apple or a banana.

(4)   Every boy ate at least two pieces of fruit.

Epistemic indefinites show a similar contrast (see e.g., Alonso-Ovalle & Menéndez-Benito 2015): while (5) has an obligatory ignorance effect, this effect is optional in an embedded position (6).

(5)   Irgendein   Student hat angerufen.   (German)
        IRGEND-one student has called
        'Some (non-specific) student called.'

(6)   Jede Studentin hat irgendein   Buch gelesen
        Each student   has IRGEND-one book read
        'Each (female) student read some (non-specific) book.'

That said, the examples (3), (4) and (6) are not free of additional pragmatic inferences. While potentially lacking the ignorance implicature, they give rise to a variation effect. For example, (3) is infelicitous in a situation in which every boy ate the same kind of fruit; that is, if every boy ate an apple or every boy ate a banana. Similar effects can be observed with modified numerals and epistemic indefinites: unless they are read with an ignorance inference, (4) implies that not every boy ate the same number of fruits and (6) implies that not every female student read the same book.

    In this paper we discuss two experiments that (i) probe the existence of variation effects with modified numerals, and (ii) study what mechanisms can be said to be responsible for such effects. In our assessment, we will try to abstract away from

specific frameworks (classical, Hamblin, inquisitive). Instead, we focus on one particular aspect of the pragmatic mechanisms involved in all these frameworks, namely the particular set of alternatives that is assumed to be responsible for the occurring inferences. The next section reviews the strategies that could potentially be responsible for variation in (3), (4) and (6). Section 3 summarises experiments targeting the variation effects. Section 4 concludes.

2   Three strategies yielding variation

In what follows we assume a very basic recipe for implicature calculation: given an assertion *a* and a set of alternatives to that assertion *A*, *a implicates* the denial of each of the alternatives in *A*. The inference from (3) that not everyone ate an apple and not everyone ate a banana, follows straightforwardly from this mechanism once we adopt the assumption that the individual disjuncts are alternatives to the disjunction itself (shown schematically in (7)).

(7)

Assertion	Alternatives	Implicatures
$\forall x[a_x \vee b_x]$	$\forall x[a_x]$	$\neg\forall x[a_x]$
	$\forall x[b_x]$	$\neg\forall x[b_x]$
	$\forall x[a_x \wedge b_x]$	$\neg\forall x[a_x \wedge b_x]$

It follows that:
$\exists x[a_x]$
$\exists x[b_x]$

The exact choice of the alternatives plays a more crucial role once we turn to more than two disjuncts. We distinguish two strategies:

(8)   Strategy 1: the alternatives of a disjunction are the disjuncts themselves, closed under disjunction and conjunction

(9)   Strategy 2: the alternatives of a disjunction are only the atomic disjunctions themselves

The rich set of alternatives in Strategy 1 corresponds to a Katzir-style theory of alternatives (Katzir 2008). Strategy 2 offers a much poorer set, one in which the implicature mechanism involves comparison to only those alternatives that are fully specific. For ternary disjunction, the two strategies produce different results, as shown in (10) and (11). For ease of exposition (10) only contains a relevant (and sufficient) subset of the full set of potential alternatives.

(10)   Strategy 1   It follows that:

Assertion	Alternatives	Implicatures
$\forall x[a_x \vee b_x \vee c_x]$	$\forall x[a_x]$	$\neg\forall x[a_x]$
	$\forall x[b_x]$	$\neg\forall x[b_x]$
	$\forall x[c_x]$	$\neg\forall x[c_x]$
	$\forall x[a_x \vee b_x]$	$\neg\forall x[a_x \vee b_x]$
	$\forall x[a_x \vee c_x]$	$\neg\forall x[a_x \vee c_x]$
	$\forall x[b_x \vee c_x]$	$\neg\forall x[b_x \vee c_x]$

It follows that:
$\exists x[a_x]$
$\exists x[b_x]$
$\exists x[c_x]$
(total variation)

(11)   Strategy 2   It follows that:

Assertion	Alternatives	Implicatures
$\forall x[a_x \vee b_x \vee c_x]$	$\forall x[a_x]$	$\neg\forall x[a_x]$
	$\forall x[b_x]$	$\neg\forall x[b_x]$
	$\forall x[c_x]$	$\neg\forall x[c_x]$

It follows that:
$\exists x[a_x \vee b_x]$
$\exists x[b_x \vee c_x]$
$\exists x[a_x \vee c_x]$
(partial variation)

The two strategies give rise to a different interpretative effect. While both yield a variation effect, the effect of Strategy 1 is a much stronger one than the effect of Strategy 2. While the former says that every disjunct has to be true for at least one x, the latter only requires this to be the case for at least two of the disjuncts. For ternary disjunction, only Strategy 1 is suitable. A sentence like (12) is infelicitous in case (for instance) no boy ate a plum.

(12)   Every boy ate an apple, a banana or a plum.

Given the logical parallel between indefinites and disjunction, one might want to posit a parallel implicature mechanism for variation effects with (epistemic) indefinites. However, indefinites like the Spanish *algún* seem to require Strategy 2. Consider (13), which illustrates this point using universal modals (from Alonso-Ovalle & Menéndez-Benito 2010: p. 6):

(13)   Juan tiene que estar en alguna habitación de la   casa.
        Juan has   to   be   in alguna room   of the house
        'Juan must be in some (non-specific) room in the house.'

The sentence conveys that there is partial variation: Juan might be in one room or another. The strong variation requirement, that all the rooms are epistemic possibilities, is not part of the interpretation of (13) according to Alonso-Ovalle & Menéndez-Benito (2010).

Like indefinites, modified numerals have an intrinsically disjunctive logical structure, too. In fact, the comparison between superlative modified numerals and disjunction has recently been quite popular (Büring 2008; Cummins & Katsos 2010; Coppock & Brochhagen 2013; Kennedy 2015; Nouwen 2015).

A quantifier like *at least 10* could be seen to convey a disjunctive choice of values: $x = 10 \lor x = 11 \lor x = 12 \lor \ldots$. But as was the case with *algún*, it is evident that modified numerals display partial rather than total variation in universally quantified contexts (cf. Schwarz 2013). When not read with an ignorance implicature, the sentence in (14) conveys that not everyone ate the same number of tacos and it clearly lacks the total variation implication, such that for every number exceeding 9, there is someone who ate that many tacos.

(14)   Everyone ate at least 10 tacos.

We can thus already conclude that Strategy 2 fares better than Strategy 1 in the case of modified numerals:

(16)

Strategy 2			It follows that:

Assertion	Alternatives	Implicatures
$\forall x[f_x \geq 10]$	$\forall x[f_x = 10]$	$\neg\forall x[f_x = 10]$
	$\forall x[f_x = 11]$	$\neg\forall x[f_x = 11]$
	$\forall x[f_x = 12]$	$\neg\forall x[f_x = 12]$
	…etc.	…etc.

It follows that:
$\neg\exists n\forall x[f_x = n]$
(partial variation)

Although the details of implementation are rather different, this is essentially a strategy that would fit naturally in the inquisitive semantics proposal of Coppock & Brochhagen (2013) or the anti-specificity proposal of Nouwen (2015).[1] Importantly, Strategy 2 is not the only route to an account of partial variation. Rather than viewing a quantifier like *at least 10* as an infinite disjunction of values, it is often assumed that *at least 10* is (at some level of description) parallel to the binary disjunction *exactly 10 or more* (Büring 2008; Cummins & Katsos 2010; Kennedy 2015). Let's call this Strategy 3:

(16)

Strategy 3			It follows that:

Assertion	Alternatives	Implicatures
$\forall x[f_x \geq 10]$	$\forall x[f_x = 10]$	$\neg\forall x[f_x = 10]$
	$\forall x[f_x > 10]$	$\neg\forall x[f_x = 10]$

It follows that:
$\exists x[f_x = 10]$
$\exists x[f_x > 10]$
(partial variation)

Strategy 3 shares with Strategy 2 the implication of partial variation, but it is stronger in the sense that it additionally requires there to be a witness for the value 10. So, while with Strategy 3 it follows from (14) that someone ate 10 tacos (and that someone ate more), with Strategy 2 it merely follows from (14) that not everyone ate the same number of tacos.

---

1 More precisely, Nouwen (2015) discusses the potential empirical shortcomings of such a proposal as well as ways to circumvent these.

We now turn to the experiments that target variation effects and can thus potentially distinguish between the two strategies that are commonly considered, Strategy 2 and Strategy 3.

3 Experiments

We discuss two experiments conducted in Dutch that address two separate goals: (i) to what extent do modified numerals display variation effects? (Experiment 1); (ii) what strategy is involved in the variation effects? (Experiment 2). Both experiments involve a felicity judgement task. Participants were presented with claims made by a researcher, followed by a question posed by someone interviewing the researcher. The participants were asked to judge to what extent the question *made sense* given the statement just made.

Here are two illustrations of this task:

(17)   **Researcher:** Some of the samples were contaminated.
       Interviewer: Were all of the samples contaminated?
       Does the interviewer's question make sense?

(18)   **Researcher:** Some of the samples were contaminated.
       Interviewer: How did you find out that not all of the samples were contaminated?
       Does the interviewer's question make sense?

Participants could indicate to which extent they thought the question made sense in light of the claim just made by selecting a score on a Likert scale. The idea was that if a participant assumes that the interviewer will calculate an implicature in the researcher's claim (that not all of the samples were contaminated), then s/he will judge the interviewer's question in (17) as relatively infelicitous, since the participant is assuming that the interviewer already knows the answer to the question s/he is posing. Reversely, in (18), calculating the implicature will lead to a higher score, since the question takes for granted that the implicature has indeed been calculated. In other words, using this method, we can measure the likelihood that a certain inference is drawn. We used the strategy in (17) in Experiment 1 and the one in (18) in Experiment 2.

3.1 Experiment 1

3.1.1 Participants

97 people participated in Experiment 1 out of whom only 68 were native speakers of Dutch who filled in the entire questionnaire (created on www.surveymonkey.com)

without making any mistakes on the practice items. We used only the data from these 68 participants (39 female, mean age: 38, age range: 20-68) in subsequent analyses. All of the participants volunteered in filling in the questionnaire and were naive as to the purpose of the study.

3.1.2 Design & material

The experimental conditions all involved statements uttered by the researcher with a modified numeral embedded in the nuclear scope of a universal quantifier. The interviewer's question asked about more details regarding the researcher's statement in a way that went against the variation effect of the researcher's statement. For instance, one of the items had the form in (19).

(19)   Example of an experimental item (in the original Dutch):

Onderzoeker: Tijdens het evenement werd elke straat door $\begin{Bmatrix} \text{meer dan zes} \\ \text{minstens zes} \\ \text{zes of meer} \end{Bmatrix}$

agenten beveiligd.
Interviewer: Werden ze allemaal door evenveel agenten beveiligd?

English version:

Researcher: During the event every street was guarded by $\begin{Bmatrix} \text{more than six} \\ \text{at least six} \\ \text{six or more} \end{Bmatrix}$

policemen.
Interviewer: Were they all guarded by the same number of policemen?

The use of the researcher-interviewer context served to weaken a potential ignorance interpretation of the researcher's statement. The assumption is that a researcher is most likely to be an authority on her/his own results.

There were two factors in the experiment: (i) the form of the modified numeral: the comparative form (*meer dan n* 'more than *n*'), the superlative form (*minstens n* 'at least *n*') or the disjunctive form (*n of meer* '*n* or more'), and (ii) the type of question the interviewer asked, which had two levels. The latter manipulation is not relevant for the present study, so we will not be discussing it here. Moreover, we will only be concerned with *more than* and *at least*; see Alexandropoulou (2015) for a discussion of the results regarding *n or more*.

The experimental task had six target items and six fillers. It also included 13 good controls and four bad controls (number of stimuli = 29).[2] In the good controls, the

---

2 Given the variation effects, it was expected that experimental items would be judged as bad, so more good controls were present to offset the asymmetry.

interviewer asks a relatively neutral follow-up question, which does not contradict the researcher's statement; in the bad controls, the interviewer's question prompts a proposition that contradicts the given statement. An example of a good and a bad control is given below.

(20)   Example of a good control:
  Onderzoeker: Elke boer in Drenthe heeft zes of meer koeien.
  Interviewer: Hoe ben je daar achtergekomen?

  English version:
  Researcher: Every farmer in Drenthe has six or more cows.
  Interviewer: How did you find that out?

(21)   Example of a bad control:
  Onderzoeker: In de binnenstad van Leiden vind je volgens de Michelin-gids de mooiste grachten.
  Interviewer: Waarom heeft Leiden eigenlijk geen grachten?

  English version:
  Researcher: According to the Michelin guide, the nicest canals are in the centre of Leiden.
  Interviewer: Why are there no canals in Leiden?

Experimental items were rotated through six lists, so that each participant saw each item only in one experimental condition.

3.1.3   Procedure

Participants were presented with items like (19-21) and were asked to judge how well the interviewer understood the researcher's claim. They did so on a –3 to 3 Likert scale, where –3 is 'the claim is not understood' and 3 is 'the claim is understood'. Four practice items were provided at the beginning of the questionnaire to familiarise participants with the task. They had the form of bad and good continuation questions; that is, two of them were to be rated towards the negative end of the scale and the other two towards the positive end of the scale. Most participants filled in the experiment on-line, although some participants completed a pen-and-paper version of the experiment.

3.1.4   Results and discussion

Part of the obtained data from good control items was not included in the statistical analysis due to a typo. The rest of the data were analysed with ordered probit models

using the `ordinal` package (Christensen 2013) in R. Two different analyses were conducted: (i) the reference level was MORE THAN, which was compared to the other numeral modifier conditions as well as to the good and bad controls; (ii) the reference level was AT LEAST, which was compared to the good and bad controls. In both analyses, there was one fixed effect: that of the experimental manipulation (five levels in analysis (i): MORE THAN, AT LEAST, N OR MORE, GOOD CONTROLS, and BAD CONTROLS, and three levels in analysis (ii): AT LEAST, GOOD CONTROLS, and BAD CONTROLS). Both models included intercept and slope random effects for subjects and random intercepts for items, which was the maximal random-effect structure.
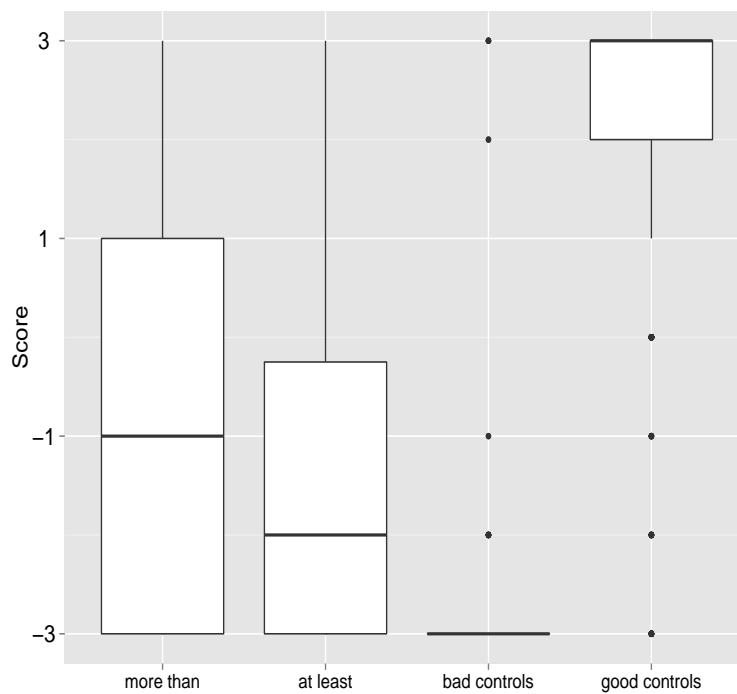
Figure 1   Experiment 1: boxplots of scores

The boxplot in Figure 1 summarises the scores of each experimental manipulation. In the first analysis, the AT LEAST items received significantly lower scores than the MORE THAN items ($\beta = -.521$, $SE = .217$, $p < .05$). Moreover, the MORE THAN items were rated significantly lower than the GOOD CONTROLS ($\beta = -2.049$, $SE = .295$, $p < .0001$) and significantly higher than the BAD CONTROLS ($\beta = 2.434$, $SE = .456$, $p < .0001$). The second analysis revealed a similar behaviour for the

AT LEAST items with respect to the two control conditions. That is, the AT LEAST items scored significantly lower than the GOOD CONTROLS ($\beta = -2.409$, $SE = .276$, $p < .0001$) and significantly higher than the BAD CONTROLS ($\beta = 1.895$, $SE = .416$, $p < .0001$).

Recall that the interviewer's questions in MORE THAN and AT LEAST would form a plausible follow-up to the researcher's statement, if they didn't go against the variation inference of modified numerals. The fact that MORE THAN and AT LEAST were judged as significantly worse than the GOOD CONTROLS supports the position that the modified numerals yield variation effects. At the same time, the stimuli with the modified numerals were judged as better than the BAD CONTROLS, which may suggest that the interviewer's violation of variation effects is not a case of semantic contradiction. Overall, it is tempting to conclude that variation effects are *pragmatic* in nature, given the fact that the scores were significantly different from those of both good and bad controls. We should note, however, that the modified numeral items did not form minimal pairs with the controls and, so, there could be other reasons for the differences we found.

We now turn to the follow-up experiment that further investigates the nature of variation effects with modified numerals.

3.2   Experiment 2

The results of Experiment 1 provide evidence that modified numerals yield variation effects. However, the experiment leaves it open whether such effects are due to Strategy 2 or 3. This issue is addressed in the second experiment.

3.2.1   Participants

38 native speakers of Dutch filled in the on-line experiment (created in Ibex and hosted on Ibex farm, Drummond 2007). All participants were undergraduate students at the Faculty of Arts, University of Groningen (most of them first-year students). They received a course credit for their participation and were naive as to the purpose of the study.

3.2.2   Design & material

The second experiment tested two types of modified numerals, *meer dan n* 'more than *n*' and *minstens n* 'at least *n*'. The modified numerals were tested in six conditions, detailed below.

In four of the six conditions, modified numerals were embedded under a universal quantifier, *alle* 'all'. In the first two of these four, ALL-LOWEST, the researcher made

a statement similar to Experiment 1, claiming that for every situation, more than $n - 1$/at least n entities had a property P. The interviewer's question then presupposed that (the researcher just provided evidence that) in some cases, n entities had the property P, as exemplified in (22).

(22) Example of an experimental item in conditions ALL-LOWEST (Dutch):

Onderzoeker: Tijdens het evenement werden alle straten door $\left\{ \begin{array}{l} \text{meer dan drie} \\ \text{minstens vier} \end{array} \right\}$ agenten beveiligd.
Interviewer: Hoe ben je er achtergekomen dat er een straat met vier agenten is?

English version:

Researcher: During the event all streets were guarded by $\left\{ \begin{array}{l} \text{more than three} \\ \text{at least four} \end{array} \right\}$ policemen.
Interviewer: How did you find out that there was a street guarded by four policemen?

The interviewer's presupposition makes use of the inference triggered by variation effects. Recall that under Strategy 3, the researcher's statement in (22) implies that there was at least one street guarded by four policemen. The inference would license the presupposition. However, if variation effects were derived through Strategy 2, it would only follow from the researcher's statement that not all streets had the same number of policemen. Hence, the presupposition would not be satisfied and some extra reasoning would be required to accommodate the felicity of the interviewer's question.

The second pair of the first four conditions, ALL-HIGHER, illustrated in (22), differed only in the number that the interviewer's question targeted (boldfaced in the example).

(23) Example of an experimental item in conditions ALL-HIGHER (Dutch):

Onderzoeker: Tijdens het evenement werden alle straten door $\left\{ \begin{array}{l} \text{meer dan drie} \\ \text{minstens vier} \end{array} \right\}$ agenten beveiligd.
Interviewer: Hoe ben je er achtergekomen dat er een straat met **vijf** agenten is?

English version:

Researcher: During the event all streets were guarded by $\left\{ \begin{array}{l} \text{more than three} \\ \text{at least four} \end{array} \right\}$ policemen.

Interviewer: How did you find out that there was a street guarded by **five** policemen?

Since neither Strategy 2 nor Strategy 3 derives the inference that there were some streets guarded by five policemen, the presupposition would have to be accommodated under either strategy.

The last two conditions, SOME, tested modified numerals embedded under existential quantifiers, as exemplified in (24).

(24)   Example of an experimental item in conditions SOME (Dutch):

Onderzoeker: Tijdens het evenement werden enkele straten door $\left\{ \begin{array}{l} \text{meer dan drie} \\ \text{minstens vier} \end{array} \right\}$ agenten beveiligd.
Interviewer: Hoe ben je er achtergekomen dat er geen straat met meer dan vier agenten was?

English version:

Researcher: During the event some streets were guarded by $\left\{ \begin{array}{l} \text{more than three} \\ \text{at least four} \end{array} \right\}$ policemen.
Interviewer: How did you find out that there was no street guarded by more than 4 policemen?

In this case, the interviewer's presupposition should be valid if modified numerals triggered scalar implicatures on the scale $\langle$more than n / at least n, more than $n + 1$ / at least $n + 1, \ldots \rangle$.

Finally, the experiment also tested variation effects with disjunction in one condition. Since this is not directly relevant for our study, we will not discuss it here.

The experiment consisted of 14 items (having seven conditions). There were an additional and comparable 14 items (with seven conditions), which tested modified numerals (and disjunctions) under universal and existential modals (not discussed here). The experiment also included 10 good controls similar to Experiment 1, and 10 contradictions. Finally, there were 10 fillers. The total number of stimuli was 58.

3.2.3   Procedure

Participants had to judge items on a -3 to 3 Likert scale, where -3 stands for 'the claim is not understood' and 3 stands for 'the claim is understood'. There were four practice items at the beginning of the questionnaire, two of which were examples of bad questions and two of which were examples of good questions. The items were rotated through seven lists so that each participant saw each item only in one condition.

3.2.4 Results and discussion

As in Experiment 1, the data were analysed with ordered probit models using the `ordinal` package in R. Two different analyses were considered: (i) conditions in which modified numerals were either embedded under universal or existential quantifiers were compared to each other; (ii) the reference level of the first analysis was compared to good and bad controls. In the first analysis, there were two fixed effects and their interaction. The effects were: the type of modified numeral (MORE THAN or AT LEAST, the former being the reference level); the type of embedding + question (ALL-LOWEST – modified numerals under 'all', the question targets the lowest possible number, cf. (22), ALL-HIGHER – modified numerals under 'all', the question targets a higher number, cf. (23), SOME – modified numerals under 'some'; ALL-LOWEST was the reference level). In the second analysis, there was one fixed effect: that of the experimental manipulation comparing comparatives with ALL-LOWEST to controls (ALL-LOWEST + MORE THAN vs. GOOD CONTROLS and BAD CONTROLS). The models included the maximal random effect structure for subjects and items.
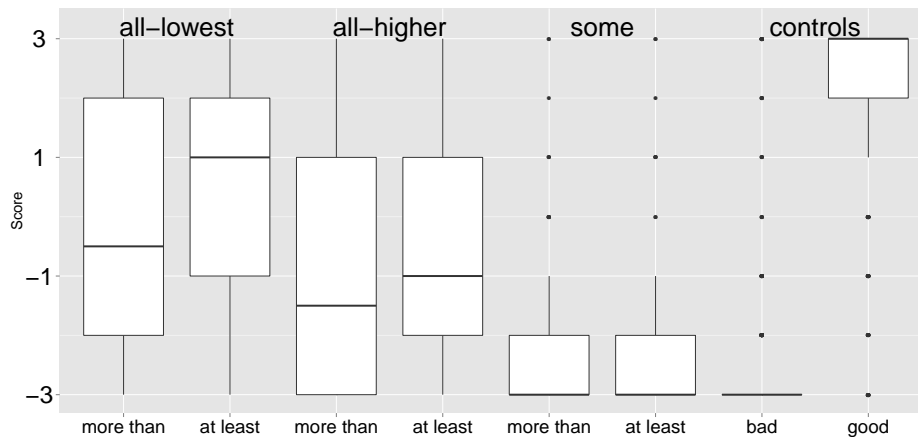

Figure 2   Experiment 2: boxplots of scores

In the first analysis, we found two significant main effects: ALL-HIGHER and SOME were rated as worse than ALL-LOWEST ($\beta = -.603, SE = .234, p = .01$ and $\beta = -1.765, SE = .338, p < .0001$, respectively). In the second analysis, GOOD CONTROLS were judged as better than ALL-LOWEST + MORE THAN, while BAD CONTROLS were judged as worse ($\beta = -1.637, SE = .303, p < .0001$ and $\beta = -2.031, SE = .343, p < .0001$, respectively). Boxplots of our data are provided in

Figure 2.

The significant difference between ALL-HIGHER and ALL-LOWEST in the expected direction shows that the interviewer's question made more sense when it targeted the lowest possible number compatible with the researcher's statement. This result is not accounted for by Strategy 2, which predicts that readers should accommodate questions in ALL-LOWEST and ALL-HIGHER in equal measures and therefore not distinguish between these two conditions. In contrast, Strategy 3 predicts this difference, since the presupposition of the interviewer's question in ALL-LOWEST follows from variation effects and does not need to be accommodated, unlike in ALL-HIGHER. Proponents of Strategy 2 would thus have to argue that the observed difference has an alternative source to the one just mentioned. We are not sure what such an alternative could be. We are not aware of evidence that statements about lower numbers are easier to accommodate than statements about higher numbers. Notice also that one cannot simply state that the interviewer repeats the same number as the reviewer in ALL-LOWEST and this repetition explains the observed finding. This explanation would work for the AT LEAST items, but it would fail to explain our data in the MORE THAN condition, since there the interviewer never repeats the number that the researcher used.

While we believe that our results support Strategy 3, they are problematic for the actual proposals adopting Strategy 3 (Büring 2008; Cummins & Katsos 2010; Kennedy 2015). This is because those proposals assume the variation effects *only* for *at least*. But this is clearly incorrect: we observe the difference ALL-LOWEST vs. ALL-HIGHER irrespective of the type of modified numeral.

The degraded status of SOME compared to ALL-LOWEST strongly suggests that while variation effects are available with modified numerals, scalar implicatures are not. This is in line with previous observations (e.g., Krifka 1999). Finally, as in Experiment 1, we also see that our items with modified numerals are worse than good controls and better than bad controls, suggesting that variation effects are available (hence, ALL-LOWEST is judged as better than bad controls), but they are not as readily available as neutral follow-up questions (good controls). The latter fact might be interpreted as evidence that some extra work (pragmatic inference) is needed for their derivation.

There was another finding in Experiment 1, which we did not mention so far: *minstens* 'at least' was judged as worse than *meer dan* 'more than' when variation was suspended. In the current experiment, this should translate into AT LEAST being judged as *better* than MORE THAN in the ALL-LOWEST condition. We observe this tendency, as can be seen in the boxplot of Figure 2. However, the tendency does not reach significance. One possibility for this discrepancy between the current and previous experiments might be due to lack of power: we only had 38 subjects in Experiment 2, compared to 68 subjects in Experiment 1. But if the numerical

tendency is to be trusted, we see again that variation effects are more strongly tied to modified numerals of the 'at least' type than to the 'more than' type. The varying strength of inference is presumably to blame for why previous literature considered only superlative modified numerals like 'at least' when discussing variation effects.

4   Conclusions

As anticipated in a great deal of the theoretical literature on modified numerals, our experiments show that *at least* triggers variation effects. A more surprising result from our experiment (but one anticipated in Mayr 2013) is that comparative modified numerals (*more than n*) display the same effects, albeit in a somewhat weaker form. The variation effects were significantly different from both good and bad controls, which may suggest that they are not semantic in nature. It is tempting to conclude from these studies that these inferences are *implicatures*, a view that is dominant in the theoretical literature. However, since the experiments did not allow us to directly compare semantic responses to the candidate pragmatic ones, we will rather stick with the conclusion that the current results are fully consistent with this view.

Experiment 2 clearly distinguishes between various implementations of the pragmatic mechanism behind variation effects. We successfully demonstrated that in sentences of the form *all streets were guarded by { at least n / more than n − 1 } policemen*, there is the specific inference that some street was guarded by exactly *n* policemen. As we explained above, the existence of this inference is evidence in favour of what we called Strategy 3 and evidence against Strategy 2.

In our description of Strategy 3, we took the simple view that modified numerals are *binary* disjunctions. This is intuitively appealing for *at least n* which then takes *exactly n* and *more than n* as its alternatives. In other words, the alternatives appear to be generated from the modifier, rather than from the numeral. However, this route is far less appealing for *more than n*. As we demonstrated, such modified numerals trigger the same inferences and so, according to the same strategy, we would need to say that *more than n* has the alternatives *exactly n + 1* and *more than n + 1*. Hence, the idea that modified numerals generate alternatives by substituting the modifier only is untenable. That said, the way we described Strategy 3 above is not the only way to get this particular set of inferences (partial variation and a witness for the lowest number). Mayr (2013) (see also Schwarz 2013) suggests that the alternatives for modified numerals are generated by substituting both the numeral and the modifier, where the only alternative for *at least* is *at most* and the only alternative for *more than* is *fewer than*. For example, *at least 3* has as its alternatives: *at least 4*, *at least 5*, *at least 6*, ..., *at most 3*, *at most 4*, ..., etc. Using this set of alternatives in the implicature mechanism is equivalent to using the simple set {*exactly 3*, *more than 3*}.

In sum, we have presented evidence that modified numerals come with partial variation inferences akin to those found with epistemic indefinites. These variation effects, however, differ from those found with epistemic indefinites in that they include the inference of a witness for the lowest number compatible with the modified numeral.

References

Alexandropoulou, Stavroula. 2015. Testing the nature of variation effects with modified numerals. In Eva Csipak & Hedde Zeijlstra (eds.), *Sinn und Bedeutung 19*, 36–53.

Alonso-Ovalle, Luis & Paula Menéndez-Benito. 2010. Modal indefinites. *Natural Language Semantics* 18. 1–31.

Alonso-Ovalle, Luis & Paula Menéndez-Benito (eds.). 2015. *Epistemic Indefinites: Exploring Modality Beyond the Verbal Domain*. Oxford: Oxford University Press.

Büring, Daniel. 2008. The least *at least* can do. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of WCCFL 26*, 114–120. Somerville, Massachusetts: Cascadilla Press.

Christensen, R. H. B. 2013. ordinal—regression models for ordinal data. R package version 2013.9-30. http://cran.r-project.org/web/packages/ordinal/index.html.

Coppock, Elizabeth & Thomas Brochhagen. 2013. Raising and resolving issues with scalar modifiers. *Semantics & Pragmatics* 6(3). 1–57.

Cummins, Chris & Napoleon Katsos. 2010. Comparative and superlative quantifiers: Pragmatic effects of comparison type. *Journal of Semantics* 27(3). 271–305.

Drummond, Alex. 2007. Ibex, experiment platform. http://spellout.net/ibexfarm/.

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586.

Geurts, Bart & Rick Nouwen. 2007. At least et al.: the semantics of scalar modifiers. *Language* 83(3). 533–559.

Katzir, Roni. 2008. *Structural Competition in Grammar*. Cambridge: MIT dissertation.

Kennedy, Christopher. 2015. A "de-fregean" semantics (and neo-gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics* 8(10). 1–44. doi:10.3765/sp.8.10.

Krifka, Manfred. 1999. At least some determiners aren't determiners. In K. Turner (ed.), *The semantics/pragmatics interface from different points of view*, vol. 1, 257–291. Elsevier.

Mayr, Clemens. 2013. Implicatures of modified numerals. In *From grammar to meaning: The spontaneous logicality of language*, 139–171.

McNabb, Yaron & Doris Penka. 2015. An experimental investigation of ignorance inferences and authoritative interpretations of superlative modifiers. Under review.

Nouwen, Rick. 2010. Two kinds of modified numerals. *Semantics and Pragmatics* 3(3). 1–41. doi:http://dx.doi.org/10.3765/sp.3.3.

Nouwen, Rick. 2015. Modified numerals: the epistemic effect. In Luis Alonso-Ovalle & Paula Benito-Ménendez (eds.), *Epistemic Indefinites*, Oxford: Oxford University Press.

Schwarz, Bernhard. 2013. At least and quantity implicature: Choices and consequences. In Maria Aloni, Michael Franke & Floris Roelofsen (eds.), *Proceedings of the 19th Amsterdam Colloquium*, Universiteit van Amsterdam.

Stavroula Alexandropoulou
Trans 10
NL-3512JK Utrecht
s.alexandropoulou@uu.nl

Yaron McNabb
Trans 10
NL-3512JK Utrecht
y.mcnabb@uu.nl

Jakub Dotlačil
Oude Kijk in 't Jatstraat 26
NL-9712EK Groningen
j.dotlacil@gmail.com

Rick Nouwen
Trans 10
NL-3512JK Utrecht
r.w.f.nouwen@uu.nl